CIRJE-F-933

Tests for Covariance Matrices in High Dimension with Less Sample Size

Muni S. Srivastava
University of Toronto

Hirokazu Yanagihara
Hiroshima University

Tatsuya Kubokawa
The University of Tokyo

June 2014

Tests for Covariance Matrices in High Dimension with Less Sample Size

Muni S. Srivastava

Department of Statistics, University of Toronto, 100 St George Street, Toronto, Ontario, CANADA M5S 3G3.

Hirokazu Yanagihara

Department of Mathematics, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8626, JAPAN.

Tatsuya Kubokawa

Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN.

Abstract

In this article, we propose tests for covariance matrices of high dimension with fewer observations than the dimension for a general class of distributions with positive definite covariance matrices. In one-sample case, tests are proposed for sphericity and for testing the hypothesis that the covariance matrix $\mathbf{\Sigma}$ is an identity matrix, by providing an unbiased estimator of $\mathrm{tr}\,[\mathbf{\Sigma}^2]$ under the general model which requires no more computing time than the one available in the literature for normal model. In the two-sample case, tests for the equality of two covariance matrices are given. The asymptotic distributions of proposed tests in one-sample case are derived under the assumption that the sample size $N = O(p^\delta)$, $1/2 < \delta < 1$, where p is the dimension of the random vector, and $O(p^\delta)$ means that N/p goes to zero as N and p go to infinity. Similar assumptions are made in the two-sample case.
AMS 2001 subject classifications: 62H15, Secondary 62F05.

Keywords: Asymptotic distributions, covariance matrix, high dimension, non-normal model, sample size smaller than dimension, test statistics.

AMS 2000 subject classifications: Primary 62H15; secondary 62F05.

1. Introduction

In analyzing data, certain assumptions made implicitly or explicitly should be ascertained. For example in comparing the performances of two groups based on observations from both groups, it is necessary to ascertain if the two groups have the same variability. For example, if they have the same variability, we can use the usual *t*-statistics to verify that both groups have the same average performance. And if the variability is not the same, we are required to use Behrens-Fisher type of statistics. When observations are taken on several characteristics of an individual, we write them as observation vectors. In this case, we are required to check if the covariance matrices of the two groups are the same by using Wilks (1946) likelihood ratio test statistics provided the number of characteristics, say, p is much smaller than the number of observations for each group, say, N_1 and N_2. In this article, we consider the case when p is larger than N_1 and N_2.

The problems of large p and very small sample size are frequently encountered in statistical data analysis these days. For example, recent advances in technology to obtain DNA microarrays have made it possible to measure quantitatively the expression of thousands of genes. These observations are, however, correlated to each other as the genes are from the same subject. Since the number of subjects available for taking the observations are so few as compared to the gene expressions, multivariate theory for large p and small sample size N needs to be applied in the analysis of such data. Alternatively, one may try to reduce the dimension by

Email addresses: srivasta@utstat.toronto.edu (Muni S. Srivastava), yanagi@math.sci.hiroshima-u.ac.jp (Hirokazu Yanagihara), tatsuya@e.u-tokyo.ac.jp (Tatsuya Kubokawa)

using the false discovery rate (FDR) proposed by Benjamini and Hochberg (1995) provided the observations are equally positively related as shown by Benjamini and Yekutieli (2001) or apply the average false discovery rate (AFD) proposed by Srivastava (2010). The above AFD or FDR methods do not guarantee that the dimension p can be reduced to a dimension which is substantially smaller than N.

The development of statistical theory for analyzing high-dimensional data has taken a jump start since the publication of a two-sample test by Bai and Saranadasa (1996) which has also included the two-sample test proposed by Dempster (1958, 1960). A substantial progress has been made in providing powerful tests in testing that the mean vectors are equal in two or several samples, see Srivastava and Du (2008), Srivastava (2009), Srivastava, Katayama and Kano (2013), Yamada and Srivastava (2012) and Srivastava and Kubokawa (2013). In the context of inference on means of high-dimensional distributions, multiple tests have also been used, see Fan, Hall and Yao (2007) and Kosorok and Ma (2007) among others. All the methods of inference on means mentioned above require some verification of the structure of a covariance matrix in one-sample case and the verification of the equality of two covariance matrices in the two-sample case. The objective of the present article is to present some methods of verification of these hypotheses. Below, we describe these problems in terms of hypotheses testing.

Consider the problem of testing the hypotheses regarding the covariance matrix $\boldsymbol{\Sigma}$ of a p-dimensional observation vector based on N independent and identically distributed (i.i.d) observation vectors $\boldsymbol{x}_j$, $j = 1, \ldots, N$. In particular, we consider the problem of testing the hypothesis that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$, $\lambda > 0$, and unknown and, that of testing that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$; the first hypothesis is called sphericity hypothesis. We also consider the problem of testing the equality of the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ of the two groups when N_1 i.i.d observation vectors are obtained from the first group and N_2 i.i.d observation vectors are obtained from the second group. It will be assumed that $N_1 \leq N_2$, $0 < N_1/N_2 \leq 1$ and $N_i/p \to 0$ as $(N_1, N_2, p) \to \infty$.

We begin with the description of the model for the one-sample case. Let $\boldsymbol{x}_j$, $j = 1, \ldots, N$ be i.i.d observation vectors with mean vector $\boldsymbol{\mu}$, and covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{F}\boldsymbol{F}$, where $\boldsymbol{F}$ is the unique factorization of $\boldsymbol{\Sigma}$, that is, $\boldsymbol{F}$ is a symmetric and positive definite matrix obtained as $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{D}_\lambda\boldsymbol{\Gamma}' = \boldsymbol{\Gamma}\boldsymbol{D}_\lambda^{1/2}\boldsymbol{\Gamma}'\boldsymbol{\Gamma}\boldsymbol{D}_\lambda^{1/2}\boldsymbol{\Gamma}' = \boldsymbol{F}\boldsymbol{F}$, where $\boldsymbol{D}_\lambda = \text{diag}(\lambda_1, \ldots, \lambda_p)$ and $\boldsymbol{\Gamma}\boldsymbol{\Gamma}' = \boldsymbol{I}_p$. We assume that the observation vectors $\boldsymbol{x}_j$ are given by,

$$\boldsymbol{x}_j = \boldsymbol{\mu} + \boldsymbol{F}\boldsymbol{u}_j,\ j = 1, \ldots, N, \tag{1.1}$$

with

$$E(\boldsymbol{u}_j) = \mathbf{0}, \ \ \mathbf{Cov}\,(\boldsymbol{u}_j) = \boldsymbol{I}_p, \tag{1.2}$$

and for integers, $\gamma_1, \ldots, \gamma_p$, $0 \leq \sum_{k=1}^p \gamma_k \leq 8$, $j = 1, \ldots, N$,

$$E\left[\prod_{k=1}^p u_{jk}^{\gamma_k}\right] = \prod_{k=1}^p E(u_{jk}^{\gamma_k}), \tag{1.3}$$

where u_{jk} is the k^{th} component of the vector $\boldsymbol{u}_j = (u_{j1}, .., u_{jk}, .., u_{jp})'$. It may be noted that the condition (1.3) implies the existence of the moments of u_{jk}, $k = 1, \ldots, p$, upto the order eight. For comparison with the normal distribution, we shall write the fourth moment of u_{jk}, namely, $E[u_{jk}^4] = K_4 + 3$. For normal distribution, $K_4 = 0$. In the general case, $K_4 \geq -2$. We may also note that instead of $\boldsymbol{\Sigma} = \boldsymbol{F}^2$, we may also consider as in Srivastava (2009) $\boldsymbol{\Sigma} = \boldsymbol{C}\boldsymbol{C}'$, where $\boldsymbol{C}$ is a $p \times p$ non-singular matrix but it increases the algebraic manipulations with no apparent gain in showing that the proposed tests can be used in non-normal situations.

We are interested in the following testing of hypothesis problems in one-sample case:

Problem (1) $H_1 : \boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p,\ \lambda > 0$ vs $A_1 : \boldsymbol{\Sigma} \neq \lambda \boldsymbol{I}_p$.
Problem (2) $H_2 : \boldsymbol{\Sigma} = \boldsymbol{I}_p$ vs $A_2 : \boldsymbol{\Sigma} \neq \boldsymbol{I}_p$.

These problems have been considered many times in the statistical literature. More recently, Onatski, Moreire and Hallin (2013) and Cai and Ma (2013) have proposed tests for testing the above problems under the assumption that the observation vectors are normally distributed. It has, however, been shown by Srivastava, Kollo and von Rosen (2011) that many of these tests are not robust against the departure from normality. The objective of this paper is to propose tests for the above two problems under the assumptions (1.1)-(1.3) which includes multivariate normal distributions as well as many others.

Onatski *et al*. (2013) test is based on the largest eigenvalue of the sample covariance matrix $\boldsymbol{S} = N^{-1}\sum_{i=1}^{N}\boldsymbol{x}_i\boldsymbol{x}_i'$, where $\boldsymbol{x}_i$ are i.i.d. from $\mathcal{N}_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ for testing $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_p$, against the alternative that $\boldsymbol{\Sigma} = \sigma^2(\boldsymbol{I}_p + \theta\boldsymbol{v}\boldsymbol{v}')$, $\boldsymbol{v}'\boldsymbol{v} = 1$. However, Berthet and Rigollet (2013) argue that the largest eigenvalue cannot discriminate between the null hypotheses and the alternative hypotheses, since $\lambda_{max}(\boldsymbol{S}) \to \infty$ as $p/N \to \infty$, and hence its fluctuations are too large and thus would require much larger θ to be able to discriminate between the two hypotheses; see also Baik and Silverstein (2006).

Cai and Ma (2013) proposed a test based on U-statistics for testing the hypothesis that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ based on N i.i.d. observations from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. For normal distribution, the assumption of the mean vector of the observations to be $\boldsymbol{0}$ makes no difference in the use of the proposed U-statistics as the observation matrix can be transformed by a known orthogonal matrix of Helmerts type to obtain $n = N - 1$ observable i.i.d. observation vectors with mean $\boldsymbol{0}$ and the same covariance matrix $\boldsymbol{\Sigma}$. But for non-normal distributions with mean vector ($\neq \boldsymbol{0}$), this U-statistics cannot be used for testing the above hypothesis. Thus, the U-statistics used by Chen, Zhang and Zhang (2010) is needed to test the above hypothesis which requires computing time of the order $O(N^4)$. Our proposed test requires computing time of the order $O(N^2)$.

In the case of two sample case we have N_1 and N_2 independently distributed p-dimensional observation vectors $\boldsymbol{x}_{ij}$, $j = 1, \ldots, N_i$, $i = 1, 2$; $N_i < p$, $N_1 \leq N_2$, $0 < N_1/N_2 \leq 1$, and $N_i/p \to 0$ as $(N_1, N_2, p) \to \infty$, with mean vectors $\boldsymbol{\mu}_i$ and covariance matrices $\boldsymbol{\Sigma}_i = \boldsymbol{F}_i^2$, $i = 1, 2$, each satisfying the conditions of the model described in (1.1)-(1.3) with $\boldsymbol{\mu}_i$ and $\boldsymbol{F}_i$ in place of $\boldsymbol{\mu}$ and $\boldsymbol{F}$ and $\boldsymbol{u}_{ij} = (u_{ij1}, .., u_{ijp})'$ in place of $\boldsymbol{u}_j$. We consider tests for testing the hypothesis H_3 vs A_3 described in the Problem 3 below:

Problem (3) $H_3: \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ vs $A_3: \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$.

Problem (3) has recently been considered by Cai, Liu and Xia (2013). Following Jiang (2004), they proposed a test against sparse alternative rather than the general alternative given above. In this article, we propose a test on the lines of Schott (2007) using the estimator of the squared Frobenius norm of $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$, under the assumptions given in (1.1)-(1.3) as has been done in Li and Chen (2012) using U-statistics. However, the computing time for the Li and Chen statistics is of the order $O(N^4)$ which for the proposed test, it is only $O(N^2)$.

For testing the hypotheses in Problems (1)-(2), in one-sample case, we make the following assumptions:

Assumption (A)

(i) $N = O(p^\delta)$, $1/2 < \delta < 1$.
(ii) $0 < a_2 < \infty$, $a_4/p = o(1)$, where $a_i = \text{tr}[\boldsymbol{\Sigma}^i]/p$, $i = 1, 2, 3, 4$.
(iii) For $\boldsymbol{\Sigma} = (\sigma_{ij})$, $p^{-2}\sum_{i,j}^{p}\sigma_{ij}^4 = o(1)$.

For testing the hypothesis given in Problem (3) for the two-sample case, the Assumption (A) applies to both the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, and the sample sizes are comparable as stated below:

Assumption (B)

(i) Assumption (A) to both the covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ with $a_{ij} = \text{tr}[\boldsymbol{\Sigma}_j^i]/p$, and $\boldsymbol{\Sigma}_j = (\sigma_{jk\ell})$, $i = 1, 2, 3, 4$, $j = 1, 2$.
(ii) For $N_1 \leq N_2$, $0 < N_1/N_2 \leq 1$.

The organization of the paper is as follows. In Section 2, we give notations and preliminaries for one-sample testing problems. In section 3, we propose tests and give their asymptotic distributions based on the asymptotic theory given in Section 6. The problem of testing the equality of two covariance matrices will be considered in Section 4. Simulation results showing power and attained significance level, the so-called ASL will be given in Section 5. Section 6 gives the general asymptotic theory under which the proposed statistics are shown to be normal. In Section 7, we give results on moments of quadratic forms for a general class of distributions. The paper concludes in Section 8.

2. Notations and Preliminaries in One-Sample Case

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be independently and identically distributed p-dimension observation vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{F}^2$ satisfying the conditions of the model (1.1)-(1.3). Let

$$\overline{\boldsymbol{x}} = \frac{1}{N}\sum_{j=1}^{N}\boldsymbol{x}_j,\ \boldsymbol{V} = \sum_{j=1}^{N}\boldsymbol{y}_j\boldsymbol{y}_j',\ \boldsymbol{y}_j = \boldsymbol{x}_j - \overline{\boldsymbol{x}}, \tag{2.1}$$

$j = 1, \ldots, N$. It is well known that $n^{-1}\boldsymbol{V}$, $n = N-1$, is an unbiased estimator of the covariance matrix $\boldsymbol{\Sigma}$ for any distribution. Since our focus in this paper is on testing the hypotheses on a covariance matrix or equality of two covariance matrices, the matrix $\boldsymbol{V}$ plays an important role. In this paper, we consider tests based on the estimator of the squared Frobenius norm, as a distance between the hypothesis $H : \boldsymbol{\Sigma} = \boldsymbol{I}_p$ against the alternative $A : \boldsymbol{\Sigma} \neq \boldsymbol{I}_p$, the squared Frobenius norm (divided by p) is given by $p^{-1}\mathrm{tr}\,[(\boldsymbol{\Sigma}-\boldsymbol{I}_p)^2] = p^{-1}\mathrm{tr}\,[\boldsymbol{\Sigma}^2] - 2p^{-1}\mathrm{tr}\,[\boldsymbol{\Sigma}] + 1$. Thus, for notational convenience, we introduce the notation

$$a_i = \frac{1}{p}\mathrm{tr}\,[\boldsymbol{\Sigma}^i],\ i = 1, \ldots, 8. \tag{2.2}$$

We estimate a_1 and a_2 by

$$\hat{a}_1 = \frac{1}{np}\mathrm{tr}\,[\boldsymbol{V}],\ n = N-1, \tag{2.3}$$

and

$$\hat{a}_{2s} = \frac{1}{(n-1)(n+2)p}\left\{\mathrm{tr}\,[\boldsymbol{V}^2] - \frac{1}{n}(\mathrm{tr}\,[\boldsymbol{V}])^2\right\}, \tag{2.4}$$

respectively. Srivastava (2005) has shown that $\hat{a}_1$ and $\hat{a}_{2s}$ are unbiased and consistent estimators of a_1 and a_2 under the assumption of normality and Assumption (A). That is,

$$E(\hat{a}_{2s}) = a_2, \qquad Var(\hat{a}_{2s}/a_2) = \frac{4}{n^2} + o(n^{-2}),$$
$$E(\hat{a}_1) = a_1, \qquad \frac{1}{a_2}Var(\hat{a}_1) = \frac{2}{np}.$$

However, for the model (1.1)-(1.3) and $\boldsymbol{\Sigma} = (\sigma_{ij})$,

$$E(\hat{a}_{2s}) = \frac{n}{N(N+1)p}K_4\sum_{i=1}^{p}\sigma_{ii}^2 + a_2,$$

as shown in Section 2.1. Hence,

$$\frac{n}{2}E\,[\hat{a}_{2s} - a_2] = O(p^{-1}\textstyle\sum_{i=1}^{p}\sigma_{ii}^2),$$

which does not go to zero even when $\boldsymbol{\Sigma} = \lambda\boldsymbol{I}_p$. Thus, $\hat{a}_{2s}$ cannot be asymptotically normally distributed. Hence, we need to find an unbiased estimator of a_2 for a general class of distributions given by (1.1)-(1.3), or an estimator with bias of the order $O(n^{-1-\varepsilon}), \varepsilon > 0$. We propose an unbiased estimator $\hat{a}_2$ defined in (2.5). Its unbiasedness will be shown in Section 2.1, and the variances of $\hat{a}_1, \hat{a}_2$, and $Cov(\hat{a}_1, \hat{a}_2)$ will be given in subsequent sections.

We define an estimator of a_2 given by,

$$\begin{aligned}\hat{a}_2 &= \frac{1}{f}\left\{(N-2)n\mathrm{tr}\,[\boldsymbol{V}^2] - Nn\mathrm{tr}\,[\boldsymbol{D}^2] + (\mathrm{tr}\,[\boldsymbol{V}])^2\right\}\\ &= \frac{1}{f}\left\{(N-2)n\mathrm{tr}\,[\boldsymbol{M}^2] - Nn\mathrm{tr}\,[\boldsymbol{D}^2] + (\mathrm{tr}\,[\boldsymbol{D}])^2\right\},\end{aligned} \tag{2.5}$$

where $f = pN(N-1)(N-2)(N-3)$, $\boldsymbol{M} = \boldsymbol{Y}'\boldsymbol{Y}$, $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N)$ and

$$\boldsymbol{D} = \mathrm{diag}(\boldsymbol{y}_1'\boldsymbol{y}_1, ..., \boldsymbol{y}_N'\boldsymbol{y}_N),$$

namely, $\boldsymbol{D}$ denotes an $N \times N$ diagonal matrix with diagonal elements given by $\boldsymbol{y}_1'\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N'\boldsymbol{y}_N$. It will be shown in the following Section 2.1 that $\hat{a}_2$ is an unbiased estimator of $a_2 = \mathrm{tr}\,[\boldsymbol{\Sigma}^2]/p$ from which an unbiased estimator of $\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$ is given by $p\hat{a}_2$. It may be noted that it takes no longer time to compute $\hat{a}_2$ given in (2.5) than to compute $\hat{a}_{2s}$ given in (2.4). It may also be noted that from computing viewpoint, the expression given in the second line of (2.5) is better suited as all the matrices are $N \times N$ matrices, while the expression in the first line is a mixture of $N \times N$ and $p \times p$ matrices.

2.1. Unbiasedness of the estimator $\hat{a}_2$

In this subsection, we show that the estimator $\hat{a}_2$ defined in (2.5) is an unbiased estimator. For this, we need to compute the expected values of $\mathrm{tr}\,[\boldsymbol{V}^2]$, $\mathrm{tr}\,[\boldsymbol{D}^2]$, and $(\mathrm{tr}\,[\boldsymbol{V}])^2$. Note that,

$$\boldsymbol{x}_j - \overline{\boldsymbol{x}} = \boldsymbol{x}_j - \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{x}_k = \frac{n}{N}(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(j)})$$

where

$$\overline{\boldsymbol{x}}_{(j)} = \frac{1}{n}(N\overline{\boldsymbol{x}} - \boldsymbol{x}_j),\ n = N-1.$$

Also, note that $\boldsymbol{x}_j - \overline{\boldsymbol{x}}$ does not depend on the mean vector $\boldsymbol{\mu}$ given in the model (1.1), and thus we will assume without any loss of generality that $\boldsymbol{\mu} = \boldsymbol{0}$. Then, $E(\mathrm{tr}\,[\boldsymbol{D}^2])$ is expressed as

$$\begin{aligned} E(\mathrm{tr}\,[\boldsymbol{D}^2]) &= \sum_{j=1}^{N} E\left[(\boldsymbol{x}_j - \overline{\boldsymbol{x}})'(\boldsymbol{x}_j - \overline{\boldsymbol{x}})\right]^2 \\ &= \left(\frac{n}{N}\right)^4 \sum_{j=1}^{N} E\left[\boldsymbol{x}_j'\boldsymbol{x}_j - 2\boldsymbol{x}_j'\overline{\boldsymbol{x}}_{(j)} + \overline{\boldsymbol{x}}_{(j)}'\overline{\boldsymbol{x}}_{(j)}\right]^2 \\ &= \left(\frac{n}{N}\right)^4 \sum_{j=1}^{N} E\left[\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_j - 2\boldsymbol{u}_j'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)} + \overline{\boldsymbol{u}}_{(j)}'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)}\right]^2, \end{aligned}$$

which is rewritten as

$$\begin{aligned} &\left(\frac{n}{N}\right)^4 \sum_{j=1}^{N} E\left[(\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2 + 4(\boldsymbol{u}_j'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})^2 + (\overline{\boldsymbol{u}}_{(j)}'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})^2\right] \\ &+ \left(\frac{n}{N}\right)^4 \sum_{j=1}^{N} E\left[-4(\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_j)(\boldsymbol{u}_j'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_j) + 2(\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_j)(\overline{\boldsymbol{u}}_{(j)}'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})\right] \\ &+ \left(\frac{n}{N}\right)^4 \sum_{j=1}^{N} E\left[-4(\boldsymbol{u}_j'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})(\overline{\boldsymbol{u}}_{(j)}'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})\right]. \end{aligned}$$

Hence, using the results on the moments of quadratic form given in Section 7, we get

$$\begin{aligned} E(\mathrm{tr}\,[\boldsymbol{D}^2]) &= \frac{n^4}{N^3} E\left\{\left[(\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2\right] + 4E\left[(\boldsymbol{u}_j'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})^2\right] + E\left[(\overline{\boldsymbol{u}}_{(j)}'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})^2\right]\right\} \\ &\quad + 2\frac{n^4}{N^3} E\left[(\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_j)(\overline{\boldsymbol{u}}_{(j)}'\boldsymbol{\Sigma}\overline{\boldsymbol{u}}_{(j)})\right] \\ &= \frac{n(n^3+1)}{N^3} K_4 \sum_{i=1}^{p} \sigma_{ii}^2 + \frac{2n^2}{N}\mathrm{tr}\,[\boldsymbol{\Sigma}^2] + \frac{n^2}{N}(\mathrm{tr}\,[\boldsymbol{\Sigma}])^2. \end{aligned} \tag{2.6}$$

Following the above derivation, we obtain

$$E(\mathrm{tr}\,[\boldsymbol{V}])^2 = \frac{n^2}{N} K_4 \sum_{i=1}^{p} \sigma_{ii}^2 + 2n\mathrm{tr}\,[\boldsymbol{\Sigma}^2] + n^2\,(\mathrm{tr}\,[\boldsymbol{\Sigma}])^2, \tag{2.7}$$

$$E(\mathrm{tr}\,[\boldsymbol{V}^2]) = \frac{n^2}{N} K_4 \sum_{i=1}^{p} \sigma_{ii}^2 + nN\mathrm{tr}\,[\boldsymbol{\Sigma}^2] + n\,(\mathrm{tr}\,[\boldsymbol{\Sigma}])^2. \tag{2.8}$$

Collecting the terms according to the formula of $\hat{a}_2$ in (2.5), we find that the coefficients of $K_4 \sum_{i=1}^p \sigma_{ii}^2$ and of $(\mathrm{tr}\,[\boldsymbol{\Sigma}])^2$ are zero. The coefficients of $\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$ is $N(N-1)(N-2)(N-3)/f$. Hence,

$$E(\hat{a}_2) = a_2,$$

proving that $\hat{a}_2$ is an unbiased estimator of a_2.

Next, we show that $\hat{a}_{2s}$ defined in (2.4) is a biased estimator of a_2. From (2.7) and (2.8), we get

$$\begin{aligned} E(\hat{a}_{2s}) &= \frac{1}{(n-1)(n+2)p} E\left\{\mathrm{tr}\,[\boldsymbol{V}^2] - \frac{1}{n}(\mathrm{tr}\,[\boldsymbol{V}])^2\right\} \\ &= \frac{1}{(n-1)(n+2)p}\left\{\frac{n^2-n}{N} K_4 a_{20} + (nN-2)\mathrm{tr}\,[\boldsymbol{\Sigma}^2]\right\} \\ &= \frac{n}{(n+2)N} K_4 a_{20} + \frac{n^2+n-2}{n^2+n-2}\, a_2 \\ &= \frac{n}{N(N+1)} K_4 a_{20} + a_2, \quad a_{20} = \frac{1}{p}\sum_{i=1}^p \sigma_{ii}^2. \end{aligned} \tag{2.9}$$

Thus, unless $K_4 = 0$, $\hat{a}_{2s}$ is not an unbiased estimator of a_2, as has been shown in Srivastava, Kollo, and von Rosen (2011) for $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$.

2.2. *Variance of* $\hat{a}_1$

In this section, we derive the variance of $\hat{a}_1$. The matrix of N independent observation vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ is given by $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$ with $E[\boldsymbol{X}] = (\boldsymbol{\mu}, \ldots, \boldsymbol{\mu}) = \boldsymbol{\mu}\mathbf{1}'$ where $\mathbf{1}$ is an N-vector of ones, $\mathbf{1} = (1, \ldots, 1)'$, and $\mathbf{Cov}\,(\boldsymbol{x}_i) = \boldsymbol{\Sigma}$, $i = 1, \ldots, N$. Let $\boldsymbol{A} = \boldsymbol{I}_N - N^{-1}\mathbf{1}\mathbf{1}'$, where $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. Then, with $n = N-1$, we define $\boldsymbol{S}$ as $n^{-1}\boldsymbol{V}$, which can be written as

$$\boldsymbol{S} = \frac{1}{n}(\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}') = \frac{1}{n}(\boldsymbol{X} - \boldsymbol{\mu}\mathbf{1}')\boldsymbol{A}(\boldsymbol{X} - \boldsymbol{\mu}\mathbf{1}')'.$$

Thus $\boldsymbol{S}$ does not depend on the mean vector $\boldsymbol{\mu}$. In the following calculations, we shall assume that $\boldsymbol{\mu} = \mathbf{0}$. Thus,

$$\boldsymbol{S} = \frac{1}{n}\left(\frac{n}{N}\sum_{i=1}^N \boldsymbol{x}_i\boldsymbol{x}_i' - \frac{1}{N}\sum_{j\neq k}^N \boldsymbol{x}_j\boldsymbol{x}_k'\right) = \frac{1}{N}\sum_{i=1}^N \boldsymbol{x}_i\boldsymbol{x}_i' - \frac{1}{nN}\sum_{j\neq k}^N \boldsymbol{x}_j\boldsymbol{x}_k',$$

and

$$\begin{aligned} \hat{a}_1 &= \frac{1}{p}\mathrm{tr}\,[\boldsymbol{S}] = \frac{1}{Np}\sum_{i=1}^N \boldsymbol{x}_i'\boldsymbol{x}_i - \frac{1}{Npn}\sum_{j\neq k}^N \boldsymbol{x}_j'\boldsymbol{x}_k \\ &= \frac{1}{Np}\sum_{i=1}^N \boldsymbol{x}_i'\boldsymbol{x}_i - \frac{2}{Npn}\sum_{j<k}^N \boldsymbol{x}_j'\boldsymbol{x}_k, \end{aligned}$$

where $\boldsymbol{x}_i$'s are i.i.d. with mean vector $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{F}\boldsymbol{F}$. We note that $Cov(\boldsymbol{x}_i'\boldsymbol{x}_i, \boldsymbol{x}_j'\boldsymbol{x}_k) = 0$ for $j \neq k$, and $Var(\sum_{j<k}^N \boldsymbol{x}_j'\boldsymbol{x}_k) = (Nn/2)\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$. Hence, from Lemma 7.1 in Section 7,

$$Var(\textstyle\sum_{i=1}^N \boldsymbol{x}_i'\boldsymbol{x}_i) = NVar(\boldsymbol{x}_i'\boldsymbol{x}_i) = NVar(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_i) = K_4 \displaystyle\sum_{i=1}^p \sigma_{ii}^2 + 2\mathrm{tr}\,[\boldsymbol{\Sigma}^2].$$

Thus, with

$$a_{20} = \frac{1}{p}\sum_{i=1}^p \sigma_{ii}^2, \quad \text{and} \quad a_2 = \frac{1}{p}\mathrm{tr}\,[\boldsymbol{\Sigma}^2], \tag{2.10}$$

we get

$$Var(\hat{a}_1) = \frac{1}{Np}K_4 a_{20} + \frac{2}{Np}a_2 + \frac{2}{Nnp}a_2$$
$$= \frac{2a_2}{np} + K_4\frac{a_{20}}{Np} = \frac{1}{np}\left(2a_2 + \frac{n}{N}K_4 a_{20}\right). \tag{2.11}$$

Thus,

$$\hat{a}_1 = \hat{a}_1^* + O_p(N^{-1}p^{-1/2}), \tag{2.12}$$

where

$$\hat{a}_1^* = \frac{1}{Np}\sum_{i=1}^{N}(\boldsymbol{x}_i - \boldsymbol{\mu})'(\boldsymbol{x}_i - \boldsymbol{\mu}). \tag{2.13}$$

We state these results in the following theorem.

Theorem 2.1 *For the general model given in* (1.1)-(1.3) *and under the assumption* (A), *the means of $\hat{a}_1$ as well as of $\hat{a}_1^*$ is a_1, and $\hat{a}_1^* - \hat{a}_1 = O_p(N^{-1}p^{-1/2})$. The variance of $\hat{a}_1$ is given by*

$$Var(\hat{a}_1) = (2a_2 + nK_4 a_{20}/N)/(np) \equiv C_{11}/(np).$$

2.3. Variance of $\hat{a}_2$

In this section, we derive the variance of $\hat{a}_2$, the estimator of a_2 given in (2.5). Since

$$\mathrm{tr}\,[\boldsymbol{V}^2] = \mathrm{tr}\left[\left(\sum_{i=1}^{N}\boldsymbol{y}_i\boldsymbol{y}_i'\right)^2\right] = \sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \mathrm{tr}\left[\sum_{i\neq j}^{N}(\boldsymbol{y}_i\boldsymbol{y}_i')(\boldsymbol{y}_j\boldsymbol{y}_j')\right]$$
$$= \sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_j)^2,$$
$$(\mathrm{tr}\,[\boldsymbol{V}])^2 = \left(\sum_{i=1}^{N}\boldsymbol{y}_i'\boldsymbol{y}_i\right)^2 = \sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)(\boldsymbol{y}_j'\boldsymbol{y}_j),$$

we can rewrite $\hat{a}_2$ with $f = pN(N-1)(N-2)(N-3)$, as

$$\hat{a}_2 = \frac{1}{f}\left[(N-2)n\sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_j)^2 + (N-2)n\sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2\right]$$
$$+ \frac{1}{f}\left[-Nn\sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)(\boldsymbol{y}_j'\boldsymbol{y}_j)\right]$$
$$= \frac{1}{f}\left[(N-2)n\sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_j)^2 - (2n-1)\sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \sum_{i\neq j}(\boldsymbol{y}_i'\boldsymbol{y}_i)(\boldsymbol{y}_j'\boldsymbol{y}_j)\right]$$
$$= \frac{1}{pN(N-3)}\sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_j)^2 - \frac{(2n-1)}{f}\sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 + \frac{1}{f}\sum_{i\neq j}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)(\boldsymbol{y}_j'\boldsymbol{y}_j).$$

From the Markov inequality we find from (2.6) that for every $\varepsilon > 0$,

$$P\left\{(2n-1)f^{-1}\textstyle\sum_{i=1}^{N}(\boldsymbol{y}_i'\boldsymbol{y}_i)^2 \;>\; \varepsilon\right\} \;\leq\; \dfrac{2n}{f\varepsilon}\,E(\mathrm{tr}\,[\boldsymbol{D}^2])$$
$$\leq \frac{2}{N(N-2)(N-3)\varepsilon}\left(NK_4 a_{20} + 2Na_2 + pNa_1^2\right)$$
$$= O(N^{-2}) + O(pN^{-2}) = O(p^{1-2\delta}) = o(1),$$

since $N = O(p^\delta)$ for $\delta > 1/2$. Similarly,

$$\begin{aligned} P\left\{f^{-1}\textstyle\sum_{i\neq j}^N (\boldsymbol{y}_i'\boldsymbol{y}_i)(\boldsymbol{y}_j'\boldsymbol{y}_j) > \epsilon\right\} &\leq \frac{1}{f\epsilon}N(N-1)(\mathrm{tr}\,[\boldsymbol{\Sigma}])^2 \\ &= \frac{p}{(N-2)(N-3)\varepsilon}a_1^2 = o(1). \end{aligned}$$

Hence,

$$\begin{aligned} \hat{a}_2 &= \frac{1}{pN(N-3)}\sum_{i\neq j}^N (\boldsymbol{y}_i'\boldsymbol{y}_j)^2 + o_p(1) \\ &= \frac{1}{pN(N-3)}\sum_{i\neq j}^N \left\{(\boldsymbol{x}_i - \overline{\boldsymbol{x}})'(\boldsymbol{x}_j - \overline{\boldsymbol{x}})\right\}^2 + o_p(1) \\ &= \frac{1}{pN(N-1)}\sum_{i\neq j}^N \left\{(\boldsymbol{x}_i - \boldsymbol{\mu})'(\boldsymbol{x}_j - \boldsymbol{\mu})\right\}^2 + o_p(1) \\ &= a_2^* + o_p(1), \end{aligned}$$

where

$$\hat{a}_2^* = \frac{1}{pnN}\sum_{i\neq j}^N \left\{(\boldsymbol{x}_i - \boldsymbol{\mu})'(\boldsymbol{x}_j - \boldsymbol{\mu})\right\}^2.$$

Thus,

$$Var(\hat{a}_2) = Var(\hat{a}_2^*)(1 + o(1)).$$

To find the variance of $\hat{a}_2^*$, we define

$$z_{ij} = \frac{1}{\sqrt{p}}(\boldsymbol{x}_i - \boldsymbol{\mu})'(\boldsymbol{x}_j - \boldsymbol{\mu}) = \frac{1}{\sqrt{p}}\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j$$

from the model (1.1). Thus, $\hat{a}_2^* = \frac{1}{Nn}\sum_{i\neq j}^N z_{ij}^2$ and $E(z_{ij}^2) = a_2$. Hence, from the results on moments given in Section 7, we get

$$\begin{aligned} Var(\hat{a}_2^*) &= E\left[\frac{1}{nN}\sum_{i\neq j}^N (z_{ij}^2 - a_2)\right]^2 \\ &= \frac{2}{N^2n^2}E\left[\sum_{i\neq j}^N (z_{ij}^2 - a_2)^2 + \frac{4}{N^2n^2}\sum_{i\neq j\neq k}^N (z_{ij}^2 - a_2)(z_{ik}^2 - a_2)\right] \\ &= \frac{2}{Nn}Var(z_{ij}^2) + \frac{4(N-2)}{Nn}Cov(z_{ij}^2, z_{ik}^2) \quad (i \neq j \neq k) \\ &= \frac{2}{Np^2}Var[(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2] + \frac{4(N-2)}{Np^2}\left\{E\left[(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_k)^2\right] - a_2^2\right\}. \ (i \neq j \neq k) \end{aligned}$$

For any $i \neq j \neq k$, the second term of the above expression is

$$\frac{4(N-2)}{Np^2}\left\{E\left[(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_k)^2\right] - a_2^2\right\} = \frac{4(N-2)}{Np^2}\left\{K_4\sum_{\ell=1}^p \{(\boldsymbol{\Sigma}^2)_{\ell\ell}\}^2 + 2\mathrm{tr}\,[\boldsymbol{\Sigma}^4]\right\},$$

since $E\left[(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_k)^2\right] - a_2^2 = E\left[\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j\boldsymbol{u}_j'\boldsymbol{\Sigma}\boldsymbol{u}_i\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_k\boldsymbol{u}_k'\boldsymbol{\Sigma}\boldsymbol{u}_i\right] - a_2^2 = E\left[\boldsymbol{u}_i'\boldsymbol{\Sigma}^2\boldsymbol{u}_i\boldsymbol{u}_i'\boldsymbol{\Sigma}^2\boldsymbol{u}_i\right] - a_2^2 = E\left[(\boldsymbol{u}_i'\boldsymbol{\Sigma}^2\boldsymbol{u}_i)^2\right] - a_2^2 = Var[(\boldsymbol{u}_i'\boldsymbol{\Sigma}^2\boldsymbol{u}_i)^2] = K_4\sum_{\ell=1}^p \{(\boldsymbol{\Sigma}^2)_{\ell\ell}\}^2 + 2\mathrm{tr}\,[\boldsymbol{\Sigma}^4]$. For any $i \neq j$, from Lemmas 7.1 and 7.2, we get

$$Var[(\boldsymbol{u}_i'\boldsymbol{\Sigma}\boldsymbol{u}_j)^2] = K_4\sum_{k,\ell}^p \sigma_{k\ell}^4 + 6K_4\sum_{k=1}^p \{(\boldsymbol{\Sigma}^2)_{kk}\}^2 + 2(\mathrm{tr}\,[\boldsymbol{\Sigma}^2])^2 + 6\mathrm{tr}\,[\boldsymbol{\Sigma}^4].$$

Hence

$$Var(\hat{a}_2^*) = \frac{2}{Nnp^2}\left[K_4^2\sum_{i,j}^{p}\sigma_{ij}^4 + 6K_4\sum_{i=1}^{p}\left\{(\mathbf{\Sigma}^2)_{ii}\right\}^2 + 2(\mathrm{tr}\,[\mathbf{\Sigma}^2])^2 + 6\mathrm{tr}\,[\mathbf{\Sigma}^4]\right]$$
$$+\frac{4(N-2)}{Nnp^2}\left[K_4\sum_{i=1}^{p}\left\{(\mathbf{\Sigma}^2)_{ii}\right\}^2 + 2\mathrm{tr}\,[\mathbf{\Sigma}^4]\right],$$

which is rewritten as

$$\begin{aligned}Var(\hat{a}_2^*) &= \frac{4a_2^2}{Nn} + \frac{4(N+1)}{Nnp^2}K_4\sum_{i=1}^{p}\{(\mathbf{\Sigma}^2)_{ii}\}^2 + \frac{4(2N+1)}{Nnp^2}\,\mathrm{tr}\,[\mathbf{\Sigma}^4]\\ &= \frac{4a_2^2}{n^2} + \frac{4}{np}K_4b_4 + \frac{8}{np}a_4 + o(n^{-2})\\ &= \frac{4a_2^2}{n^2} + \frac{4}{np}(K_4b_4 + 2a_4) + o(n^{-2})\\ &\equiv \frac{1}{n^2}C_{22} + o(n^{-2}),\end{aligned}$$

where $C_{22} = 4[a_2^2 + (n/p)(K_4b_4 + 2a_4)]$, $b_4 = p^{-1}\sum_{i=1}^{p}\left\{(\mathbf{\Sigma}^2)_{ii}\right\}^2$ and $a_4 = p^{-1}\mathrm{tr}\,[\mathbf{\Sigma}^4]$. The above result is stated in the following theorem.

Theorem 2.2 *Under the general model given in* (1.1)-(1.3) *and under the assumption* (A)*, the variance of* $\hat{a}_2^*$ *is approximated as*

$$Var(\hat{a}_2^*) = \frac{4}{n^2}\left\{a_2^2 + \frac{n}{p}(K_4b_4 + 2a_4)\right\} + o(n^{-2}) = \frac{C_{22}}{n^2} + o(n^{-2}),$$

where C_{22}, b_4 *and* a_4 *have been defined above.*

2.4. Covariance between $\hat{a}_1$ *and* $\hat{a}_2$

In this section, we derive an expression for the covariance between $\hat{a}_1$ and $\hat{a}_2$ which is needed to obtain the joint distribution of $\hat{a}_1$ and $\hat{a}_2$. Since $\hat{a}_1$ and $\hat{a}_2$ are asymptotically equivalent to $\hat{a}_1^*$ and $\hat{a}_2^*$ respectively, we obtain the $Cov(\hat{a}_1^*, \hat{a}_2^*)$. In terms of model (1.1),

$$\hat{a}_1^* = \frac{1}{Np}\sum_{\ell=1}^{N}\boldsymbol{u}_\ell'\mathbf{\Sigma}\boldsymbol{u}_\ell \;\text{ and }\; \hat{a}_2^* = \frac{1}{pnN}\sum_{j\neq k}^{N}(\boldsymbol{u}_j'\mathbf{\Sigma}\boldsymbol{u}_k)^2$$

Thus,

$$\begin{aligned}Cov(\hat{a}_1^*, \hat{a}_2^*) &= \frac{1}{N^2np^2}E\left[\sum_{\ell=1}^{N}\boldsymbol{u}_\ell'\mathbf{\Sigma}\boldsymbol{u}_\ell\left(\sum_{j\neq k}^{N}\boldsymbol{u}_j'\mathbf{\Sigma}\boldsymbol{u}_k\right)^2\right] - a_1a_2\\ &= \frac{1}{N^2np^2}E\left[\sum_{\ell\neq j\neq k}^{N}(\boldsymbol{u}_\ell'\mathbf{\Sigma}\boldsymbol{u}_\ell)(\boldsymbol{u}_j'\mathbf{\Sigma}\boldsymbol{u}_k)^2 + 2\sum_{j\neq k}^{N}(\boldsymbol{u}_j'\mathbf{\Sigma}\boldsymbol{u}_j)(\boldsymbol{u}_j'\mathbf{\Sigma}\boldsymbol{u}_k)^2\right] - a_1a_2\\ &= \frac{N-2}{N}a_1a_2 + \frac{2}{Np^2}\left\{K_4\sum_{r=1}^{p}\sigma_{rr}(\mathbf{\Sigma}^2)_{rr} + 2\mathrm{tr}\,[\mathbf{\Sigma}^3] + \mathrm{tr}\,[\mathbf{\Sigma}]\mathrm{tr}\,[\mathbf{\Sigma}^2]\right\} - a_1a_2\\ &= \frac{2}{np}\,(K_4b_3 + 2a_3)\\ &\equiv \frac{1}{np}C_{12}.\end{aligned}$$

where $C_{12} = 2(K_4b_3 + 2a_3)$, $b_3 = p^{-1}\sum_{r=1}^{p}\sigma_{rr}(\mathbf{\Sigma}^2)_{rr}$ and $a_3 = p^{-1}\mathrm{tr}\,[\mathbf{\Sigma}^3]$ The above result is stated in the following theorem.

Theorem 2.3 *Under the general model given in* (1.1)-(1.3) *and under the assumption* (A)*, the covariance between* $\hat{a}_1^*$ *and* $\hat{a}_2^*$ *is given by*

$$Cov(\hat{a}_1^*, \hat{a}_2^*) = \frac{2}{np}(K_4 b_3 + 2a_3) = \frac{C_{12}}{np},$$

where b_3 *and* a_3 *are defined above.*

Corollary 2.1 *From the results of Theorems 2.1-2.3, the covariance matrix of* $(\hat{a}_1^*, \hat{a}_2^*)'$ *is approximated as*

$$\mathbf{Cov}\left[\begin{pmatrix}\hat{a}_1^*\\ \hat{a}_2^*\end{pmatrix}\right] = \begin{pmatrix} C_{11}/(np) & C_{12}/(np)\\ C_{12}/(np) & C_{22}/(n^2)\end{pmatrix} + o(n^{-2}) = \boldsymbol{C} + o(n^{-2}). \tag{2.14}$$

It will be shown in Section 6 that as $(N, p) \to \infty$, $(\hat{a}_1^*, \hat{a}_2^*)'$ is asymptotically distributed as bivariate normal with mean vector $(a_1, a_2)'$ and covariance matrix $\boldsymbol{C}$ as given above in (2.14).

3. Tests for Testing that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_p$

In this section, we consider the model described in (1.1)-(1.3), and propose tests for the two hypotheses, namely for testing sphericity and for testing the hypothesis that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$.
The sphericity hypothesis will be considered first in Section 3.1, and the hypothesis that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$ will be considered in Section 3.2.

3.1. Testing sphericity

For finite p, and $N \to \infty$, John (1972) proposed a locally best invariant test based on the statistic

$$U = \frac{\mathrm{tr}\,[\boldsymbol{S}^2]/p}{(\mathrm{tr}\,[\boldsymbol{S}]/p)^2} - 1,\ \boldsymbol{S} = \frac{1}{n}\boldsymbol{V} \tag{3.1}$$

and showed that for finite p, as $n \to \infty$, $NpU/2 \xrightarrow{d} \chi^2_d$ under the hypothesis that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$, where $d = p(p+1)/2-1$, and $\xrightarrow{d}$ denotes a convergence in distribution. Ledoit and Wolf (2002) showed that for $(n, p) \to \infty$ such that $p/N \to c$, the following modified statistic,

$$T_{LW} = (NU - p) \xrightarrow{d} \mathcal{N}(1, 4), \tag{3.2}$$

under the hypothesis that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$; the distribution of this statistic when $\boldsymbol{\Sigma} \neq \lambda \boldsymbol{I}_p$ was not given, but later given in Srivastava (2005).

Srivastava (2005) showed that from Cauchy-Schwartz inequality

$$\frac{a_2}{a_1^2} = \frac{\left(p^{-1}\sum_{i=1}^p \lambda_i^2\right)}{\left(p^{-1}\sum_{i=1}^p \lambda_i^2\right)^2} \geq 1, \text{ and } = 1 \text{ iff } \lambda_i = \lambda, \tag{3.3}$$

where λ_i are the eigenvalues of $\boldsymbol{\Sigma}$. Thus a measure of sphericity is given by

$$M_1 = \frac{a_2}{a_1^2} - 1, \tag{3.4}$$

which is equal to zero if and only if $\lambda_1 = \cdots = \lambda_p = \lambda$. Thus, Srivastava (2005) proposed a test based on unbiased and consistent estimators of a_1 and a_2, namely $\hat{a}_1$ and $\hat{a}_{2s}$ defined in (2.3) and (2.4) respectively under the assumption that the observations are i.i.d $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. This test statistic is given by

$$T_{1s} = \frac{n}{2}\left(\frac{\hat{a}_{2s}}{\hat{a}_1^2} - 1\right) \tag{3.5}$$

Srivastava (2005) showed that as $(n, p) \to \infty$ $T_{1s} \xrightarrow{d} \mathcal{N}(0, 1)$ under the hypothesis that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$. The asymptotic distribution of T_{1s} when $\boldsymbol{\Sigma} \neq \lambda \boldsymbol{I}_p$ is also given. It has been shown in (2.9) that the estimator $\hat{a}_{2s}$ is not unbiased for the general model given in (1.1)-(1.3).

An unbiased estimator of a_2 or $\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$ can be obtained by using Hoeffding's (1948) U-statistics, for details, see Fraser (1957, chapter 4), Serfling (1980) and Lee (1990). For example, if the mean vector $\boldsymbol{\mu}$ is zero, $\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$ can be estimated by

$$\mathrm{tr}\,[\boldsymbol{\Sigma}^2] = \frac{1}{N(N-1)} \sum_{i \neq j}^{N} (\boldsymbol{x}_i' \boldsymbol{x}_j)^2,$$

as has been done by Ahmad, Werner and Brunner (2008) in connection with testing mean vectors in high dimensional data. The computation of this estimator takes time of order $O(N^2)$, same as in calculating $\hat{a}_{2s}$. But when $\boldsymbol{\mu}$ is not known, an unbiased estimator of $\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$, using Hoeffding's U-statistic is given by

$$\begin{aligned}
\widehat{\mathrm{tr}\,[\boldsymbol{\Sigma}^2]} &= \frac{1}{4N(N-1)(N-2)(N-3)} \sum_{i \neq j \neq k \neq \ell}^{N} [(\boldsymbol{x}_i - \boldsymbol{x}_j)'(\boldsymbol{x}_k - \boldsymbol{x}_\ell)]^2 \\
&= \frac{1}{N(N-1)} \sum_{i \neq j}^{N} (\boldsymbol{x}_i' \boldsymbol{x}_j)^2 - \frac{2}{N(N-1)(N-2)} \sum_{i \neq j \neq k}^{N} \boldsymbol{x}_i' \boldsymbol{x}_j \boldsymbol{x}_j' \boldsymbol{x}_k \\
&\quad + \frac{1}{N(N-1)(N-2)(N-3)} \sum_{i \neq j \neq k \neq \ell}^{N} \boldsymbol{x}_i' \boldsymbol{x}_j \boldsymbol{x}_k' \boldsymbol{x}_\ell,
\end{aligned}$$

as used by Chen, Zhang and Zhong (2010) in replacing $\hat{a}_{2s}$ in Srivastava's statistic T_{1s} by using the above estimator divided by p. The above estimator of $\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$, however, has summation over four indices, and thus requires computing time of $O(N^4)$, which is not easy to compute.

Thus, in this paper, we use the estimator $\hat{a}_2$ in place of $\hat{a}_{2s}$, and propose the statistic T_1 given by

$$T_1 = \frac{n}{2} \left(\frac{\hat{a}_2}{\hat{a}_1^2} - 1 \right), \tag{3.6}$$

which is a one-sided test since we are testing the hypothesis $H_1 : M = 0$ against the alternative $A_1 : M_1 > 0$. In Section 6, we show that as $(N, p) \to \infty$, $(\hat{a}_1, \hat{a}_2)$ has a bivariate normal distribution with covariance matrix $\boldsymbol{C}$ given in (2.14). Thus, following the delta method for obtaining the asymptotic distribution as in Srivastava (2005) or Srivastava and Khatri (1979, page 59, Theorem 2.10.2), we can see that the variance of $(\hat{a}_2/\hat{a}_1^2)$ is approximated as $Var(\hat{a}_2/\hat{a}_1^2) = \tau^2(n, p) + o(n^{-2})$, where

$$\begin{aligned}
\tau^2(n, p) &= \left(\frac{-2a_2}{a_1^3}, \frac{1}{a_1^2} \right) \begin{pmatrix} C_{11}/(np) & C_{12}/(np) \\ C_{12}/(np) & C_{22}/(n^2) \end{pmatrix} \begin{pmatrix} -2a_2/a_1^3 \\ 1/a_1^2 \end{pmatrix} \\
&= \left(\frac{-2a_2 C_{11}}{npa_1^3} + \frac{C_{12}}{npa_1^2}, \frac{-2a_2 C_{12}}{npa_1^3} + \frac{C_{22}}{n^2 a_1^2} \right) \begin{pmatrix} -2a_2/a_1^3 \\ 1/a_1^2 \end{pmatrix} \\
&= \frac{4a_2^2 C_{11}}{npa_1^6} - \frac{2a_2 C_{12}}{npa_1^5} - \frac{2a_2 C_{12}}{npa_1^5} + \frac{C_{22}}{n^2 a_1^4} \\
&= \frac{4a_2^2}{n^2 a_1^4} + \frac{1}{np} \left[\frac{4a_2^2(K_4 a_{20} + 2a_2)}{a_1^6} - \frac{8a_2}{a_1^5}(K_4 b_3 + 2a_3) + \frac{4}{a_1^4}(K_4 b_4 + 2a_4) \right] \\
&= \frac{4a_2^2}{n^2 a_1^4} + \frac{1}{np} K_4 \left[\frac{4a_2^2 a_{20}}{a_1^6} - \frac{8a_2 b_3}{a_1^5} + 4\frac{b_4}{a_1^4} \right] + \frac{1}{np} \left[\frac{8a_2^3}{a_1^6} - \frac{16a_2 a_3}{a_1^5} + \frac{8a_4}{a_1^4} \right].
\end{aligned} \tag{3.7}$$

Thus, we get the following result, stated as Theorem 3.1.

Theorem 3.1 *As* $(n, p) \to \infty$, *the asymptotic distribution of the statistics* $\hat{a}_2/\hat{a}_1^2$ *under the assumption* (A) *and for the distribution given in* (1.1)-(1.3) *is given by*

$$\left(\hat{a}_2/\hat{a}_1^2 - a_2/a_1^2 \right) / \tau(n, p) \xrightarrow{d} \mathcal{N}(0, 1),$$

where $\tau^2(n, p)$ *is defined in* (3.7).

Under the hypothesis that $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$, $a_2/a_1^2 = 1$, $a_{20}/a_1^2 = 1$, $b_3/a_1^3 = 1$, $b_4/a_1^4 = 1$ and $a_4/a_1^4 = 1$. Hence, under the hypothesis, the variance of $(\hat{a}_2/\hat{a}_1^2)$ denoted by Var_0 is given by,

$$Var_0(\hat{a}_2/\hat{a}_1^2) = \frac{4}{n^2}. \tag{3.8}$$

Hence, we get the following result, stated as Corollary 3.1.

Corollary 3.1 *The asymptotic distribution of the test statistics T_1 when $\boldsymbol{\Sigma} = \lambda \boldsymbol{I}_p$, $\lambda > 0$, as $(n, p) \to \infty$, is given by*

$$T_1 \xrightarrow{d} \mathcal{N}(0, 1)$$

3.2. Testing $\boldsymbol{\Sigma} = \boldsymbol{I}_p$

In this section, we consider the problem of testing the hypothesis that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$. Nagao (1973) proposed the locally most powerful test given by,

$$\tilde{T}_1 = \frac{1}{p}\mathrm{tr}\,[(\boldsymbol{S} - \boldsymbol{I}_p)^2], \tag{3.9}$$

and showed that as N goes to infinity while p remains fixed, the limiting null distribution of

$$\frac{Np}{2}\tilde{T}_1 \xrightarrow{d} Y_{p(p+1)/2} \tag{3.10}$$

where Y_d denotes a χ^2 with d degrees of freedom. Ledoit and Wolf (2002) modified this statistic and proposed the statistic

$$W = \frac{1}{p}\mathrm{tr}\,[(\boldsymbol{S} - \boldsymbol{I}_p)^2] - \frac{p}{N}\hat{a}_1^2 + \frac{p}{N} \tag{3.11}$$

and showed that as $(N, p) \to \infty$ in a manner that $p/N \to c$,

$$nW - p \xrightarrow{d} \mathcal{N}(1, 4)$$

under the hypothesis that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$.

The test proposed by Nagao was based on an estimate of the distance

$$\begin{aligned} M_2 &= \frac{1}{p}\mathrm{tr}\,[(\boldsymbol{\Sigma} - \boldsymbol{I}_p)^2] \\ &= \frac{1}{p}(\mathrm{tr}\,[\boldsymbol{\Sigma}^2] - 2\mathrm{tr}\,[\boldsymbol{\Sigma}] + p) \\ &= a_2 - 2a_1 + 1. \end{aligned} \tag{3.12}$$

Srivastava (2005) proposed a test based on unbiased and consistent estimators $\hat{a}_1$ and $\hat{a}_{2s}$ under normality. This test is given by,

$$T_{2s} = \frac{n}{2}(\hat{a}_{2s} - 2\hat{a}_1 + 1). \tag{3.13}$$

As $(N, p) \to \infty$, it has been shown to be normally distributed under the assumption that the observations are normally distributed.

We propose the statistic

$$T_2 = \frac{n}{2}(\hat{a}_2 - 2\hat{a}_1 + 1) \tag{3.14}$$

and show that asymptotically as $(N, p) \to \infty$, T_2 is normally distributed. It may be noted that T_2 is also a one-sided test.

From the covariance matrix of $(\hat{a}_1, \hat{a}_2)$ given in (2.14), the variance of $(\hat{a}_2 - 2\hat{a}_1)$ can be approximated as $Var(\hat{a}_2 - 2\hat{a}_1) = \eta^2(n, p) + o(n^{-2})$, where

$$\begin{aligned} \eta^2(n, p) &= Var(\hat{a}_2) + 4Var(\hat{a}_1) - 4Cov(\hat{a}_1, \hat{a}_2) \\ &= \frac{4a_2^2}{n^2} + \frac{8a_4}{np} + \frac{8a_2}{np} - \frac{16a_3}{np} + \frac{4K_4}{np}(b_4 + a_{20} - 2b_3). \end{aligned} \tag{3.15}$$

We state these results in Theorem 3.2.

Theorem 3.2 *Under the assumption* (A) *and for the distribution given in* (1.1)-(1.3)*, the asymptotic distribution of the statistic* $\hat{a}_2 - 2\hat{a}_1$ *as* $(n, p) \to \infty$ *is given by,*

$$\{(\hat{a}_2 - 2\hat{a}_1) - (a_2 - 2a_2)\} / \eta(n, p) \xrightarrow{d} \mathcal{N}(0, 1), \tag{3.16}$$

where $\eta^2(n, p) = Var(\hat{a}_2 - 2\hat{a}_1)$ *given in* (3.15).

Since when the hypothesis that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, $a_2 - 2a_1 = -1$, and $\eta^2(n, p) = 4/n^2$, we get the following Corollary.

Corollary 3.2 *The asymptotic distribution of the best statistic* T_2 *when* $\boldsymbol{\Sigma} = \boldsymbol{I}_p$*, is given by*

$$T_2 \xrightarrow{d} \mathcal{N}(0, 1),$$

as $(n, p) \to \infty$.

4. Tests for the Equality of Two Covariance Matrices

In this section, we consider the problem of testing the hypothesis of the equality of two covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ when N_1 i.i.d p-dimensional observation vectors $\boldsymbol{x}_{ij}$, $j = 1, \ldots, N_1$ are obtained from the first group following the model (1.1) - (1.3) with $\boldsymbol{F}$ replaced by $\boldsymbol{F}_1$, $\boldsymbol{\mu}$ by $\boldsymbol{\mu}_1$ and $\boldsymbol{u}_j$ by $\boldsymbol{u}_{1j}$ where $\boldsymbol{\Sigma}_1 = \boldsymbol{F}_1^2$. And, similarly N_2 i.i.d p-dimensional vectors $\boldsymbol{x}_{2j}$, $j = 1, \ldots, N_2$ are obtained from the second group following the model (1.1)-(1.3) with $\boldsymbol{F}$ replaced by $\boldsymbol{F}_2$, μ by $\boldsymbol{\mu}_2$ and $\boldsymbol{u}_j$ by $\boldsymbol{u}_{2j}$, where $\boldsymbol{\Sigma}_2 = \boldsymbol{F}_2^2$. The sample mean vectors are now given by

$$\overline{\boldsymbol{x}}_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} \boldsymbol{x}_{1j}, \text{ and } \overline{\boldsymbol{x}}_2 = \frac{1}{N_2} \sum_{j=1}^{N_2} \boldsymbol{x}_{2j}.$$

Similarly, we define sample covariance matrices $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ through $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ given by,

$$\boldsymbol{V}_1 = \sum_{j=1}^{N_1} \boldsymbol{y}_{1j}\boldsymbol{y}'_{1j}, \ \boldsymbol{V}_1 = \sum_{j=1}^{N_2} \boldsymbol{y}_{2j}\boldsymbol{y}'_{2j}$$

$$\boldsymbol{S}_1 = \frac{1}{n_1}\boldsymbol{V}_1, \text{ and } \boldsymbol{S}_2 = \frac{1}{n_2}\boldsymbol{V}_2, \ n_i = N_i - 1, \ i = 1, 2,$$

where

$$\boldsymbol{y}_{1j} = \boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_1, j = 1, \ldots, N_1, \ \boldsymbol{y}_{2j} = \boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_2, j = 1, \ldots, N_2.$$

Under normality assumption, the unbiased and consistent estimators of $a_{1i} = \mathrm{tr}\,[\boldsymbol{\Sigma}_i]/p$ and $a_{2i} = \mathrm{tr}\,[\boldsymbol{\Sigma}_i^2]/p$ will be denoted by $\hat{a}_{1i}$ and $\hat{a}_{2is}$ respectively by using $\boldsymbol{V}_i$ in place of $\boldsymbol{V}_1$ and N_i or $n_i = N_i - 1$ in place of N, $i = 1, 2$. The unbiased estimator of a_{2i} under the general model will be denoted by $\hat{a}_{2i}$. Thus,

$$\hat{a}_{1i} = \frac{1}{n_i p}\mathrm{tr}\,[\boldsymbol{V}_i], \ \hat{a}_{2is} = \frac{1}{(n_i - 1)(n_i + 2)p} \left\{ \mathrm{tr}\,[\boldsymbol{V}_i^2] - \frac{1}{n_i}(\mathrm{tr}\,[\boldsymbol{V}_i])^2 \right\},$$

and

$$\hat{a}_{2i} = \frac{1}{f_i} \left\{ (N_i - 2)n_i \mathrm{tr}\,[\boldsymbol{V}_i^2] - Nn\mathrm{tr}\,[\boldsymbol{D}_i^2] + \mathrm{tr}\,[\boldsymbol{V}_i^2] \right\},$$

where for i = 1,2,

$$f_i = pN_i(N_i - 1)(N_i - 2)(N_i - 3)$$

$$\boldsymbol{D}_i = \mathrm{diag}(\boldsymbol{y}'_{i1}\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}'_{iN}\boldsymbol{y}_{iN_i}) : \ N_i \times N_i.$$

To test the hypothesis stated in Problem (3), namely testing the hypothesis $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, say, against the alternative $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$, Schott (2007) proposed the statistic

$$T_{Sc} = \frac{\hat{a}_{21s} + \hat{a}_{22s} - 2\mathrm{tr}\,[\boldsymbol{V}_1\boldsymbol{V}_2]/(pn_1n_2)}{2\hat{a}_{2s}\,(1/n_1 + 1/n_2)}, \tag{4.1}$$

where

$$\hat{a}_{2s} = \frac{1}{n_1 + n_2}(n_1\hat{a}_{21s} + n_2\hat{a}_{22s}), \quad (4.2)$$

is the estimator of $a_2 = p^{-1}\text{tr}[\boldsymbol{\Sigma}^2]$ under the hypothesis that $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. It may be noted that the square of the expression in the denominator of T_{Sc} is an estimate of the variance of the statistic in the numerator. Using the new unbiased estimator of a_{2i}, we obtain the statistic

$$T_3 = \frac{\hat{a}_{21} + \hat{a}_{22} - 2\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]/(pn_1n_2)}{\sqrt{\widehat{Var_0(\hat{q}_3)}}}, \quad (4.3)$$

where $\widehat{Var_0(\hat{q}_3)}$ denotes the estimated variance of the numerator of (4.3), namely, the estimated variance of

$$\hat{q}_3 = \hat{a}_{21} + \hat{a}_{22} - \frac{2}{pn_1n_2}\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2],$$

under the hypothesis that $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. The variance of $\hat{q}_3$ is as shown in Sections 4.1 and 4.2, is given by

$$\begin{aligned} Var_0(\hat{q}_3) &= Var(\hat{a}_{21}) + Var(\hat{a}_{22}) + \left(\frac{2}{pn_1n_2}\right)^2 Var(\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]) \\ &\quad - \frac{4}{pn_1n_2}\sum_{i=1}^{2} Cov(\hat{a}_{2i}, \text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]). \\ &= \frac{4a_2^2}{n_1^2} + \frac{4a_2^2}{n_2^2} + \frac{8a_2^2}{n_1n_2} + o(N_1^{-1}),\ N_1 \le N_2. \end{aligned}$$

Thus, assuming that $N_i/p \to 0$, $i = 1, 2$, as $(N_1, N_2, p) \to \infty$, the test statistic T_3 defined in (4.3) is given by

$$T_3 = \frac{\hat{a}_{21} + \hat{a}_{22} - 2\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]/(pn_1n_2)}{2\hat{a}_2\,(1/n_1 + 1/n_2)}, \quad (4.4)$$

where

$$\hat{a}_2 = \frac{1}{n_1 + n_2}(n_1\hat{a}_{21} + n_2\hat{a}_{22}). \quad (4.5)$$

It may be noted that T_3 is a one-sided test. That is, the hypothesis is rejected if $T_3 > z_{1-\alpha}$, where $z_{1-\alpha}$ is the upper $100(1-\alpha)\%$ point of the standard normal distribution.

4.1. Evaluation of Variance of $\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]$

To evaluate the variance of $\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]$,we note that $\boldsymbol{F}_1 = \boldsymbol{F}_2 = \boldsymbol{F}$, under the hypothesis that $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} = \boldsymbol{F}^2$, and asymptotically

$$\boldsymbol{V}_1 = \boldsymbol{F}\left(\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}\boldsymbol{u}'_{1i}\right)\boldsymbol{F}, \quad \boldsymbol{V}_2 = \boldsymbol{F}\left(\sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}'_{2j}\right)\boldsymbol{F}.$$

Thus

$$\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2] = \text{tr}\left[\left(\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}\boldsymbol{u}'_{1i}\right)\boldsymbol{\Sigma}\left(\sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}'_{2j}\right)\boldsymbol{\Sigma}\right] = \sum_{i=1}^{N_1}\boldsymbol{u}'_{1i}\boldsymbol{C}\boldsymbol{u}_{1i},$$

where $\boldsymbol{C} = (c_{ij}) = \boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\Sigma}$ and $\boldsymbol{B} = \sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}'_{2j}$. Hence,

$$\begin{aligned} Var_0(\text{tr}[\boldsymbol{V}_1\boldsymbol{V}_2]) &= Var\left(\textstyle\sum_{i=1}^{N_1}\boldsymbol{u}'_{1i}\boldsymbol{C}\boldsymbol{u}_{1i}\right) \\ &= E\left[Var\left(\textstyle\sum_{i=1}^{N_1}\boldsymbol{u}'_{1i}\boldsymbol{C}\boldsymbol{u}_{1i}\middle|\boldsymbol{C}\right)\right] + Var\left(E\left[\textstyle\sum_{i=1}^{N_1}\boldsymbol{u}'_{1i}\boldsymbol{C}\boldsymbol{u}_{1i}\middle|\boldsymbol{C}\right]\right). \end{aligned}$$

Note that

$$E\left[\textstyle\sum_{i=1}^{N_1}\boldsymbol{u}'_{1i}\boldsymbol{C}\boldsymbol{u}_{1i}\middle|\boldsymbol{C}\right] = N_1\text{tr}[\boldsymbol{C}] = N_1\sum_{j=1}^{N_2}\boldsymbol{u}'_{2j}\boldsymbol{\Sigma}^2\boldsymbol{u}_{2j}.$$

Hence

$$Var\left(E\left[\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}'\boldsymbol{C}\boldsymbol{u}_{1i}\middle|\boldsymbol{C}\right]\right) = N_1^2N_2\left[K_4\sum_{i=1}^{p}\{(\boldsymbol{\Sigma}^2)_{ii}\}^2 + 2\mathrm{tr}\,[\boldsymbol{\Sigma}^4]\right].$$

Thus, under the assumption (A),

$$\frac{4}{p^2n_1^2n_2^2}Var\left(E\left[\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}'\boldsymbol{C}\boldsymbol{u}_{1i}\middle|\boldsymbol{C}\right]\right) \le \frac{1}{N_2p^2}(K_4+2)\mathrm{tr}\,[\boldsymbol{\Sigma}^4] = o((np)^{-1}),$$

and hence

$$\frac{1}{p^2n_1^2n_2^2}Var(\mathrm{tr}\,[\boldsymbol{V}_1\boldsymbol{V}_2]) = \frac{1}{p^2n_1^2n_2^2}E\left[Var\left(\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}'\boldsymbol{C}\boldsymbol{u}_{1i}\middle|\boldsymbol{C}\right)\right] + o((np)^{-1}).$$

For a given $\boldsymbol{C} = \boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\Sigma}$, where $\boldsymbol{B} = \sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}_{2j}'$,

$$\begin{aligned}Var\left(\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}'\boldsymbol{u}_{1i}\right) &= N_1\left(K_4\sum_{i=1}^{p}c_{ii}^2 + 2\mathrm{tr}\,[\boldsymbol{C}^2]\right)\\ &= N_1\left(K_4\sum_{i=1}^{p}c_{ii}^2 + 2\mathrm{tr}\,[\boldsymbol{\Sigma}^2\boldsymbol{B}\boldsymbol{\Sigma}^2\boldsymbol{B}]\right).\end{aligned}$$

Let $\boldsymbol{\Sigma} = (\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_p)$. Then,

$$\begin{aligned}\sum_{i=1}^{p}c_{ii}^2 = \sum_{i=1}^{p}(\boldsymbol{\sigma}_i'\boldsymbol{B}\boldsymbol{\sigma}_i)^2 &= \sum_{i=1}^{p}\left\{\boldsymbol{\sigma}_i'\left(\sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}_{2j}'\right)\boldsymbol{\sigma}_i\right\}^2\\ &= \sum_{i=1}^{p}\left\{\sum_{j=1}^{N_2}(\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j})^2\right\}^2 = \sum_{i=1}^{p}\left\{\sum_{j=1}^{N_2}(\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j})^4 + \sum_{j\neq k}^{N_2}\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j}\boldsymbol{u}_{2k}'\boldsymbol{\sigma}_i\right\}\\ &= \sum_{i=1}^{p}\left(\sum_{j=1k}^{N_2}\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j}\boldsymbol{u}_{2k}'\boldsymbol{\sigma}_i + \sum_{j\neq k}^{N_2}\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j}\boldsymbol{u}_{2k}'\boldsymbol{\sigma}_i\right).\end{aligned}$$

Hence,

$$\begin{aligned}E\left(\sum_{i=1}^{p}c_{ii}^2\right) &= N_2\sum_{i=1}^{p}E[(\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j}\boldsymbol{u}_{2j}'\boldsymbol{\sigma}_i)^2] = N_2\sum_{i=1}^{p}E[(\boldsymbol{u}_{2j}'\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i'\boldsymbol{u}_{2j})^2]\\ &= N_2\left[K_4\sum_{i=1}^{p}\sum_{j=1}^{p}\{(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i')_{jj}\}^2 + 3\sum_{i=1}^{p}(\boldsymbol{\sigma}_i'\boldsymbol{\sigma}_i)^2\right]\\ &\le N_2(K_4+3)\sum_{i=1}^{p}(\boldsymbol{\sigma}_i'\boldsymbol{\sigma}_i)^2 \le N_2(K_4+3)\,\mathrm{tr}\,[\boldsymbol{\Sigma}^4],\end{aligned}$$

since $\boldsymbol{\Sigma}^4 = \sum_{i=1}^{p}\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i'\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i' + \sum_{i\neq j}^{p}\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i'\boldsymbol{\sigma}_j\boldsymbol{\sigma}_j'$. Similarly,

$$\begin{aligned}E(\mathrm{tr}\,[\boldsymbol{C}^2]) = E(\mathrm{tr}\,[\boldsymbol{\Sigma}^2\boldsymbol{B}\boldsymbol{\Sigma}^2\boldsymbol{B}]) &= E\left(\mathrm{tr}\left[\boldsymbol{\Sigma}^2\left(\sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}_{2j}'\right)\boldsymbol{\Sigma}^2\left(\sum_{j=1}^{N_2}\boldsymbol{u}_{2j}\boldsymbol{u}_{2j}'\right)\right]\right)\\ &= E\left[\sum_{j=1}^{N_2}\left(\boldsymbol{u}_{2j}\boldsymbol{\Sigma}^2\boldsymbol{u}_{2j}'\right)^2 + \sum_{j\neq k}^{N_2}\left(\boldsymbol{u}_{2j}'\boldsymbol{\Sigma}^2\boldsymbol{u}_{2k}\right)^2\right]\\ &= N_2\left[K_4\sum_{r=1}^{p}\{(\boldsymbol{\Sigma}^2)_{rr}\}^2 + \mathrm{tr}\,[\boldsymbol{\Sigma}^4] + (\mathrm{tr}\,[\boldsymbol{\Sigma}^2])^2\right] + N_2n_2\mathrm{tr}\,[\boldsymbol{\Sigma}^4].\end{aligned}$$

Thus,

$$\frac{1}{p^2n_1^2n_2^2}Var_0(\mathrm{tr}\,[\boldsymbol{V}_1\boldsymbol{V}_2]) = \frac{1}{p^2n_1n_2}\left[K_4^2\sum_{i=1}^{p}\sum_{j=1}^{p}(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i')_{jj} + 3K_4\sum_{i=1}^{p}(\boldsymbol{\sigma}_i'\boldsymbol{\sigma}_i)^2\right]$$
$$+ \frac{2}{p^2n_1n_2}\left[K_4\sum_{r=1}^{p}\{(\boldsymbol{\Sigma}^2)_{rr}\}^2 + 2\mathrm{tr}\,[\boldsymbol{\Sigma}^4] + (\mathrm{tr}\,[\boldsymbol{\Sigma}^2])^2 + n_2\mathrm{tr}\,[\boldsymbol{\Sigma}^4]\right]$$
$$= \frac{2a_2^2}{n_1n_2} + o(n^{-2}).$$

4.2. Evaluation of covariance between $\hat{a}_{2i}$ *and* $\mathrm{tr}\,[\boldsymbol{V}_1\boldsymbol{V}_2]$

The covariance between $\hat{a}_{21}$ and $\mathrm{tr}\,[\boldsymbol{V}_1\boldsymbol{V}_2]$ under the hypothesis is given by

$$C_{12}^{(1)} \equiv Cov_0[\hat{a}_{21}, \mathrm{tr}\,[\boldsymbol{V}_1\boldsymbol{V}_2]/(N_1N_2p)]$$
$$= \frac{1}{N_1^2N_2n_1p^2}E\left[\left\{\sum_{j\neq k}^{N_1}(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}\boldsymbol{u}_{1k})^2\right\}\mathrm{tr}\left[\boldsymbol{\Sigma}\left(\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}\boldsymbol{u}_{1i}'\right)\boldsymbol{\Sigma}\left(\sum_{\ell=1}^{N_2}\boldsymbol{u}_{2l}\boldsymbol{u}_{2l}'\right)\right]\right] - a_2^2$$
$$= \frac{N_2}{N_1^2N_2n_1p^2}E\left[\left\{\sum_{j\neq k}^{N_1}(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}\boldsymbol{u}_{1k})^2\right\}\mathrm{tr}\left[\boldsymbol{\Sigma}\left(\sum_{i=1}^{N_1}\boldsymbol{u}_{1i}\boldsymbol{u}_{1i}'\right)\right]\right] - a_2^2,$$

since $\boldsymbol{u}_{1j}$ and $\boldsymbol{u}_{2\ell}$ are independently distributed and $\hat{a}_{21}$ is independently distributed of $\boldsymbol{u}_{2\ell}$. Hence,

$$C_{12}^{(1)} = \frac{1}{N_1^2n_1p^2}E\left[\sum_{i\neq j\neq k}^{N_1}(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}\boldsymbol{u}_{1k})(\boldsymbol{u}_{1i}'\boldsymbol{\Sigma}^2\boldsymbol{u}_{1i}) + 2\sum_{j\neq k}^{N_1}(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}\boldsymbol{u}_{1k})^2(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}\boldsymbol{u}_{1j})\right] - a_2^2$$
$$= \frac{N_1n_1(n_1-1)}{N_1^2n_1}a_2^2 + \frac{2N_1n_1}{N_1^2n_1}\left[\frac{K_4\sum_{i=1}^{p}\{(\boldsymbol{\Sigma}^2)_{ii}\}^2}{p^2} + \frac{2\mathrm{tr}\,[\boldsymbol{\Sigma}^4]}{p^2} + a_2^2\right] - a_2^2$$
$$= \left(\frac{n_1-1}{N_1} + \frac{2}{N_1} - 1\right)a_2^2 + \frac{2}{N_1}\times o(1) = o(N_1^{-1}).$$

5. Attained Significance Level and Power

In this section we compare the performance of the proposed test statistics with the tests given under the normality assumption. The attained significance level to the nominal value $\alpha = 0.05$ and the power are investigated in finite samples by simulation.

The attained significance level (ASL) is defined by $\hat{\alpha}_T = \#(T_H > z_\alpha)/r_1$ for Problems (1) and (2), and by $\hat{\alpha}_T = \#(T_H^2 > \chi_{1,\alpha}^2)/r_1$ for Problem (3), where T_H are values of the test statistic T computed from data simulated under the null hypothesis H, r_1 is the number of replications, z_α is the $100(1-\alpha)\%$ quantile of the standard normal distribution and $\chi_{1,\alpha}^2$ is the $100(1-\alpha)\%$ quantile of the chi-square distribution. The ASL assesses how close the null distribution of T is to its limiting null distribution. From the same simulation, we also obtain $\widehat{z_\alpha}$ for Problems (1) and (2) and $\widehat{\chi_{1,\alpha}^2}$ for Problem (3) as the $100(1-\alpha)\%$ sample quantile of the empirical null distribution, and define the attained power by $\widehat{\beta_T} = \#(T_A > \widehat{z_\alpha})/r_2$ for Problems (1) and(2), $\widehat{\beta_T} = \#(T_A^2 > \widehat{\chi_{1,\alpha}^2})/r_2$ for Problem (3), where T_A are values of T computed from data simulated under the alternative hypothesis A. In our simulation, we set $r_1 = 10,000$ and $r_2 = 5,000$.

It may be noted that irrespective of the ASL of any statistic, the power has been computed when all the statistics in the comparison have the same specified significance level as the cut off points have been obtained by simulation. The ASL gives an idea as to how close it is to the specified significance level. If it is not close, the only choice left is to obtain it from simulation, not from the asymptotic distribution. It is common in practice, although not recommended, to depend on the asymptotic distribution, rather than relying on simulations to determine the ASL.

Through the simulation, let $\boldsymbol{\mu} = \boldsymbol{0}$ without loss of generality. For $j = 1, \ldots, N$, $\boldsymbol{u}_j = (u_{ij})$ given in the model (1.1) is generated with the four cases: one is the normal case and the others are the non-normal cases.

(Case 1) $u_{ij} \sim \mathcal{N}(0,1)$,
(Case 2) $u_{ij} = (v_{ij} - 32)/8$ for $v_{ij} \sim \chi^2_{32}$,
(Case 3) $u_{ij} = (v_{ij} - 8)/4$ for $v_{ij} \sim \chi^2_8$,
(Case 4) $u_{ij} = (v_{ij} - 2)/2$ for $v_{ij} \sim \chi^2_2$,

where χ^2_m denotes the chi-square distribution with m degrees of freedom, and u_{ij} are standardized. Since the skewness and kurtosis $(K_4 + 3)$ of χ^2_m is, respectively, $(8/m)^{1/2}$ and $3 + 12/m$, it is noted that χ^2_2 has higher skewness and kurtosis than χ^2_8 and χ^2_{32}. Following (1.1), $\boldsymbol{x}_j$ is generated by $\boldsymbol{x}_j = \boldsymbol{F}\boldsymbol{u}_j$ for $\boldsymbol{\Sigma} = \boldsymbol{F}^2$.

[1] Testing problems (1) and (2). For these testing problems, the null and alternative hypotheses we treat are $H : \boldsymbol{\Sigma} = \boldsymbol{I}_p$ and $A : \boldsymbol{\Sigma} = \mathrm{diag}(d_1, \dots, d_p)$, $d_i = 1 + (-1)^{i+1}(p-i)/(2p)$. We compare the two tests T_{1s} and T_1, given in (3.5) and (3.6) for Problem (1), and the tests T_{2s} and T_2, given in (3.13) and (3.15) for Problem (2). It is noted that the 95% point of the standard normal distribution is 1.64485. The simulation results are reported in Tables 1-4 for Problems (1) and (2), respectively.

From the tables, it is observed that the attained significance level (ASL) of the proposed tests T_1 and T_2 are close to the specified level while the ASL values of the tests T_{1s} and T_{2s} proposed under normal distribution are much inflated in Cases 3 and 4. Concerning the powers, both tests have similar performances although the proposed tests are slightly more powerful in Case 4.

[2] Testing problem (3). For this testing problem, the covariance matrix $\boldsymbol{\Sigma}$ we treat here is of the form

$$\boldsymbol{\Sigma}_{(\rho)} = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix} \begin{pmatrix} \rho^{|1-1|^{\frac{1}{10}}} & \rho^{|1-2|^{\frac{1}{10}}} & \cdots & \rho^{|1-p|^{\frac{1}{10}}} \\ \rho^{|2-1|^{\frac{1}{10}}} & \rho^{|2-2|^{\frac{1}{10}}} & \cdots & \rho^{|2-p|^{\frac{1}{10}}} \\ & \cdots & & \\ \rho^{|p-1|^{\frac{1}{10}}} & \rho^{|p-2|^{\frac{1}{10}}} & \cdots & \rho^{|p-p|^{\frac{1}{10}}} \end{pmatrix} \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix},$$

where $\sigma_i = 1 + (-1)^{i+1}U_i/2$ for a random variable U_i having uniform distribution $U(0,1)$. Then we consider the null and alternative hypotheses given by $H : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{(0.1)}$ for $\rho = 0.1$ and $A : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{(0.1)}$ for $\rho = 0.1$, $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{(0.3)}$ for $\rho = 0.3$. We compare the two tests T_{Sc} and T_3, given in (4.1) and (4.4). The simulation results are reported in Table 5.

For the table, it is revealed that the ASL of the proposed test T_3 are closer to the nominal level than T_{Sc} in Cases 3 and 4. Concerning the powers, both tests have similar performances although the proposed test has slightly more powerful in Case 4.

6. Asymptotic Distributions

In this section, we show that all the test statistics proposed in Sections 3 and 4 are asymptotically normally distributed as (N_1, N_2, p) go to infinity. The test statistic T_i depends on $\hat{a}_{2i}$ and $\hat{a}_{1i}$, $i = 1, 2$ or simply on $(\hat{a}_2, \hat{a}_1)$ in one-sample case. We shall consider $(\hat{a}_{21}, \hat{a}_{11})$ or equivalently $(\hat{a}^*_{21}, \hat{a}^*_{11})$ in probability. To obtain asymptotic normality, we consider a linear combination $l_1\hat{a}^*_{21} + l_2\hat{a}^*_{11}$ of $\hat{a}^*_{21}$ and $\hat{a}^*_{11}$, where we assume without any loss of generality that $l_1^2 + l_2^2 = 1$. We shall know that for all l_1 and l_2, $l_1\hat{a}_{21} + l_2\hat{a}_{11}$ is normally distributed from which the joint normality of $\hat{a}_{21}$ and $\hat{a}_{11}$ follow. Note that

$$\begin{aligned}
\hat{a}^*_{21} - a_2 &= \frac{1}{N_1 n_1 p} \sum_{i \neq j}^{N_1} (\boldsymbol{u}'_{1i}\boldsymbol{\Sigma}\boldsymbol{u}_{1j})^2 - a_2 \\
&= \frac{1}{N_1 n_1 p} \sum_{i \neq j}^{N_1} \left\{ (\boldsymbol{u}'_{1i}\boldsymbol{\Sigma}\boldsymbol{u}_{1j})^2 - \boldsymbol{u}'_{1j}\boldsymbol{\Sigma}^2\boldsymbol{u}_{1j} + \boldsymbol{u}'_{1j}\boldsymbol{\Sigma}^2\boldsymbol{u}_{1j} \right\} - a_2 \\
&= \frac{1}{N_1 n_1 p} \sum_{i \neq j}^{N_1} \left\{ (\boldsymbol{u}'_{1i}\boldsymbol{\Sigma}\boldsymbol{u}_{1j})^2 - \boldsymbol{u}'_{1j}\boldsymbol{\Sigma}^2\boldsymbol{u}_{1j} \right\} + A,
\end{aligned}$$

where $A = (N_1 n_1 p)^{-1} \sum_{i\neq j}^{N_1} \left(\boldsymbol{u}'_{1j} \boldsymbol{\Sigma}^2 \boldsymbol{u}_{1j} - \mathrm{tr}\,[\boldsymbol{\Sigma}^2] \right) = (N_1 p)^{-1} \sum_{i=1}^{N_1} \left(\boldsymbol{u}'_{1i} \boldsymbol{\Sigma}^2 \boldsymbol{u}_{1i} - \mathrm{tr}\,[\boldsymbol{\Sigma}^2] \right)$ and

$$\begin{aligned} Var(A) &= \frac{1}{(N_1 p)^2} \sum_{i=1}^{N_1} Var(\boldsymbol{u}'_{1i} \boldsymbol{\Sigma}^2 \boldsymbol{u}_{1i}) \\ &\leq \frac{1}{(N_1 p)^2} 4N_1 (K_4 + 2) \mathrm{tr}\,[\boldsymbol{\Sigma}^4] = o(N_1^{-1}), \end{aligned}$$

since $\mathrm{tr}\,[\boldsymbol{\Sigma}^4]/p^2 = o(1)$ from Assumption A. Thus, $A \to 0$ in probability. Hence,

$$n_1(\hat{a}_{21}^* - a_2) \overset{p}{=} \frac{2}{N_1 p} \sum_{i=2}^{N_1} \sum_{j=1}^{i-1} \left\{ (\boldsymbol{u}'_{1i} \boldsymbol{\Sigma} \boldsymbol{u}_{1j})^2 - \boldsymbol{u}'_{1j} \boldsymbol{\Sigma}^2 \boldsymbol{u}_{1j} \right\} = \sum_{i=2}^{N_1} \xi_i,$$

where $\xi_i \equiv 2(N_1 p)^{-1} \sum_{j=1}^{i-1} \left\{ (\boldsymbol{u}'_{1i} \boldsymbol{\Sigma} \boldsymbol{u}_{1j})^2 - \boldsymbol{u}'_{1j} \boldsymbol{\Sigma}^2 \boldsymbol{u}_{1j} \right\}$. Let $\mathfrak{I}_i^{(b)}$ be a σ-algebra generated by random vectors $\boldsymbol{u}_{11}, \ldots, \boldsymbol{u}_{1i}$, $i = 1, ...N_1$, and let $(\Omega, \mathfrak{I}, P)$ be the probability space, where Ω is the whole space and P is the probability measure. Let $\emptyset$ be the null set. Then, with $\mathfrak{I}_0^{(p)} = (\emptyset, \Omega) = \mathfrak{I}_{-1}$, we find that $\mathfrak{I}_0^{(p)} \subset \mathfrak{I}_1^{(p)} \subset \cdots \subset \mathfrak{I}_{N_1}^{(p)} \subset \mathfrak{I}$, and $E(\xi_i | \mathfrak{I}_{i-1}) = 0$. Let $\boldsymbol{B}_\ell = \boldsymbol{\Sigma} \boldsymbol{u}_{1\ell} \boldsymbol{u}'_{1\ell} \boldsymbol{\Sigma}$ for $\ell = j,\ j \neq k$. Then

$$\begin{aligned} \left(\frac{N_1 p}{2} \right)^2 E(\xi_i^2 | \mathfrak{I}_{i-1}) &= \sum_{j=1}^{i-1} Var(\boldsymbol{u}'_{1i} \boldsymbol{B}_j \boldsymbol{u}_{1i}) + 2 \sum_{j<k}^{i-1} Cov(\boldsymbol{u}'_{1i} \boldsymbol{B}_j \boldsymbol{u}_{1i}, \boldsymbol{u}'_{1i} \boldsymbol{B}_k \boldsymbol{u}_{1i}) \\ &= \sum_{j=1}^{i-1} \left[K_4 \sum_{\ell=1}^{p} \{(\boldsymbol{B}_j)_{\ell\ell}\}^2 + 2\mathrm{tr}\,[\boldsymbol{B}_j^2] \right] + 2 \sum_{j<k}^{i-1} \left\{ K_4 \sum_{\ell=1}^{p} (\boldsymbol{B}_j)_{\ell\ell} (\boldsymbol{B}_k)_{\ell\ell} + 2\mathrm{tr}\,[\boldsymbol{B}_j \boldsymbol{B}_k] \right\}, \end{aligned}$$

where $(\boldsymbol{B}_j)_{\ell\ell}$ is the $(\ell, \ell)t^{th}$ diagonal element of the matrix $\boldsymbol{B}_j = ((\boldsymbol{B}_j)_{\ell r})$. Thus,

$$\begin{aligned} E(\xi_i^2) &\leq 4 \frac{(K_4 + 2)}{N_1^2 p^2} \left\{ (i-1) E(\boldsymbol{u}'_j \boldsymbol{\Sigma}^2 \boldsymbol{u}_j)^2 + \frac{(i-1)(i-2)}{2} \mathrm{tr}\,[\boldsymbol{\Sigma}]^4 \right\} \\ &\leq \frac{4(K_4 + 2)}{p^2} \left\{ (K_4 + 2) \mathrm{tr}\,[\boldsymbol{\Sigma}]^4 + (\mathrm{tr}\,[\boldsymbol{\Sigma}^2])^2 + \ \mathrm{tr}\,[\boldsymbol{\Sigma}^4] \right\} = O((p^{-1} \mathrm{tr}\,[\boldsymbol{\Sigma}^2])^2). \end{aligned}$$

Hence, the sequence $\{\xi_i, \mathfrak{I}_i\}$ is a sequence of square integrable martingale difference, see Shiryaev (1984) or Hall and Heyde (1980). Similarly, it can be shown that

$$\sum_{i=0}^{N_1} E(\xi_i^2 | \mathfrak{I}_{i-1}) \xrightarrow{p} \sigma_0^2$$

for some finite constant σ_0^2. Thus, it remains to show that the Lindberg's condition, namely

$$\sum_{i=0}^{N_1} E[\xi_i^2 I(|\xi_i| > \epsilon | \mathfrak{I}_{i-1})] \xrightarrow{p} 0.$$

is satisfied. It is known, see, e.g, Srivastava (2009), that this condition will be satisfied if we show that

$$\sum_{i=0}^{N_1} E(\xi_i^4) \to 0 \text{ as } N_1 \to \infty.$$

Next, we evaluate $E(\xi_i^4)$. Note that

$$\xi_i^2 = \left(\frac{2}{N_1 p} \right)^2 \left[\sum_{j=1}^{i-1} c_{ij}^2 + 2 \sum_{j<k}^{i-1} c_{ij} c_{ik} \right],$$

where

$$c_{ij} = \boldsymbol{u}'_{1i} \boldsymbol{B}_j \boldsymbol{u}_{1i} - \mathrm{tr}\,[\boldsymbol{B}_j],\ \ \boldsymbol{B}_j = \boldsymbol{\Sigma} \boldsymbol{u}_{1j} \boldsymbol{u}'_{1j} \boldsymbol{\Sigma}.$$

Hence, from an inequality in Rao (2002),

$$\xi_i^4 \leq 2\left(\frac{2}{N_1 p}\right)^4 \left[\left(\sum_{j=1}^{i-1} c_{ij}^2\right)^2 + 4\left(\sum_{j<k}^{i-1} c_{ij}c_{ik}\right)^2\right]$$
$$= 2\left(\frac{2}{N_1 p}\right)^4 \left[\sum_{j=1}^{i-1} c_{ij}^4 + 6\sum_{j<k}^{i-1} c_{ij}^2 c_{ik}^2 + 4\sum_{j<k,l<r}^{i-1} c_{ij}c_{ik}c_{il}c_{ir}\right],$$

and

$$E(\xi_i^4) \leq 2\left(\frac{2}{N_1 p}\right)^4 E\left[\sum_{j=1}^{i-1} c_{ij}^4 + 6\sum_{j<k}^{i-1} c_{ij}^2 c_{ik}^2\right]$$
$$\leq 32\left(\frac{1}{N_1 p}\right)^4 \left[\sum_{j=1}^{i-1} E(c_{ij}^4) + 6\sum_{j<k}^{i-1} E(c_{ij}^2 c_{ik}^2)\right]$$
$$\leq 96\frac{1}{(N_1 p)^4}\left(\sum_{j=1}^{i-1} c_{ij}^2\right)^2.$$

Hence, using again an inequality in Rao (2002), we get

$$\sum_{i=1}^{N_1} E(\xi_i^4) \leq 192\frac{E(c_{ij}^4)}{N_1^3 p^4} \ (i \neq j),$$

which is of order $O(N_1^{-1})$ or converges to zero as $(N_1, p) \to \infty$ under the Assumptions (A1)-(A3) by using the results on the moments given in Section 7. For example, for some constant γ, we get from Corollary 7.2, and the fact that $N = O(p^\delta)$, $\delta > 1/2$,

$$\frac{1}{p^4}E[(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}^2\boldsymbol{u}_{1j})^4]$$
$$\leq \frac{1}{p^4}\gamma[(\text{tr}\,[\boldsymbol{\Sigma}^2])^4 + (\text{tr}\,[\boldsymbol{\Sigma}^2])^2\text{tr}\,[\boldsymbol{\Sigma}^4] + (\text{tr}\,[\boldsymbol{\Sigma}^4])^2 + \text{tr}\,[\boldsymbol{\Sigma}^2]\text{tr}\,[\boldsymbol{\Sigma}^6]) + \text{tr}\,[\boldsymbol{\Sigma}^8]],$$

which is of orer $O(1)$ under the Assumptions (A1)-(A4), because

$$\frac{1}{p^3}\text{tr}\,[\boldsymbol{\Sigma}^6] \leq \frac{1}{p^3}\text{tr}\,[\boldsymbol{\Sigma}^2]\text{tr}\,[\boldsymbol{\Sigma}^4] = a_2\frac{a_4}{p} \to 0, \quad \frac{1}{p^4}\text{tr}\,[\boldsymbol{\Sigma}^8] \leq \frac{1}{p^4}(\text{tr}\,[\boldsymbol{\Sigma}^4])^2 = \left(\frac{a_4}{p}\right)^2 \to 0.$$

Similarly, for some constant γ_1,

$$\frac{1}{p^4}E\{E[(\boldsymbol{u}_{1i}'\boldsymbol{B}_j\boldsymbol{u}_{1i})^4|\boldsymbol{B}_j]\} \leq \frac{1}{p^4}\gamma_1 E[2(\text{tr}\,[\boldsymbol{B}_j])^4 + (\text{tr}\,[\boldsymbol{B}_j])^2(\text{tr}\,[\,\boldsymbol{B}_j^2]) + (\text{tr}\,[\boldsymbol{B}_j^2])^2 + \text{tr}\,[\boldsymbol{B}_j^4]]$$
$$= \frac{5}{p^4}\gamma_1 E[(\boldsymbol{u}_{1j}'\boldsymbol{\Sigma}^2\boldsymbol{u}_{1j})^4].$$

Hence, $E[(\boldsymbol{u}_{1i}'\boldsymbol{B}_j\boldsymbol{u}_{1i})^4]/p^4 = O(1)$. Similarly,

$$\sqrt{N_1 p}(\hat{a}_{11}^* - a_1) = \frac{1}{\sqrt{N_1 p}}\left\{\sum_{i=1}^{N_1}(\boldsymbol{u}_{1i}'\boldsymbol{\Sigma}\boldsymbol{u}_{1i} - \text{tr}\,[\boldsymbol{\Sigma}])\right\} = \sum_{i=1}^{N_1}\xi_{2i},$$

where $\xi_{2i} = (N_1 p)^{-1/2}[\boldsymbol{u}_{1i}'\boldsymbol{\Sigma}\boldsymbol{u}_{1i} - \text{tr}\,[\boldsymbol{\Sigma}]]$. It can be checked that $E(\xi_{2i}|\Im_{i-1}) = E(\xi_{2i}) = 0$ and

$$E(\xi_{2i}^2|\Im_{i-1}) = Var(\xi_{2i}) = \frac{1}{N_1 p}Var(\boldsymbol{u}_{1i}'\boldsymbol{\Sigma}\boldsymbol{u}_{1i})$$
$$= \frac{1}{N_1 p}\left(K_4\sum_{j=1}^{p}\sigma_{jj}^2 + 2\text{tr}\,[\boldsymbol{\Sigma}^2]\right).$$

Hence, $\boldsymbol{\Sigma}_{i=1}^{N_1} E(\xi_i^2|\Im_{i-1}) = p^{-1}[K_4 \sum_{j=1}^{p} \sigma_{jj}^2 + 2\text{tr}[\boldsymbol{\Sigma}^2]] < \infty$. Similarly it can be shown that $\boldsymbol{\Sigma}_{i=1}^{N_1} E(\xi_i^4) \to 0$ as $(N_1, p) \to \infty$. Thus, asymptotically, as $(N_1, N_2, p) \to \infty$,

$$\boldsymbol{\Omega}_1^{-1/2} \begin{pmatrix} \hat{a}_{11} - a_1 \\ \hat{a}_{21} - a_2 \end{pmatrix} \xrightarrow{d} \mathcal{N}_2(\boldsymbol{0}, \boldsymbol{I}_2),$$

where $\boldsymbol{\Omega} = \boldsymbol{\Omega}_1^{1/2} \boldsymbol{\Omega}_1^{1/2}$ given by

$$\boldsymbol{\Omega} = \begin{pmatrix} C_{11}^{(1)}/(N_1 p) & C_{12}^{(1)}/(N_1 p) \\ C_{12}^{(1)}/(N_1 p) & C_{22}^{(1)}/(N_1^2) \end{pmatrix}$$

Thus, the asymptotic distribution of $(\hat{a}_{21}/\hat{a}_{11}^2)$ is normal with mean a_{21}/a_{11}^2 and variance ν_1^2 given by

$$\begin{aligned} \nu_1^2 &= (-2a_2/a_1^3, 1/a_1^2)\boldsymbol{\Omega}(-2a_2/a_1^3, 1/a_1^2)' \\ &= \frac{4a_2^2 C_{11}^{(1)}}{a_1^6 N_1 p} - \frac{4a_2 C_{12}^{(1)}}{a_1^5 N_1 p} + \frac{C_{22}^{(1)}}{a_1^4 N_1^2} \\ &= \frac{1}{N_1^2 a_1^4} \left(C_{22}^{(1)} - \frac{4N_1}{p} \frac{a_2}{a_1} C_{12}^{(1)} + \frac{4N_1}{p} \frac{a_2^2}{a_1^2} C_{11}^{(1)} \right). \end{aligned}$$

7. Moments of Quadratic forms

In this section we give the moments of quadratic forms, which are useful for evaluating the varaiances of $\hat{a}_1$ and $\hat{a}_2$. For proofs, see Srivastava (2009) and Srivastava and Kubokawa (2013).

Lemma 7.1 *Let $\boldsymbol{u} = (u_1, \ldots, u_p)'$ be a p-dimensional random vector such that $E(\boldsymbol{u}) = \boldsymbol{0}$, $\textbf{Cov}(\boldsymbol{u}) = \boldsymbol{I}_p$, $E(u_i^4) = K_4 + 3$, $i = 1, \ldots, p$, and*

$$E(u_i^a u_j^b u_k^c u_\ell^d) = E(u_i^a) E(u_j^b) E(u_k^c) E(u_\ell^d),$$

$0 \leq a+b+c+d \leq 4$ for all i, j, k, ℓ. Then for any $p \times p$ symmetric matrices $\boldsymbol{A} = (a_{ij})$ and $\boldsymbol{B} = (b_{ij})$ of constants, we have

$$\begin{aligned} E(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u})^2 &= K_4 \sum_{i=1}^{p} a_{ii}^2 + 2\text{tr}[\boldsymbol{A}^2] + (\text{tr}[\boldsymbol{A}])^2, \\ Var(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}) &= K_4 \sum_{i=1}^{p} a_{ii}^2 + 2\text{tr}[\boldsymbol{A}^2], \\ E[\boldsymbol{u}'\boldsymbol{A}\boldsymbol{u}\boldsymbol{u}'\boldsymbol{B}\boldsymbol{u}] &= K_4 \sum_{i=1}^{p} a_{ii} b_{ii} + 2\text{tr}[\boldsymbol{A}\boldsymbol{B}] + \text{tr}[\boldsymbol{A}]\text{tr}[\boldsymbol{B}], \end{aligned}$$

for any symmetric matrix $\boldsymbol{B} = (b_{ij})$ of constants.

Corollary 7.1 *Let $\overline{\boldsymbol{u}} = N^{-1} \sum_{i=1}^{N} \boldsymbol{u}_i$, where $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ are independently and identically distributed. Then*

$$Var(\overline{\boldsymbol{u}}'\boldsymbol{A}\overline{\boldsymbol{u}}) = \frac{K_4}{N^3} \sum_{i=1}^{p} a_{ii}^2 + \frac{2}{N^2} \text{tr}[\boldsymbol{A}^2].$$

Lemma 7.2 *Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be independently and identically distributed random vectors with zeroes mean vector and covariance matrix $\boldsymbol{I}_p$. Then for any $p \times p$ symmetric matrix $\boldsymbol{A} = (a_{ij})$,*

$$Var[(\boldsymbol{u}'\boldsymbol{A}\boldsymbol{v})^2] = K_4^2 \sum_{i,j}^{p} a_{ij}^4 + 6K_4 \sum_{i=1}^{p} \{(\boldsymbol{A}^2)_{ii}\}^2 + 6\text{tr}[\boldsymbol{A}^4] + 2(\text{tr}[\boldsymbol{A}^2])^2.$$

Note that for any symmetric matrix $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p) = \boldsymbol{A}'$, $(\boldsymbol{A}^2)_{ii} = \boldsymbol{a}_i'\boldsymbol{a}_i = \sum_{j=1}^{p} a_{ij}^2$, and $\sum_{i=1}^{p} \{(\boldsymbol{A}^2)_{ii}\}^2 = \sum_{i=1}^{p} (\sum_{j=1}^{p} a_{ij}^2)^2 = \sum_{i,j,k}^{p} a_{ij}^2 a_{ik}^2$.

8. Concluding Remarks

In this paper, we have proposed a new estimator of $p^{-1}\mathrm{tr}\,[\boldsymbol{\Sigma}^2]$ which is unbiased and consistent for a general class of distributions which includes normal distribution. The computing time for this estimator is the same as the one used in the literature under normality assumption. Using this new estimator we modified the tests proposed by Srivastava (2005) for testing the sphericity of the covariance matrix $\boldsymbol{\Sigma}$, and for testing $\boldsymbol{\Sigma} = \boldsymbol{I}_p$. The performance of these two modified tests are compared by simulation. It is shown that the attained significance level (ASL) of the proposed tests are close to the specified level while the tests proposed under normal distribution, the ASL is 83.61% for chi-square with 2 degrees of freedom. Thus, the modified proposed test is robust against departure from normality without losing power.

Acknowledgments.

We thanks the associate editor and the two reviewers for their kind suggestions and for pointing out a misprint and providing some recent references. The research of M.S. Srivastava was supported in part by NSERC of Canada, and that of T. Kubokawa by Grant-in-Aid for Scientific Research (21540114 and 23243039) from Japan Society for the Promotion of Science.

References

Ahmad, M., Werner, C., and Brunner, E. (2008). Analysis of high dimensional repeated measures designs: the one sample case. *Comput. Statist. Data Anal.*, **53**, 416–427.

Baik, J., and Silverstein, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Analysis*, **97**, 1382–1408.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Royal Statist. Soc.*, **B57**, 289–300.

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.*, **29**, 1165–1188.

Berthet, Q., and Rigollet, P. (2013). Optimal detection of sparse principal components in high dimension. *Ann. Statist.*, **41**, 1780–1815.

Cai, T.T., and Ma, Z. (2013). Optimal hypotheses testing for high dimensional covariance matrices. *Bernoulli*, **19**, 2359–2388.

Cai, T., Liu, W., and Xia, J. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc.*, **108**, 265–277.

Chen, S.S., Zhang, L., and Zhong, P. (2010). Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.*, **105**, 810–819.

Dempster, A.P. (1958). A high dimensional two-sample significance test. *Ann. Math. Statist.*, **29**, 995–1010.

Dempster, A.P. (1960). A significance test for the separation of two highly multivariate small samples. *Biometrics*, **16**, 41–50.

Fan, J., Hall, P., and Yao, Q. (2007). How many simultaneous hypothesis tests can normal, Student's t or bootstrap calibration be applied. *J. Amer. Statist. Assoc.*, **102**, 1282–1288.

Fraser, D.A.S. (1957). *Nonparametric Methods in Statistics*. Wiley, New York.

Hall, P., and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, New York.

Hoeffding, H. (1948). A class of statistics with asymptotically normal distributions. *Ann. Math. Statist.*, **19**, 293–325.

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

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

Kosorok, M., and Ma, S. (2007). Marginal asymptotics for the large p, small n paradigm: With applications to microarray data. *Ann. Statist.*, **35**, 1456–1486.

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

Lee, A.J. (1990). *U-Statistics: Theory and Practice*. Marcel Dekker.

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

Nagao, H. (1973). On some test criteria for covariance matrix. *Ann. Statist.*, **1**, 700–709.

Onatski, A., Moreira, M.J., and Hallin, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Ann. Statist.*, **41**, 1204–1231.

Rao, C.R. (2002). *Linear Statistical Inference and Its applications* (Paperback ed.). John Wiley & Sons.

Schott, J.R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample size. *Comput. Statsit. Data Anal.*, **51**, 6535–6542.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Shiryaev, A.N. (1984). *Probability* (2nd. ed.). Springer-Verlag.

Srivastava, M.S. (2005). Some tests concerning the covariance matrix in high-dimensional data. *J. Japan Statist. Soc.*, **35**, 251–272.

Srivastava, M.S. (2009). A test of the mean vector with fewer observations than the dimension under non-normality. *J. Multivariate Anal.*, **100**, 518–532.

Srivastava, M.S. (2010). Controlling the average false discovery in large-scale multiple testing. *J. Statist. Res.*, **44**, 85–102.

Srivastava, M.S., and Du, M. (2008). A test for the mean vector with fewer observations than the dimension. *J. Multivariate Anal.*, **99**, 386–402.

Srivastava, M.S., Katayama, S., and Kano, Y. (2013). A two sample test in high dimensional data. *J. Multivariate Anal.*, **114**, 249–358.

Srivastava, M.S., and Khatri, C.G. (1979). *An Introduction to Multivariate Statistics*, North-Holland, New York.

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. *J. Multivariate Anal.*, **102**, 1090–1103.

Srivastava, M.S., and Kubokawa, T. (2013). Tests for multivariate analysis of variance in high dimension under non-normality. *J. Multivariate Anal.*, **115**, 204–216.

Yamada, T., and Srivastava, M.S. (2012). A test for the multivariate analysis of variance in high-dimension. *Comm. Statist. Theory Methods*, **41**, 2602–2612.

Wilks, S.S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Ann. Math. Statist.*, **17**, 257–281.

Table 1: ASL and powers given in percentage (%) of the tests T_{1s} and T_1 for Problem (1), as well as the tests T_{2s} and T_2 for Problem (2), under Case 1, $\mathcal{N}(0,1)$

		ASL in H				**Power in A**			
p	N	T_{1s}	T_1	T_{2s}	T_2	T_{1s}	T_1	T_{2s}	T_2
20	10	5.10	6.71	5.50	7.39	11.02	10.53	10.86	10.43
	20	5.18	6.38	6.02	7.04	20.98	18.69	19.76	17.89
	40	5.36	5.78	5.68	6.23	48.92	47.22	47.47	45.11
	60	5.37	5.79	5.97	6.35	76.25	74.91	74.06	72.94
40	10	5.32	7.14	6.02	7.68	10.86	10.34	10.15	9.65
	20	5.51	6.01	5.88	6.43	19.59	18.88	19.28	18.29
	40	4.85	5.23	5.00	5.69	49.18	48.02	48.80	46.81
	60	5.27	5.33	5.69	5.74	76.00	75.00	74.29	74.24
60	10	5.05	7.05	5.36	7.51	11.45	9.94	11.17	10.01
	20	5.12	5.97	5.52	6.36	20.73	20.23	19.83	19.39
	40	4.85	5.30	5.09	5.54	50.13	48.27	49.51	48.13
	60	5.05	5.20	5.23	5.44	76.77	75.65	76.12	75.27
100	10	5.13	7.19	5.28	7.36	10.40	9.96	10.39	9.63
	20	5.08	6.08	5.29	6.26	19.99	18.60	19.88	18.29
	40	5.33	5.81	5.33	5.92	47.83	46.49	47.32	46.15
	60	4.70	5.15	4.70	5.25	77.43	75.93	77.01	75.48
200	10	5.38	7.09	5.55	7.20	10.91	10.74	10.72	10.62
	20	5.08	6.17	5.27	6.24	21.07	18.86	20.52	18.69
	40	4.66	5.21	4.76	5.32	50.66	48.19	49.96	47.86
	60	5.00	5.39	5.01	5.46	77.39	75.61	77.18	75.75

Table 2: ASL and powers given in percentage (%) of the tests T_{1s} and T_1 for Problem (1), as well as the tests T_{2s} and T_2 for Problem (2), under Case 2, χ^2_{32}

		ASL in H				**Power in A**			
p	N	T_{1s}	T_1	T_{2s}	T_2	T_{1s}	T_1	T_{2s}	T_2
20	10	6.50	6.41	7.37	7.19	10.82	10.89	10.50	9.85
	20	7.43	6.72	8.33	7.40	20.51	18.56	18.63	17.75
	40	7.50	6.10	8.37	6.90	47.54	45.89	44.61	44.10
	60	7.60	5.74	8.49	6.45	74.22	74.09	71.90	72.41
40	10	7.19	7.19	7.94	7.56	11.17	10.43	10.50	9.78
	20	7.81	6.60	8.40	6.95	19.20	18.09	18.24	18.04
	40	7.48	5.85	8.16	6.39	47.25	45.83	45.22	43.98
	60	7.53	5.73	7.96	6.00	74.05	73.03	72.40	72.07
60	10	6.85	7.18	7.27	7.37	11.40	10.38	11.40	10.24
	20	7.54	6.12	7.90	6.42	19.89	18.79	19.16	18.69
	40	7.72	5.80	7.91	5.84	47.34	46.42	46.02	45.57
	60	7.38	5.42	7.65	5.71	75.40	75.11	74.46	74.22
100	10	6.82	6.63	6.96	6.98	12.21	11.33	11.95	10.86
	20	6.93	5.88	7.07	6.11	21.77	20.36	21.71	20.33
	40	7.49	5.88	7.68	6.05	46.88	45.52	46.03	44.86
	60	7.17	5.39	7.35	5.56	75.89	75.74	75.26	75.03
200	10	6.89	7.07	7.15	7.23	11.48	10.64	11.45	10.70
	20	7.47	6.30	7.39	6.33	19.44	18.93	19.61	18.76
	40	7.46	5.88	7.50	5.94	46.13	45.32	45.95	45.44
	60	7.21	5.32	7.31	5.30	77.05	76.27	76.89	76.19

Table 3: ASL and powers given in percentage (%) of the tests T_{1s} and T_1 for Problem (1), as well as the tests T_{2s} and T_2 for Problem (2), under Case 3, χ^2_{8}

		ASL in H				**Power in A**			
p	N	T_{1s}	T_1	T_{2s}	T_2	T_{1s}	T_1	T_{2s}	T_2
20	10	12.82	6.89	14.84	7.91	10.78	11.26	9.79	10.67
	20	15.03	6.72	17.22	7.75	19.25	18.58	16.47	17.15
	40	17.87	6.98	19.62	8.02	42.33	42.21	36.66	40.35
	60	17.86	6.68	19.98	7.72	66.10	68.09	61.49	65.93
40	10	14.38	7.14	15.74	7.46	10.41	10.50	10.27	10.08
	20	17.37	6.98	18.28	7.30	18.25	18.22	17.43	17.77
	40	18.07	6.57	19.48	7.22	43.15	42.81	40.59	41.32
	60	18.55	6.36	19.61	6.83	69.05	69.60	66.57	68.01
60	10	14.63	7.55	15.76	7.74	9.76	9.74	9.60	9.74
	20	16.71	6.33	17.20	6.86	18.12	18.80	17.39	18.59
	40	18.04	6.21	18.86	6.56	44.34	44.09	42.93	43.56
	60	18.50	5.89	19.20	6.09	72.60	73.01	71.32	72.41
100	10	15.27	7.17	15.59	7.53	10.65	10.28	10.11	10.19
	20	16.69	5.73	17.08	5.89	19.77	19.72	18.85	19.86
	40	17.55	5.89	18.17	6.02	45.93	45.71	45.68	44.94
	60	18.25	5.86	18.68	6.09	74.29	73.42	73.24	73.27
200	10	15.57	7.10	15.76	7.36	9.90	10.04	9.85	10.02
	20	16.68	6.01	16.93	6.07	20.91	19.72	20.44	19.79
	40	17.23	5.85	17.61	5.89	46.73	45.57	46.02	45.92
	60	18.09	5.22	18.22	5.48	75.52	75.41	74.91	75.04

Table 4: ASL and powers given in percentage (%) of the tests T_{1s} and T_1 for Problem (1), as well as the tests T_{2s} and T_2 for Problem (2), under Case 4, χ^2_2

		ASL in H				**Power in** A			
p	N	T_{1s}	T_1	T_{2s}	T_2	T_{1s}	T_1	T_{2s}	T_2
20	10	42.16	10.24	42.93	11.31	8.19	8.51	7.41	8.53
	20	55.24	10.59	56.78	11.96	11.84	13.91	9.98	12.89
	40	63.40	11.00	65.79	12.94	22.42	28.02	16.68	26.08
	60	66.69	11.01	69.74	13.27	37.48	45.73	27.85	43.16
40	10	50.09	9.19	49.20	9.60	8.76	10.30	8.33	10.32
	20	61.01	9.37	61.32	10.48	13.03	15.36	11.25	14.56
	40	71.25	9.22	72.11	10.36	25.41	32.21	20.78	31.30
	60	73.99	9.02	75.36	10.44	44.74	54.76	36.57	52.50
60	10	51.85	9.30	51.20	9.34	8.57	9.43	8.51	8.81
	20	64.54	8.67	64.22	9.09	12.84	16.41	11.39	15.99
	40	72.86	8.98	73.41	9.72	28.19	34.88	23.34	33.96
	60	76.84	8.35	77.50	9.27	49.91	59.88	42.20	58.27
100	10	53.80	8.44	53.48	8.64	9.26	10.05	9.18	10.26
	20	67.71	7.99	67.35	8.63	13.54	16.74	12.54	16.46
	40	75.83	7.76	75.94	8.37	31.83	39.01	27.61	37.99
	60	79.22	7.42	79.63	7.81	54.03	63.45	49.08	62.65
200	10	56.32	7.82	55.93	8.11	8.82	9.86	8.84	9.74
	20	69.45	7.12	69.67	7.24	14.26	17.59	13.51	17.93
	40	78.90	6.50	78.91	6.67	33.76	41.56	31.50	41.35
	60	81.15	6.17	81.18	6.32	61.95	70.88	59.13	70.42

Table 5: Critical values, ASL and powers given in percentage (%) of the tests T_{Sc} and T_3 for Problem (3)

		Critical Value		**ASL in** H		**Power in** A	
N	p	T_{Sc}	T_3	T_{Sc}	T_3	T_{Sc}	T_3
		Case 1 : $\mathcal{N}(0,1)$					
20	40	1.4359	1.4740	2.64	2.99	25.96	25.32
40	80	1.5489	1.5919	3.92	4.29	69.48	68.50
60	120	1.6208	1.6501	4.70	5.07	92.46	92.02
80	200	1.5983	1.6186	4.60	4.74	99.30	99.24
		Case 2 : χ^2_{32}					
20	40	1.5380	1.5127	3.46	3.49	24.98	24.54
40	80	1.6945	1.6178	5.65	4.75	68.74	68.58
60	120	1.7807	1.6894	6.54	5.48	91.92	91.98
80	200	1.7424	1.6257	6.07	4.80	99.04	99.04
		Case 3 : χ^2_8					
20	40	1.8263	1.5557	7.98	3.83	20.04	23.10
40	80	2.0472	1.6527	11.48	5.08	64.74	67.54
60	120	2.1271	1.6485	13.13	5.07	89.26	91.08
80	200	2.1575	1.6757	12.86	5.32	98.48	98.78
		Case 4 : χ^2_2					
20	40	2.6124	1.7723	31.78	6.46	11.64	19.58
40	80	3.4193	1.8795	51.70	7.21	43.84	60.76
60	120	3.6721	1.9193	59.15	7.95	78.16	88.58
80	200	3.6834	1.7761	62.21	6.48	96.42	98.72