Multi-sample problem in sphericity test with monotone missing data

Tetsuya Sato*, Ayaka Yagi*, and Takashi Seo*

* *Department of Applied Mathematics Tokyo University of Science*
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

Abstract

In this study, we consider the sphericity test under monotone missing data for a multi-sample problem. We derive the likelihood ratio (LR) and an asymptotic expansion of the likelihood ratio test (LRT) statistic as well as modified LRT statistic for the null distribution. We also provide approximate upper percentiles under the null hypothesis. Furthermore, we propose an unbiased estimator and evaluate estimators using the mean squared error (MSE). For complete data, we derive a similar result. We use Monte Carlo simulations to numerically evaluate the actual Type I error rates for the approximate upper percentiles and estimators. Additionally, we provide examples of the LRT statistic and modified LRT statistic as well as approximate upper percentiles under monotone missing data.

Key Words and Phrases: Asymptotic expansion, Chi-square distribution, Likelihood ratio test statistic, Maximum likelihood estimator, Mean squared error, Modified likelihood ratio test statistic, Multi sample problem.

1 Introduction

In this study, we consider the sphericity test, a test of variance-covariance matrices under monotone missing data for a multi-sample problem:

$$H_0 : \boldsymbol{\Sigma}^{(1)} = \cdots = \boldsymbol{\Sigma}^{(m)} = \sigma^2 \boldsymbol{I}_p \text{ vs. } H_1 : \text{not } H_0, \tag{1}$$

where σ^2 is unknown parameter and $\sigma^2 > 0$, $m \in \mathbb{N}$.

This study assumes that data are missing completely at random (MCAR) and follow multivariate normal distribution.

Box (1949) introduced a general distribution theory for a class of likelihood criteria, and Muirhead (1982) and Anderson (2003) discussed its general distribution theory. For a one-sample problem, Muirhead (1982) discussed the sphericity test under complete data and then provided an asymptotic expansion of the modified LRT statistic for the null distribution using the general distribution proposed by Box (1949). Under two-step monotone missing data, Chang and Richards (2010) discussed the null moment of the LR. Sato, Yagi, and Seo (2025) provided an asymptotic expansion of the LRT statistic and modified LRT statistic for the null distribution under complete data and monotone missing data. Kanda and Fujikoshi (1998) discussed the maximum likelihood estimators (MLEs) of the mean vector and variance-covariance matrix provided monotone missing data. For a multi-sample problem, Mendoza (1980) provided the MLE of σ^2 under the null hypothesis and limiting null distribution of the modified LRT statistic under complete data. Tsukada (2014) discussed equivalence testing for mean vectors and variance-covariance matrices under two-step monotone missing data, and then derived the LRT statistic and its limiting null distribution. In addition, Tsukada (2024) provided an asymptotic expansion of the modified LRT statistic for the null distribution using the general distribution proposed by Box (1949).

The remainder of this paper is organized as follows. In Section 2, we describe the LR for the sphericity test (1) under monotone missing data. In Section 3, we derive an asymptotic expansion of the LRT statistic and modified LRT statistic for the null distribution by comparing the moments of a random variable proposed by Box (1949) with the null moment of the LR. Furthermore, we describe the upper percentiles of the LRT statistic and modified LRT statistic under the null hypothesis and provide some approximate upper percentiles. In Section 4, we examine estimators of σ^2. In Section 5, we prove that the LR under monotone missing data is affine invariant under the null hypothesis. Section 6 describes an asymptotic expansion of the LRT statistic and modified LRT statistic for the null distribution and estimators of σ^2 under complete data. In Section 7, we numerically evaluate the actual type I error rates for the approximate upper percentiles using Monte Carlo simulation. In Section 8, we present numerical examples under monotone missing data. Finally, Section 9 concludes the study.

In this study, we examine the sphericity test (1) under two-step and general k-step monotone missing data, respectively.

2 LR under monotone missing data

Here, we consider the LR in the sphericity test (1) under monotone missing data for a multi-sample problem.

2.1 Two-step case

We consider the following data matrix $\boldsymbol{X}^{(\ell)}$ from the ℓth population ($\ell = 1, \ldots, m$).

$$\boldsymbol{X}^{(\ell)} = \begin{pmatrix} x_{11}^{(\ell)} & \cdots & x_{1p_1}^{(\ell)} & x_{1,p_1+1}^{(\ell)} & \cdots & x_{1p}^{(\ell)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{N_1^{(\ell)}1}^{(\ell)} & \cdots & x_{N_1^{(\ell)}p_1}^{(\ell)} & x_{N_1^{(\ell)},p_1+1}^{(\ell)} & \cdots & x_{N_1^{(\ell)}p}^{(\ell)} \\ x_{N_1^{(\ell)}+1,1}^{(\ell)} & \cdots & x_{N_1^{(\ell)}+1,p_1}^{(\ell)} & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{N^{(\ell)}1}^{(\ell)} & \cdots & x_{N^{(\ell)}p_1}^{(\ell)} & * & \cdots & * \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_{11}^{(\ell)\prime} & \boldsymbol{x}_{21}^{(\ell)\prime} \\ \vdots & \vdots \\ \boldsymbol{x}_{1N_1^{(\ell)}}^{(\ell)\prime} & \boldsymbol{x}_{2N_1^{(\ell)}}^{(\ell)\prime} \\ \boldsymbol{x}_{1,N_1^{(\ell)}+1}^{(\ell)\prime} & * \\ \vdots & \vdots \\ \boldsymbol{x}_{1N^{(\ell)}}^{(\ell)\prime} & * \end{pmatrix},$$

where $N^{(\ell)} = N_1^{(\ell)} + N_2^{(\ell)}, p = p_1 + p_2, N_1^{(\ell)} > p$ and "$*$" indicates a missing observation. Suppose that the data matrix $\boldsymbol{X}^{(\ell)}$ is multivariate normally distributed as follows.

$$\boldsymbol{x}_1^{(\ell)}, \ldots, \boldsymbol{x}_{N_1^{(\ell)}}^{(\ell)} \overset{i.i.d.}{\sim} N_p(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)}), \boldsymbol{x}_{1,N_1^{(\ell)}+1}^{(\ell)}, \ldots, \boldsymbol{x}_{1N^{(\ell)}}^{(\ell)} \overset{i.i.d.}{\sim} N_{p_1}(\boldsymbol{\mu}_1^{(\ell)}, \boldsymbol{\Sigma}_{11}^{(\ell)}),$$

where

$$\boldsymbol{x}_j^{(\ell)} = \left(\boldsymbol{x}_{1j}^{(\ell)\prime}, \boldsymbol{x}_{2j}^{(\ell)\prime}\right)', \ j = 1, \ldots, N_1^{(\ell)},$$

and

$$\boldsymbol{\mu}^{(\ell)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(\ell)} \\ \boldsymbol{\mu}_2^{(\ell)} \end{pmatrix}, \boldsymbol{\Sigma}^{(\ell)} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(\ell)} & \boldsymbol{\Sigma}_{12}^{(\ell)} \\ \boldsymbol{\Sigma}_{21}^{(\ell)} & \boldsymbol{\Sigma}_{22}^{(\ell)} \end{pmatrix}.$$

Furthermore, we define the sample mean vectors and Wishart matrices as follows.

$$\overline{\boldsymbol{x}}_{(1)1}^{(\ell)} = \frac{1}{N_1^{(\ell)}} \sum_{j=1}^{N_1^{(\ell)}} \boldsymbol{x}_{1j}^{(\ell)}, \ \overline{\boldsymbol{x}}_{(1)2}^{(\ell)} = \frac{1}{N_1^{(\ell)}} \sum_{j=1}^{N_1^{(\ell)}} \boldsymbol{x}_{2j}^{(\ell)}, \ \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} = \frac{1}{N_2^{(\ell)}} \sum_{j=N_1^{(\ell)}+1}^{N^{(\ell)}} \boldsymbol{x}_{1j}^{(\ell)},$$

$$\boldsymbol{W}_{(1)11}^{(\ell)} = \sum_{j=1}^{N_1^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(1)1}^{(\ell)})(\boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(1)1}^{(\ell)})',$$

$$\boldsymbol{W}_{(1)12}^{(\ell)} = (\boldsymbol{W}_{(1)21}^{(\ell)})' = \sum_{j=1}^{N_1^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(1)1}^{(\ell)})(\boldsymbol{x}_{2j}^{(\ell)} - \overline{\boldsymbol{x}}_{(1)2}^{(\ell)})',$$

$$\boldsymbol{W}^{(\ell)}_{(1)22} = \sum_{j=1}^{N_1^{(\ell)}} (\boldsymbol{x}^{(\ell)}_{2j} - \overline{\boldsymbol{x}}^{(\ell)}_{(1)2})(\boldsymbol{x}^{(\ell)}_{2j} - \overline{\boldsymbol{x}}^{(\ell)}_{(1)2})',$$

$$\boldsymbol{W}^{(\ell)}_{(2)11} = \sum_{j=N_1^{(\ell)}+1}^{N^{(\ell)}} (\boldsymbol{x}^{(\ell)}_{1j} - \overline{\boldsymbol{x}}^{(\ell)}_{(2)1})(\boldsymbol{x}^{(\ell)}_{1j} - \overline{\boldsymbol{x}}^{(\ell)}_{(2)1})' + \frac{N_1^{(\ell)} N_2^{(\ell)}}{N^{(\ell)}} (\overline{\boldsymbol{x}}^{(\ell)}_{(1)1} - \overline{\boldsymbol{x}}^{(\ell)}_{(2)1})(\overline{\boldsymbol{x}}^{(\ell)}_{(1)1} - \overline{\boldsymbol{x}}^{(\ell)}_{(2)1})',$$

$$\boldsymbol{W}^{(\ell)}_{(1)22\cdot 1} = \boldsymbol{W}^{(\ell)}_{(1)22} - \boldsymbol{W}^{(\ell)}_{(1)21} (\boldsymbol{W}^{(\ell)}_{(1)11})^{-1} \boldsymbol{W}^{(\ell)}_{(1)12}.$$

Subsequently, the LR of the test (1) under two-step monotone missing data is

$$\lambda_{(m)} = \left(\prod_{\ell=1}^{m} |\widehat{\boldsymbol{\Sigma}}^{(\ell)}_{11}|^{\frac{N^{(\ell)}}{2}} |\widehat{\boldsymbol{\Sigma}}^{(\ell)}_{22\cdot 1}|^{\frac{N_1^{(\ell)}}{2}} \right) (\widetilde{\sigma}^2)^{-\frac{Np_1 + N_1 p_2}{2}},$$

where $\widehat{\boldsymbol{\Sigma}}^{(\ell)}_{11}$ and $\widehat{\boldsymbol{\Sigma}}^{(\ell)}_{22\cdot 1}$ are the MLEs of $\boldsymbol{\Sigma}^{(\ell)}_{11}$ and $\boldsymbol{\Sigma}^{(\ell)}_{22\cdot 1} (= \boldsymbol{\Sigma}^{(\ell)}_{22} - \boldsymbol{\Sigma}^{(\ell)}_{21} {\boldsymbol{\Sigma}^{(\ell)}_{11}}^{-1} \boldsymbol{\Sigma}^{(\ell)}_{12})$, respectively, $\widetilde{\sigma}^2$ is the MLE of σ^2 under H_0, and

$$\widehat{\boldsymbol{\Sigma}}^{(\ell)}_{11} = \frac{1}{N^{(\ell)}} \left(\boldsymbol{W}^{(\ell)}_{(1)11} + \boldsymbol{W}^{(\ell)}_{(2)11} \right), \widehat{\boldsymbol{\Sigma}}^{(\ell)}_{22\cdot 1} = \frac{1}{N_1^{(\ell)}} \boldsymbol{W}^{(\ell)}_{(1)22\cdot 1},$$

$$\widetilde{\sigma}^2 = \frac{1}{Np_1 + N_1 p_2} \sum_{\ell=1}^{m} \left(\mathrm{tr}(\boldsymbol{W}^{(\ell)}_{(1)11} + \boldsymbol{W}^{(\ell)}_{(2)11}) + \mathrm{tr} \boldsymbol{W}^{(\ell)}_{(1)22} \right),$$

$$N = \sum_{\ell=1}^{m} N^{(\ell)}, N_1 = \sum_{\ell=1}^{m} N_1^{(\ell)}.$$

Furthermore, the LR $\lambda_{(m)}$ can written as

$$\lambda_{(m)} = \frac{\prod_{\ell=1}^{m} \left| \frac{1}{N^{(\ell)}} \left(\boldsymbol{W}^{(\ell)}_{(1)11} + \boldsymbol{W}^{(\ell)}_{(2)11} \right) \right|^{\frac{N^{(\ell)}}{2}} \left| \frac{1}{N_1^{(\ell)}} \boldsymbol{W}^{(\ell)}_{(1)22\cdot 1} \right|^{\frac{N_1^{(\ell)}}{2}}}{\left\{ \frac{1}{Np_1 + N_1 p_2} \sum_{\ell=1}^{m} \left(\mathrm{tr}(\boldsymbol{W}^{(\ell)}_{(1)11} + \boldsymbol{W}^{(\ell)}_{(2)11}) + \mathrm{tr} \boldsymbol{W}^{(\ell)}_{(1)22} \right) \right\}^{\frac{Np_1 + N_1 p_2}{2}}}.$$

2.2 k-step case ($k \geq 2$)

We extend Section 2.1 to genenal k-step monotone missing data. For $j = 1, \ldots, k, \ell = 1, \ldots, m$, we assume that the data are multivariate normally distributed as follows.

$$\boldsymbol{x}^{(\ell)}_{(1\ldots k-j+1), N^{(\ell)}_{(1\ldots j-1)}+1}, \ldots, \boldsymbol{x}^{(\ell)}_{(1\ldots k-j+1), N^{(\ell)}_{(1\ldots j)}} \overset{i.i.d.}{\sim} N_{p_{(1\ldots k-j+1)}}(\boldsymbol{\mu}^{(\ell)}_{(1\ldots k-j+1)}, \boldsymbol{\Sigma}^{(\ell)}_{(1\ldots k-j+1)(1\ldots k-j+1)}),$$

where we define $N^{(\ell)}_{(1\ldots j-1)} = 0$ when $j = 1$, and

$$\boldsymbol{\mu}^{(\ell)}_{(1\ldots k-j+1)} = \begin{pmatrix} \boldsymbol{\mu}^{(\ell)}_1 \\ \vdots \\ \boldsymbol{\mu}^{(\ell)}_{k-j+1} \end{pmatrix}, \boldsymbol{\mu}^{(\ell)}_{(1\ldots k)} = \boldsymbol{\mu}^{(\ell)},$$

$$\boldsymbol{\Sigma}^{(\ell)}_{(1\ldots k-j+1)(1\ldots k-j+1)} = \begin{pmatrix} \boldsymbol{\Sigma}^{(\ell)}_{11} & \cdots & \boldsymbol{\Sigma}^{(\ell)}_{1,k-j+1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}^{(\ell)}_{k-j+1,1} & \cdots & \boldsymbol{\Sigma}^{(\ell)}_{k-j+1,k-j+1} \end{pmatrix}, \boldsymbol{\Sigma}^{(\ell)}_{(1\ldots k)(1\ldots k)} = \boldsymbol{\Sigma}^{(\ell)},$$

$$\boldsymbol{x}^{(\ell)}_{(1\ldots k-j+1),i} = \begin{pmatrix} \boldsymbol{x}^{(\ell)}_{1i} \\ \vdots \\ \boldsymbol{x}^{(\ell)}_{k-j+1,i} \end{pmatrix}, \ i = N^{(\ell)}_{(1\ldots j-1)} + 1, \ldots, N^{(\ell)}_{(1\ldots j)},$$

$$N^{(\ell)}_{(1\ldots j)} = N^{(\ell)}_1 + \cdots + N^{(\ell)}_j, p_{(1\ldots j)} = p_1 + \cdots + p_j,$$

$$N^{(\ell)} = N^{(\ell)}_{(1\ldots k)}, p = p_{(1\ldots k)}, N^{(\ell)}_1 > p.$$

Furthermore, for $a, b, g = 1, \ldots, k-j+1; j = 1, \ldots, k, \ell = 1, \ldots, m$, we difine

$$\overline{\boldsymbol{x}}^{(\ell)}_{(j)g} = \frac{1}{N^{(\ell)}_j} \sum_{i=N^{(\ell)}_{(1\ldots j-1)}+1}^{N^{(\ell)}_{(1\ldots j)}} \boldsymbol{x}^{(\ell)}_{gi}, \overline{\boldsymbol{x}}^{(\ell)}_{[j]g} = \frac{1}{N^{(\ell)}_{(1\ldots j)}} (N^{(\ell)}_1 \overline{\boldsymbol{x}}^{(\ell)}_{(1)g} + \cdots + N^{(\ell)}_j \overline{\boldsymbol{x}}^{(\ell)}_{(j)g}),$$

$$\overline{\boldsymbol{x}}^{(\ell)}_{[j][g]} = \begin{pmatrix} \overline{\boldsymbol{x}}^{(\ell)}_{[j]1} \\ \vdots \\ \overline{\boldsymbol{x}}^{(\ell)}_{[j]g} \end{pmatrix}, \overline{\boldsymbol{x}}^{(\ell)}_{(j)[g]} = \begin{pmatrix} \overline{\boldsymbol{x}}^{(\ell)}_{(j)1} \\ \vdots \\ \overline{\boldsymbol{x}}^{(\ell)}_{(j)g} \end{pmatrix},$$

$$\begin{aligned} \boldsymbol{W}^{(\ell)}_{(j)} &= \sum_{i=N^{(\ell)}_{(1\ldots j-1)}+1}^{N^{(\ell)}_{(1\ldots j)}} (\boldsymbol{x}^{(\ell)}_{(1\ldots k-j+1),i} - \overline{\boldsymbol{x}}^{(\ell)}_{(j)[k-j+1]})(\boldsymbol{x}^{(\ell)}_{(1\ldots k-j+1),i} - \overline{\boldsymbol{x}}^{(\ell)}_{(j)[k-j+1]})' \\ &\quad + \frac{N^{(\ell)}_{(1\ldots j-1)} N^{(\ell)}_j}{N^{(\ell)}_{(1\ldots j)}} (\overline{\boldsymbol{x}}^{(\ell)}_{(j)[k-j+1]} - \overline{\boldsymbol{x}}^{(\ell)}_{[j-1][k-j+1]})(\overline{\boldsymbol{x}}^{(\ell)}_{(j)[k-j+1]} - \overline{\boldsymbol{x}}^{(\ell)}_{[j-1][k-j+1]})' \\ &= \begin{pmatrix} \boldsymbol{W}^{(\ell)}_{(j)11} & \cdots & \boldsymbol{W}^{(\ell)}_{(j)1,k-j+1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{W}^{(\ell)}_{(j)k-j+1,1} & \cdots & \boldsymbol{W}^{(\ell)}_{(j)k-j+1,k-j+1} \end{pmatrix}, \\ \boldsymbol{W}^{(\ell)}_{(k-j+1)} &= \left(\begin{array}{c|c} \boldsymbol{W}^{(\ell)}_{(k-j+1)(1\ldots j-1)(1\ldots j-1)} & \boldsymbol{W}^{(\ell)}_{(k-j+1)(1\ldots j-1)j} \\ \hline \boldsymbol{W}^{(\ell)}_{(k-j+1)j(1\ldots j-1)} & \boldsymbol{W}^{(\ell)}_{(k-j+1)jj} \end{array} \right), \ j = 2, \ldots, k, \\ \boldsymbol{W}^{(\ell)}_{[j]ab} &= \boldsymbol{W}^{(\ell)}_{(1)ab} + \cdots + \boldsymbol{W}^{(\ell)}_{(j)ab}, \\ \boldsymbol{W}^{(\ell)}_{[k-j+1]jj\cdot 1\ldots j-1} &= \boldsymbol{W}^{(\ell)}_{[k-j+1]jj} \\ &\quad - \boldsymbol{W}^{(\ell)}_{[k-j+1]j(1\ldots j-1)} \{\boldsymbol{W}^{(\ell)}_{[k-j+1](1\ldots j-1)(1\ldots j-1)}\}^{-1} \boldsymbol{W}^{(\ell)}_{[k-j+1](1\ldots j-1)j}, \\ j &= 2, \ldots, k. \end{aligned}$$

Subsequently, the LR of the test (1) under k-step monotone missing data is

$$\lambda_{(m)}^{(k)} = \left(\prod_{\ell=1}^{m} |\widehat{\mathbf{\Sigma}}_{11}^{(\ell)}|^{\frac{N^{(\ell)}}{2}} \prod_{j=2}^{k} |\widehat{\mathbf{\Sigma}}_{jj\cdot 1\ldots j-1}^{(\ell)}|^{\frac{N_{(1\ldots k-j+1)}^{(\ell)}}{2}} \right) (\widetilde{\sigma}^2)^{-\frac{1}{2}\sum_{j=1}^{k} N_{(1\ldots k-j+1)} p_j},$$

where $\widehat{\mathbf{\Sigma}}_{11}^{(\ell)}$ and $\widehat{\mathbf{\Sigma}}_{jj\cdot 1\ldots j-1}^{(\ell)}$ are the MLEs of $\mathbf{\Sigma}_{11}^{(\ell)}$ and $\mathbf{\Sigma}_{jj\cdot 1\ldots j-1}^{(\ell)} (= \mathbf{\Sigma}_{jj}^{(\ell)} - \mathbf{\Sigma}_{j(1\ldots j-1)}^{(\ell)} \mathbf{\Sigma}_{(1\ldots j-1)(1\ldots j-1)}^{(\ell)-1} \mathbf{\Sigma}_{(1\ldots j-1)j}^{(\ell)})$, $j = 2, \ldots, k$, respectively; $\widetilde{\sigma}^2$ is the MLE of σ^2 under H_0, and

$$\begin{aligned}
\widehat{\mathbf{\Sigma}}_{11}^{(\ell)} &= \frac{1}{N^{(\ell)}} \boldsymbol{W}_{[k]11}^{(\ell)}, \widehat{\mathbf{\Sigma}}_{jj\cdot 1\ldots j-1}^{(\ell)} = \frac{1}{N_{(1\ldots k-j+1)}^{(\ell)}} \boldsymbol{W}_{[k-j+1]jj\cdot 1\ldots j-1}^{(\ell)}, \ j = 2, \ldots, k, \\
\widetilde{\sigma}^2 &= \left(\sum_{j=1}^{k} N_{(1\ldots k-j+1)} p_j \right)^{-1} \sum_{\ell=1}^{m} \sum_{j=1}^{k} \mathrm{tr} \boldsymbol{W}_{[k-j+1]jj}^{(\ell)}, \\
N &= \sum_{\ell=1}^{m} N^{(\ell)}, N_j = \sum_{\ell=1}^{m} N_j^{(\ell)}, N_{(1\ldots j)} = N_1 + \cdots + N_j, \ j = 1, \ldots, k.
\end{aligned}$$

Furthermore, the LR $\lambda_{(m)}^{(k)}$ can be written as

$$\lambda_{(m)}^{(k)} = \frac{\prod_{\ell=1}^{m} \left| \frac{1}{N^{(\ell)}} \boldsymbol{W}_{[k]11}^{(\ell)} \right|^{\frac{N^{(\ell)}}{2}} \prod_{j=2}^{k} \left| \frac{1}{N_{(1\ldots k-j+1)}^{(\ell)}} \boldsymbol{W}_{[k-j+1]jj\cdot 1\ldots j-1}^{(\ell)} \right|^{\frac{N_{(1\ldots k-j+1)}^{(\ell)}}{2}}}{\left\{ \left(\sum_{j=1}^{k} N_{(1\ldots k-j+1)} p_j \right)^{-1} \sum_{\ell=1}^{m} \sum_{j=1}^{k} \mathrm{tr} \boldsymbol{W}_{[k-j+1]jj}^{(\ell)} \right\}^{\frac{1}{2}\sum_{j=1}^{k} N_{(1\ldots k-j+1)} p_j}}.$$

3 Main results

In this section, we derive an asymptotic expansion of the LRT statistic and modified LRT statistic for the the null distribution using the distribution in Box (1949), and provide an asymptotic expansion of the upper percentiles when H_0 holds. Finally, we provide some approximate upper percentiles.

3.1 Two-step case

Extending Sato, Yagi, and Seo (2025) to a multi-sample problem, we obtain the following theorem.

Theorem 1

When H_0 holds and $\gamma_1 = N_1/N \to \delta_1 \in (0, 1]$ $(N_1 \to \infty, N \to \infty)$, $\gamma_1^{(\ell)} = N_1^{(\ell)}/N \to$

$\delta_1^{(\ell)} \in (0,1]$ $(N_1^{(\ell)} \to \infty, N \to \infty)$,$\gamma_N^{(\ell)} = N^{(\ell)}/N \to \delta_N^{(\ell)} \in (0,1]$ $(N^{(\ell)} \to \infty, N \to \infty)$ *for* $\ell = 1, \ldots, m$*, the distribution function of the LRT statistic can be expanded as*

$$
\begin{aligned}
\Pr(-2\log\lambda_{(m)} \le x) = G_f(x) &+ \frac{\beta_{(2,m)}}{N}\{G_{f+2}(x) - G_f(x)\} \\
&+ \frac{\gamma_{(2,m)}}{N^2}\{G_{f+4}(x) - G_f(x)\} + O(N^{-3}),
\end{aligned} \tag{2}
$$

where $G_f(x)$ *is the distribution function of the* χ^2 *distribution with* f *degrees of freedom and*

$$
\begin{aligned}
f &= \frac{1}{2}(mp^2 + mp - 2), \\
\beta_{(2,m)} &= \frac{1}{24}\left[\left(\sum_{\ell=1}^{m}\frac{1}{\gamma_N^{(\ell)}}\right)p_1(2p_1^2 + 9p_1 + 11)\right. \\
&\quad + \left(\sum_{\ell=1}^{m}\frac{1}{\gamma_1^{(\ell)}}\right)\{p_2(2p_2^2 + 9p_2 + 11) + 6p_1p_2(p+3)\} \\
&\quad \left. - \frac{2}{p_1 + \gamma_1 p_2}(3m^2p^2 + 6mp + 2)\right], \\
\gamma_{(2,m)} &= \frac{1}{48}\left[\left(\sum_{\ell=1}^{m}\frac{1}{\gamma_N^{(\ell)2}}\right)p_1(p_1+1)(p_1+2)(p_1+3)\right. \\
&\quad + \left(\sum_{\ell=1}^{m}\frac{1}{\gamma_1^{(\ell)2}}\right)\Big(p_2(p_2+1)(p_2+2)(p_2+3) \\
&\quad + 2p_1p_2\{(p_2+1)(2p+p_1+7) + 2(p_1+1)(p_1+2)\}\Big) \\
&\quad \left. - \frac{4}{(p_1+\gamma_1p_2)^2}mp(mp+1)(mp+2)\right],
\end{aligned}
$$

and the distribution function of the modified LRT statistic can be expressed as follows.

$$
\Pr(-2\rho\log\lambda_{(m)} \le x) = G_f(x) + \frac{\gamma^*_{(2,m)}}{M^2}\{G_{f+4}(x) - G_f(x)\} + O(M^{-3}), \tag{3}
$$

where $M = \rho N$ *and*

$$
\begin{aligned}
\rho &= 1 - \frac{4}{(mp^2 + mp - 2)N}\beta_{(2,m)}, \\
\gamma^*_{(2,m)} &= -\frac{2}{mp^2 + mp - 2}\beta^2_{(2,m)} + \gamma_{(2,m)}.
\end{aligned}
$$

For the proof of **Theorem 1**, see Appendix.A.

Subsequently, we derive an asymptotic expansion of the upper percentiles. From (2) and (3), when H_0 holds, the upper 100α percentiles of $-2\log\lambda_{(m)}$ and $-2\rho\log\lambda_{(m)}$ can be expanded as follows:

$$
\begin{aligned}
q_{(2,m)}(\alpha) &= \chi_f^2(\alpha) + \frac{1}{N}\frac{2\beta_{(2,m)}}{f}\chi_f^2(\alpha) + \frac{1}{N^2}\frac{2\gamma_{(2,m)}}{f(f+2)}\chi_f^2(\alpha)\{\chi_f^2(\alpha)+f+2\} + O(N^{-3}),\\
q^*_{(2,m)}(\alpha) &= \chi_f^2(\alpha) + \frac{1}{M^2}\frac{2\gamma^*_{(2,m)}}{f(f+2)}\chi_f^2(\alpha)\{\chi_f^2(\alpha)+f+2\} + O(M^{-3}),
\end{aligned}
$$

respectively, where $\chi_f^2(\alpha)$ is the upper 100α percentile of the χ^2 distribution with f degrees of freedom.

3.2 k-step case ($k \geq 2$)

In the general k-step case, we have the following theorem.

Theorem 2

For $i = 1, \ldots, k$, $j = 2, \ldots, k$, $\ell = 1, \ldots, m$, let $\gamma_i = N_i/N, \gamma_i^{(\ell)} = N_i^{(\ell)}/N$ and

$$
\begin{aligned}
\gamma_{(1\ldots k-j+1)} &= \gamma_1 + \cdots + \gamma_{k-j+1}\\
&= \frac{N_{(1\ldots k-j+1)}}{N} \to \delta_{k-j+1} \in (0,1]\ (N_{(1\ldots k-j+1)} \to \infty, N \to \infty),\\
\gamma_{(1\ldots k)} &= 1,\\
\gamma^{(\ell)}_{(1\ldots k-j+1)} &= \gamma^{(\ell)}_1 + \cdots + \gamma^{(\ell)}_{k-j+1}\\
&= \frac{N^{(\ell)}_{(1\ldots k-j+1)}}{N} \to \delta^{(\ell)}_{k-j+1} \in (0,1]\ (N^{(\ell)}_{(1\ldots k-j+1)} \to \infty, N \to \infty),\\
\gamma^{(\ell)}_{(1\ldots k)} &= \frac{N^{(\ell)}}{N} \to \delta^{(\ell)}_N \in (0,1]\ (N^{(\ell)} \to \infty, N \to \infty).
\end{aligned}
$$

When H_0 holds, the distribution functions of the LRT statistic and modified LRT statistic can be expanded as

$$
\begin{aligned}
\Pr(-2\log\lambda^{(k)}_{(m)} \leq x) &= G_f(x) + \frac{\beta_{(k,m)}}{N}\{G_{f+2}(x) - G_f(x)\}\\
&\quad + \frac{\gamma_{(k,m)}}{N^2}\{G_{f+4}(x) - G_f(x)\} + O(N^{-3}), \qquad (4)\\
\Pr(-2\rho\log\lambda^{(k)}_{(m)} \leq x) &= G_f(x) + \frac{\gamma^*_{(k,m)}}{M^2}\{G_{f+4}(x) - G_f(x)\} + O(M^{-3}), \qquad (5)
\end{aligned}
$$

respectively, where we define $p_{(1\ldots j-1)} = 0$ when $j = 1$, and

$$
f = \frac{1}{2}(mp^2 + mp - 2),
$$

$$
\begin{aligned}
M &= \rho N, \\
\beta_{(k,m)} &= \frac{1}{24}\Bigg[\sum_{\ell=1}^{m}\sum_{j=1}^{k}\frac{1}{\gamma^{(\ell)}_{(1\ldots k-j+1)}}\left\{p_j(2p_j^2+9p_j+11)+6p_{(1\ldots j-1)}p_j(p_{(1\ldots j)}+3)\right\} \\
&\quad -\frac{2}{\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j}(3m^2p^2+6mp+2)\Bigg], \\
\gamma_{(k,m)} &= \frac{1}{48}\Bigg[\sum_{\ell=1}^{m}\sum_{j=1}^{k}\frac{1}{\gamma^{(\ell)2}_{(1\ldots k-j+1)}}\Big(p_j(p_j+1)(p_j+2)(p_j+3) \\
&\quad +2p_{(1\ldots j-1)}p_j\big\{(p_j+1)(2p_{(1\ldots j)}+p_{(1\ldots j-1)}+7)+2(p_{(1\ldots j-1)}+1)(p_{(1\ldots j-1)}+2)\big\}\Big) \\
&\quad -\frac{4}{\left(\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right)^2}mp(mp+1)(mp+2)\Bigg], \\
\rho &= 1-\frac{4}{(mp^2+mp-2)N}\beta_{(k,m)}, \\
\gamma^*_{(k,m)} &= -\frac{2}{mp^2+mp-2}\beta^2_{(k,m)}+\gamma_{(k,m)}.
\end{aligned}
$$

For the proof of **Theorem 2**, see Appendix.B. From (4) and (5), when H_0 holds, the upper 100α percentiles of $-2\log\lambda^{(k)}_{(m)}$ and $-2\rho\log\lambda^{(k)}_{(m)}$ are, respectively,

$$
\begin{aligned}
q_{(k,m)}(\alpha) &= \chi^2_f(\alpha)+\frac{1}{N}\frac{2\beta_{(k,m)}}{f}\chi^2_f(\alpha)+\frac{1}{N^2}\frac{2\gamma_{(k,m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}+O(N^{-3}), \\
q^*_{(k,m)}(\alpha) &= \chi^2_f(\alpha)+\frac{1}{M^2}\frac{2\gamma^*_{(k,m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}+O(M^{-3}).
\end{aligned}
$$

Finally, we propose $q_1(\alpha), q_2(\alpha), q_3(\alpha)$ for the approximate upper 100α percentiles of $-2\log\lambda^{(k)}_{(m)}$ and $q_1(\alpha), q^\dagger(\alpha)$ for the approximate upper 100α percentiles of $-2\rho\log\lambda^{(k)}_{(m)}$, where

$$
q_1(\alpha) = \chi^2_f(\alpha), \tag{6}
$$

$$
q_2(\alpha) = \chi^2_f(\alpha)+\frac{1}{N}\frac{2\beta_{(k,m)}}{f}\chi^2_f(\alpha), \tag{7}
$$

$$
q_3(\alpha) = \chi^2_f(\alpha)+\frac{1}{N}\frac{2\beta_{(k,m)}}{f}\chi^2_f(\alpha)+\frac{1}{N^2}\frac{2\gamma_{(k,m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}, \tag{8}
$$

$$q^{\dagger}(\alpha) = \chi_f^2(\alpha) + \frac{1}{M^2}\frac{2\gamma_{(k,m)}^*}{f(f+2)}\chi_f^2(\alpha)\{\chi_f^2(\alpha) + f + 2\}. \tag{9}$$

Note that $q_1(\alpha), q_2(\alpha), q_3(\alpha)$ split $q_{(k,m)}(\alpha)$, and $q_1(\alpha), q^{\dagger}(\alpha)$ split $q_{(k,m)}^*(\alpha)$, respectively.

4 Estimators of σ^2

In this section, we provide the MLE of σ^2 under H_0 and propose an unbiased estimator of σ^2. Furthermore, we evaluate which estimator is ideal using the MSE.

4.1 Two-step case

The MLE of σ^2 under H_0 can be written as

$$\widetilde{\sigma}^2 = \frac{1}{N(p_1 + \gamma_1 p_2)} \sum_{\ell=1}^{m} \left(\mathrm{tr}(\boldsymbol{W}_{(1)11}^{(\ell)} + \boldsymbol{W}_{(2)11}^{(\ell)}) + \mathrm{tr}\boldsymbol{W}_{(1)22}^{(\ell)} \right).$$

The expectation of $\widetilde{\sigma}^2$ is

$$\mathrm{E}[\widetilde{\sigma}^2] = \left(1 - \frac{mp}{N(p_1 + \gamma_1 p_2)}\right)\sigma^2,$$

where $N(p_1 + \gamma_1 p_2) - mp > 0$. Therefore, we propose an unbiased estimator of σ^2 as follows:

$$\widetilde{\sigma}_{\mathrm{U}}^2 = \left(1 - \frac{mp}{N(p_1 + \gamma_1 p_2)}\right)^{-1}\widetilde{\sigma}^2.$$

4.2 k-step case ($k \geq 2$)

The MLE of σ^2 under H_0 can be written as

$$\widetilde{\sigma}^2 = \left(N\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right)^{-1} \sum_{\ell=1}^{m}\sum_{j=1}^{k}\mathrm{tr}\boldsymbol{W}_{[k-j+1]jj}^{(\ell)}.$$

The expectation of $\widetilde{\sigma}^2$ can be expressed as

$$\mathrm{E}[\widetilde{\sigma}^2] = \left\{1 - mp\left(N\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right)^{-1}\right\}\sigma^2,$$

where $N(\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j) - mp > 0$. Therefore, we propose an unbiased estimator of σ^2 as follows:

$$\widetilde{\sigma}_{\mathrm{U}}^2 = \left\{1 - mp\left(N\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right)^{-1}\right\}^{-1}\widetilde{\sigma}^2.$$

The MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ are, respectively,

$$\mathrm{MSE}[\widetilde{\sigma}^2] = \frac{2\left\{N\left(\sum_{j=1}^{k}\gamma_{(1\dots k-j+1)}p_j\right) - mp\right\} + m^2p^2}{\left(N\sum_{j=1}^{k}\gamma_{(1\dots k-j+1)}p_j\right)^2}\sigma^4,$$

$$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}] = \frac{2}{N\left(\sum_{j=1}^{k}\gamma_{(1\dots k-j+1)}p_j\right) - mp}\sigma^4.$$

Considering the relationship between $\mathrm{MSE}[\widetilde{\sigma}^2]$ and $\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$ in terms of the magnitude, we obtain the following.

$$\begin{cases} \mathrm{MSE}[\widetilde{\sigma}^2] \leq \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}] & \text{if } mp \leq 4, \\ \mathrm{MSE}[\widetilde{\sigma}^2] \geq \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}] & \text{if } mp > 4, \end{cases} \tag{10}$$

where the equality relations are satisfied at $N_{(1\dots k-j+1)} \to \infty, N \to \infty$. For details on these inequalities, see Appendix.C.

Therefore, $\widetilde{\sigma}^2$ is a better estimator than $\widetilde{\sigma}^2_{\mathrm{U}}$ when $mp \leq 4$, and $\widetilde{\sigma}^2_{\mathrm{U}}$ is a better estimator than $\widetilde{\sigma}^2$ when $mp > 4$. Furthermore, when $N_{(1\dots k-j+1)} \to \infty, N \to \infty$, $\widetilde{\sigma}^2_{\mathrm{U}}$ approaches $\widetilde{\sigma}^2$ owing to the convergence property of $\gamma_{(1\dots k-j+1)}$ in **Theorem 2**.

5 Affine transformation

In this section, we consider the affine transformation of the LR and derive the following theorem.

Theorem 3

The LR of the sphericity test (1) *is invariant under H_0 for the affine transformation.*

This theorem extends Sato, Yagi, and Seo (2025). We prove **Theorem 3** under two-step and k-step monotone missing data.

5.1 Two-step case

In the ℓth population ($\ell = 1, \dots, m$), let $\boldsymbol{\Lambda}^{(\ell)}_{11}, \boldsymbol{\Lambda}^{(\ell)}_{22}$ be a $p_1 \times p_1, p_2 \times p_2$ positive definite symmetric matrices, $\boldsymbol{\Lambda}^{(\ell)}_{21}$ a $p_2 \times p_1$ matrix, and let $\boldsymbol{\nu}^{(\ell)}_1, \boldsymbol{\nu}^{(\ell)}_2$ a $p_1 \times 1, p_2 \times 1$ vectors, respectively. Subsequently, we consider the following affine transformation:

$$\begin{pmatrix} \boldsymbol{x}_{1j}^{(\ell)*} \\ \boldsymbol{x}_{2j}^{(\ell)*} \end{pmatrix} = \boldsymbol{\Lambda}^{(\ell)} \boldsymbol{C}^{(\ell)} \begin{pmatrix} \boldsymbol{x}_{1j}^{(\ell)} \\ \boldsymbol{x}_{2j}^{(\ell)} \end{pmatrix} + \boldsymbol{\nu}^{(\ell)},\ j = 1, \ldots, N_1^{(\ell)},$$
$$\boldsymbol{x}_{1j}^{(\ell)*} = \boldsymbol{\Lambda}_{11}^{(\ell)} \boldsymbol{x}_{1j}^{(\ell)} + \boldsymbol{\nu}_1^{(\ell)},\ j = N_1^{(\ell)} + 1, \ldots, N^{(\ell)},$$

where

$$\boldsymbol{\Lambda}^{(\ell)} = \begin{pmatrix} \boldsymbol{\Lambda}_{11}^{(\ell)} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{\Lambda}_{22}^{(\ell)} \end{pmatrix}, \boldsymbol{C}^{(\ell)} = \begin{pmatrix} \boldsymbol{I}_{p_1} & \boldsymbol{O} \\ \boldsymbol{\Lambda}_{21}^{(\ell)} & \boldsymbol{I}_{p_2} \end{pmatrix}, \boldsymbol{\nu}^{(\ell)} = \begin{pmatrix} \boldsymbol{\nu}_1^{(\ell)} \\ \boldsymbol{\nu}_2^{(\ell)} \end{pmatrix}.$$

The LR $\lambda_{(m)}^*$ after the affine transformation is

$$\lambda_{(m)}^* = \frac{\prod_{\ell=1}^{m} \left| \frac{1}{N^{(\ell)}} \left(\boldsymbol{W}_{(1)11}^{(\ell)*} + \boldsymbol{W}_{(2)11}^{(\ell)*} \right) \right|^{\frac{N^{(\ell)}}{2}} \left| \frac{1}{N_1^{(\ell)}} \boldsymbol{W}_{(1)22\cdot 1}^{(\ell)*} \right|^{\frac{N_1^{(\ell)}}{2}}}{\left\{ \frac{1}{Np_1 + N_1 p_2} \sum_{\ell=1}^{m} \left(\mathrm{tr}(\boldsymbol{W}_{(1)11}^{(\ell)*} + \boldsymbol{W}_{(2)11}^{(\ell)*}) + \mathrm{tr} \boldsymbol{W}_{(1)22}^{(\ell)*} \right) \right\}^{\frac{Np_1 + N_1 p_2}{2}}},$$

where

$$\overline{\boldsymbol{x}}_{(1)1}^{(\ell)*} = \frac{1}{N_1^{(\ell)}} \sum_{j=1}^{N_1^{(\ell)}} \boldsymbol{x}_{1j}^{(\ell)*},\ \overline{\boldsymbol{x}}_{(1)2}^{(\ell)*} = \frac{1}{N_1^{(\ell)}} \sum_{j=1}^{N_1^{(\ell)}} \boldsymbol{x}_{2j}^{(\ell)*},\ \overline{\boldsymbol{x}}_{(2)1}^{(\ell)*} = \frac{1}{N_2^{(\ell)}} \sum_{j=N_1^{(\ell)}+1}^{N^{(\ell)}} \boldsymbol{x}_{1j}^{(\ell)*},$$

$$\boldsymbol{W}_{(1)11}^{(\ell)*} = \sum_{j=1}^{N_1^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(1)1}^{(\ell)*})(\boldsymbol{x}_{1j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(1)1}^{(\ell)*})',$$

$$\boldsymbol{W}_{(1)12}^{(\ell)*} = (\boldsymbol{W}_{(1)21}^{(\ell)*})' = \sum_{j=1}^{N_1^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(1)1}^{(\ell)*})(\boldsymbol{x}_{2j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(1)2}^{(\ell)*})',$$

$$\boldsymbol{W}_{(1)22}^{(\ell)*} = \sum_{j=1}^{N_1^{(\ell)}} (\boldsymbol{x}_{2j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(1)2}^{(\ell)*})(\boldsymbol{x}_{2j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(1)2}^{(\ell)*})',$$

$$\boldsymbol{W}_{(2)11}^{(\ell)*} = \sum_{j=N_1^{(\ell)}+1}^{N^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)*})(\boldsymbol{x}_{1j}^{(\ell)*} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)*})' + \frac{N_1^{(\ell)} N_2^{(\ell)}}{N^{(\ell)}} (\overline{\boldsymbol{x}}_{(1)1}^{(\ell)*} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)*})(\overline{\boldsymbol{x}}_{(1)1}^{(\ell)*} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)*})',$$

$$\boldsymbol{W}_{(1)22\cdot 1}^{(\ell)*} = \boldsymbol{W}_{(1)22}^{(\ell)*} - \boldsymbol{W}_{(1)21}^{(\ell)*} (\boldsymbol{W}_{(1)11}^{(\ell)*})^{-1} \boldsymbol{W}_{(1)12}^{(\ell)*}.$$

Subsequently, we can write

$$\boldsymbol{W}_{(1)11}^{(\ell)*} + \boldsymbol{W}_{(2)11}^{(\ell)*} = \boldsymbol{\Lambda}_{11}^{(\ell)} (\boldsymbol{W}_{(1)11}^{(\ell)} + \boldsymbol{W}_{(2)11}^{(\ell)}) \boldsymbol{\Lambda}_{11}^{(\ell)},$$

$$
\begin{aligned}
\boldsymbol{W}_{(1)22}^{(\ell)*} &= \boldsymbol{\Lambda}_{22}^{(\ell)}\boldsymbol{\Lambda}_{21}^{(\ell)}\boldsymbol{W}_{(1)11}^{(\ell)}\boldsymbol{\Lambda}_{21}^{(\ell)'}\boldsymbol{\Lambda}_{22}^{(\ell)} + \boldsymbol{\Lambda}_{22}^{(\ell)}\boldsymbol{\Lambda}_{21}^{(\ell)}\boldsymbol{W}_{(1)12}^{(\ell)}\boldsymbol{\Lambda}_{22}^{(\ell)} \\
&\quad + \boldsymbol{\Lambda}_{22}^{(\ell)}\boldsymbol{W}_{(1)21}^{(\ell)}\boldsymbol{\Lambda}_{21}^{(\ell)'}\boldsymbol{\Lambda}_{22}^{(\ell)} + \boldsymbol{\Lambda}_{22}^{(\ell)}\boldsymbol{W}_{(1)22}^{(\ell)}\boldsymbol{\Lambda}_{22}^{(\ell)}, \\
\boldsymbol{W}_{(1)22\cdot1}^{(\ell)*} &= \boldsymbol{\Lambda}_{22}^{(\ell)}\boldsymbol{W}_{(1)22\cdot1}^{(\ell)}\boldsymbol{\Lambda}_{22}^{(\ell)}.
\end{aligned}
$$

We consider the mean vector $\boldsymbol{\mu}^{(\ell)^*}$ and variance-covariance matrix $\boldsymbol{\Sigma}^{(\ell)^*}$ after the affine transformation, where

$$
\boldsymbol{\mu}^{(\ell)^*} = \boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)}\boldsymbol{\mu}^{(\ell)} + \boldsymbol{\nu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)^*} = \boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)}\boldsymbol{\Sigma}^{(\ell)}(\boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)})'.
$$

Let $\boldsymbol{\nu}^{(\ell)}$ and $\boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)}$ be

$$
\boldsymbol{\nu}^{(\ell)} = -\boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)}\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)} = \begin{pmatrix} \boldsymbol{\Lambda}_{11}^{(\ell)} & \boldsymbol{O} \\ \boldsymbol{\Lambda}_{22}^{(\ell)}\boldsymbol{\Lambda}_{21}^{(\ell)} & \boldsymbol{\Lambda}_{22}^{(\ell)} \end{pmatrix} = (\sigma^2)^{\frac{1}{2}} \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(\ell)^{-\frac{1}{2}}} & \boldsymbol{O} \\ -\boldsymbol{\Sigma}_{22\cdot1}^{(\ell)^{-\frac{1}{2}}}\boldsymbol{\Sigma}_{21}^{(\ell)}\boldsymbol{\Sigma}_{11}^{(\ell)^{-1}} & \boldsymbol{\Sigma}_{22\cdot1}^{(\ell)^{-\frac{1}{2}}} \end{pmatrix},
$$

respectively. Thereafter, $\boldsymbol{\mu}^{(\ell)^*} = \boldsymbol{0}$ and $\boldsymbol{\Sigma}^{(\ell)^*} = \sigma^2\boldsymbol{I}_p$. Therefore, consider

$$
\boldsymbol{\mu}^{(\ell)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(\ell)} \\ \boldsymbol{\mu}_2^{(\ell)} \end{pmatrix}, \boldsymbol{\Sigma}^{(\ell)} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{(\ell)} & \boldsymbol{\Sigma}_{12}^{(\ell)} \\ \boldsymbol{\Sigma}_{21}^{(\ell)} & \boldsymbol{\Sigma}_{22}^{(\ell)} \end{pmatrix},
$$

we can assume $\boldsymbol{\mu}^{(\ell)} = \boldsymbol{0}$ and $\boldsymbol{\Sigma}^{(\ell)} = \sigma^2\boldsymbol{I}_p$ without loss of generality in deriving the null distribution of $-2\log\lambda_{(m)}$ and $-2\rho\log\lambda_{(m)}$ by transforming using $\boldsymbol{\Lambda}^{(\ell)}\boldsymbol{C}^{(\ell)}$. Because $\boldsymbol{\Lambda}_{11}^{(\ell)} = \boldsymbol{I}_{p_1}, \boldsymbol{\Lambda}_{22}^{(\ell)} = \boldsymbol{I}_{p_2}, \boldsymbol{\Lambda}_{21}^{(\ell)} = \boldsymbol{O}$ under H_0,

$$
\begin{aligned}
\boldsymbol{W}_{(1)11}^{(\ell)^*} + \boldsymbol{W}_{(2)11}^{(\ell)^*} &= \boldsymbol{W}_{(1)11}^{(\ell)} + \boldsymbol{W}_{(2)11}^{(\ell)}, \\
\boldsymbol{W}_{(1)22}^{(\ell)^*} &= \boldsymbol{W}_{(1)22}^{(\ell)}, \\
\boldsymbol{W}_{(1)22\cdot1}^{(\ell)^*} &= \boldsymbol{W}_{(1)22\cdot1}^{(\ell)}.
\end{aligned}
$$

Therefore, $\lambda_{(m)}^* = \lambda_{(m)}$.

From the above, the LR $\lambda_{(m)}$ is invariant under H_0 for the affine transformation, and this completes the proof.

5.2 k-step case ($k \geq 2$)

In the ℓth population ($\ell = 1, \ldots, m$), let $\boldsymbol{\Lambda}_{jj}^{(\ell)}$ be $p_j \times p_j$ positive definite symmetric matrices for $j = 1, \ldots, k$, $\boldsymbol{\Lambda}_{j(1\ldots j-1)}^{(\ell)}$ be $p_j \times p_{(1\ldots j-1)}$ matrices for $j = 2, \ldots, k$, and let $\boldsymbol{\nu}_j^{(\ell)}$ be $p_j \times 1$ vectors for $j = 1, \ldots, k$, respectively. For $j = 1, \ldots, k$, we consider the following affine transformation.

$$\boldsymbol{x}^{(\ell)*}_{(1\ldots k-j+1),i} = \boldsymbol{\Lambda}^{(\ell)}_{(j)}\boldsymbol{C}^{(\ell)}_{(j)}\boldsymbol{x}^{(\ell)}_{(1\ldots k-j+1),i} + \boldsymbol{\nu}^{(\ell)}_{(j)},\ i = N^{(\ell)}_{(1\ldots j-1)} + 1, \ldots, N^{(\ell)}_{(1\ldots j)},$$

where $\boldsymbol{C}^{(\ell)}_{(k)} = \boldsymbol{I}_{p_1}$ and

$$\boldsymbol{\Lambda}^{(\ell)}_{(j)} = \begin{pmatrix} \boldsymbol{\Lambda}^{(\ell)}_{11} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \ddots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{\Lambda}^{(\ell)}_{k-j+1,k-j+1} \end{pmatrix}, \boldsymbol{C}^{(\ell)}_{(j)} = \left(\begin{array}{ccc|c} \boldsymbol{I}_{p_1} & \boldsymbol{O} & \boldsymbol{O} & \\ \boldsymbol{\Lambda}^{(\ell)}_{21} & \boldsymbol{I}_{p_2} & \boldsymbol{O} & \boldsymbol{O} \\ \vdots & & \ddots & \\ \hline \multicolumn{3}{c|}{\boldsymbol{\Lambda}^{(\ell)}_{k-j+1(1\ldots k-j)}} & \boldsymbol{I}_{p_{k-j+1}} \end{array}\right),$$

$$\boldsymbol{\nu}^{(\ell)}_{(j)} = \begin{pmatrix} \boldsymbol{\nu}^{(\ell)}_{1} \\ \vdots \\ \boldsymbol{\nu}^{(\ell)}_{k-j+1} \end{pmatrix}.$$

The LR $\lambda^{(k)*}_{(m)}$ after the affine transformation is

$$\lambda^{(k)*}_{(m)} = \frac{\prod_{\ell=1}^{m} \left| \frac{1}{N^{(\ell)}} \boldsymbol{W}^{(\ell)*}_{[k]11} \right|^{\frac{N^{(\ell)}}{2}} \prod_{j=2}^{k} \left| \frac{1}{N^{(\ell)}_{(1\ldots k-j+1)}} \boldsymbol{W}^{(\ell)*}_{[k-j+1]jj\cdot 1\ldots j-1} \right|^{\frac{N^{(\ell)}_{(1\ldots k-j+1)}}{2}}}{\left\{ \left(\sum_{j=1}^{k} N_{(1\ldots k-j+1)} p_j \right)^{-1} \sum_{\ell=1}^{m} \sum_{j=1}^{k} \mathrm{tr} \boldsymbol{W}^{(\ell)*}_{[k-j+1]jj} \right\}^{\frac{1}{2}\sum_{j=1}^{k} N_{(1\ldots k-j+1)} p_j}},$$

where for $a, b, g = 1, \ldots, k-j+1; j = 1, \ldots, k, \ell = 1, \ldots, m,$

$$\overline{\boldsymbol{x}}^{(\ell)*}_{(j)g} = \frac{1}{N^{(\ell)}_j} \sum_{i=N^{(\ell)}_{(1\ldots j-1)}+1}^{N^{(\ell)}_{(1\ldots j)}} \boldsymbol{x}^{(\ell)*}_{gi}, \overline{\boldsymbol{x}}^{(\ell)*}_{[j]g} = \frac{1}{N^{(\ell)}_{(1\ldots j)}} (N^{(\ell)}_1 \overline{\boldsymbol{x}}^{(\ell)*}_{(1)g} + \cdots + N^{(\ell)}_j \overline{\boldsymbol{x}}^{(\ell)*}_{(j)g}),$$

$$\overline{\boldsymbol{x}}^{(\ell)*}_{[j][g]} = \begin{pmatrix} \overline{\boldsymbol{x}}^{(\ell)*}_{[j]1} \\ \vdots \\ \overline{\boldsymbol{x}}^{(\ell)*}_{[j]g} \end{pmatrix}, \overline{\boldsymbol{x}}^{(\ell)*}_{(j)[g]} = \begin{pmatrix} \overline{\boldsymbol{x}}^{(\ell)*}_{(j)1} \\ \vdots \\ \overline{\boldsymbol{x}}^{(\ell)*}_{(j)g} \end{pmatrix},$$

$$\begin{aligned} \boldsymbol{W}^{(\ell)*}_{(j)} &= \sum_{i=N^{(\ell)}_{(1\ldots j-1)}+1}^{N^{(\ell)}_{(1\ldots j)}} (\boldsymbol{x}^{(\ell)*}_{(1\ldots k-j+1),i} - \overline{\boldsymbol{x}}^{(\ell)*}_{(j)[k-j+1]})(\boldsymbol{x}^{(\ell)*}_{(1\ldots k-j+1),i} - \overline{\boldsymbol{x}}^{(\ell)*}_{(j)[k-j+1]})' \\ &\quad + \frac{N^{(\ell)}_{(1\ldots j-1)} N^{(\ell)}_j}{N^{(\ell)}_{(1\ldots j)}} (\overline{\boldsymbol{x}}^{(\ell)*}_{(j)[k-j+1]} - \overline{\boldsymbol{x}}^{(\ell)*}_{[j-1][k-j+1]})(\overline{\boldsymbol{x}}^{(\ell)*}_{(j)[k-j+1]} - \overline{\boldsymbol{x}}^{(\ell)*}_{[j-1][k-j+1]})' \\ &= \begin{pmatrix} \boldsymbol{W}^{(\ell)*}_{(j)11} & \cdots & \boldsymbol{W}^{(\ell)*}_{(j)1,k-j+1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{W}^{(\ell)*}_{(j)k-j+1,1} & \cdots & \boldsymbol{W}^{(\ell)*}_{(j)k-j+1,k-j+1} \end{pmatrix}, \end{aligned}$$

$$\boldsymbol{W}^{(\ell)^*}_{(k-j+1)} = \left(\begin{array}{c|c} \boldsymbol{W}^{(\ell)^*}_{(k-j+1)(1\ldots j-1)(1\ldots j-1)} & \boldsymbol{W}^{(\ell)^*}_{(k-j+1)(1\ldots j-1)j} \\ \hline \boldsymbol{W}^{(\ell)^*}_{(k-j+1)j(1\ldots j-1)} & \boldsymbol{W}^{(\ell)^*}_{(k-j+1)jj} \end{array} \right),\ j = 2, \ldots, k,$$

$$\begin{aligned} \boldsymbol{W}^{(\ell)^*}_{[j]ab} &= \boldsymbol{W}^{(\ell)^*}_{(1)ab} + \cdots + \boldsymbol{W}^{(\ell)^*}_{(j)ab}, \\ \boldsymbol{W}^{(\ell)^*}_{[k-j+1]jj\cdot 1\ldots j-1} &= \boldsymbol{W}^{(\ell)^*}_{[k-j+1]jj} \\ &\quad - \boldsymbol{W}^{(\ell)^*}_{[k-j+1]j(1\ldots j-1)} \{\boldsymbol{W}^{(\ell)^*}_{[k-j+1](1\ldots j-1)(1\ldots j-1)}\}^{-1} \boldsymbol{W}^{(\ell)^*}_{[k-j+1](1\ldots j-1)j}, \\ & j = 2, \ldots, k. \end{aligned}$$

Subsequently, we can write

$$\begin{aligned} \boldsymbol{W}^{(\ell)^*}_{[k]11} &= \boldsymbol{\Lambda}^{(\ell)}_{11} \boldsymbol{W}^{(\ell)}_{[k]11} \boldsymbol{\Lambda}^{(\ell)}_{11}, \\ \boldsymbol{W}^{(\ell)^*}_{[k-j+1]jj} &= \boldsymbol{\Lambda}^{(\ell)}_{jj} \boldsymbol{\Lambda}^{(\ell)}_{j(1\ldots j-1)} \boldsymbol{W}^{(\ell)}_{[k-j+1](1\ldots j-1)(1\ldots j-1)} \boldsymbol{\Lambda}^{(\ell)'}_{j(1\ldots j-1)} \boldsymbol{\Lambda}^{(\ell)}_{jj} \\ &\quad + \boldsymbol{\Lambda}^{(\ell)}_{jj} \boldsymbol{\Lambda}^{(\ell)}_{j(1\ldots j-1)} \boldsymbol{W}^{(\ell)}_{[k-j+1](1\ldots j-1)j} \boldsymbol{\Lambda}^{(\ell)}_{jj} + \boldsymbol{\Lambda}^{(\ell)}_{jj} \boldsymbol{W}^{(\ell)}_{[k-j+1]j(1\ldots j-1)} \boldsymbol{\Lambda}^{(\ell)'}_{j(1\ldots j-1)} \boldsymbol{\Lambda}^{(\ell)}_{jj} \\ &\quad + \boldsymbol{\Lambda}^{(\ell)}_{jj} \boldsymbol{W}^{(\ell)}_{[k-j+1]jj} \boldsymbol{\Lambda}^{(\ell)}_{jj},\ j = 2, \ldots, k, \\ \boldsymbol{W}^{(\ell)^*}_{[k-j+1]jj\cdot 1\ldots j-1} &= \boldsymbol{\Lambda}^{(\ell)}_{jj} \boldsymbol{W}^{(\ell)}_{[k-j+1]jj\cdot 1\ldots j-1} \boldsymbol{\Lambda}^{(\ell)}_{jj},\ j = 2, \ldots, k. \end{aligned}$$

The mean vector $\boldsymbol{\mu}^{(\ell)^*}$ and variance-covariance matrix $\boldsymbol{\Sigma}^{(\ell)^*}$ after the affine transformation can be expressed

$$\boldsymbol{\mu}^{(\ell)^*} = \boldsymbol{\Lambda}^{(\ell)}_{(1)} \boldsymbol{C}^{(\ell)}_{(1)} \boldsymbol{\mu}^{(\ell)} + \boldsymbol{\nu}^{(\ell)}_{(1)}, \boldsymbol{\Sigma}^{(\ell)^*} = \boldsymbol{\Lambda}^{(\ell)}_{(1)} \boldsymbol{C}^{(\ell)}_{(1)} \boldsymbol{\Sigma}^{(\ell)} (\boldsymbol{\Lambda}^{(\ell)}_{(1)} \boldsymbol{C}^{(\ell)}_{(1)})',$$

respectively. Let $\boldsymbol{\nu}^{(\ell)}_{(1)}$ and $\boldsymbol{\Lambda}^{(\ell)}_{(1)} \boldsymbol{C}^{(\ell)}_{(1)}$ be

$$\begin{aligned} \boldsymbol{\nu}^{(\ell)}_{(1)} &= -\boldsymbol{\Lambda}^{(\ell)}_{(1)} \boldsymbol{C}^{(\ell)}_{(1)} \boldsymbol{\mu}^{(\ell)}, \\ \boldsymbol{\Lambda}^{(\ell)}_{(1)} \boldsymbol{C}^{(\ell)}_{(1)} &= \left(\begin{array}{ccc|c} \boldsymbol{\Lambda}^{(\ell)}_{11} & \boldsymbol{O} & \boldsymbol{O} & \\ \boldsymbol{\Lambda}^{(\ell)}_{22} \boldsymbol{\Lambda}^{(\ell)}_{21} & \boldsymbol{\Lambda}^{(\ell)}_{22} & \boldsymbol{O} & \boldsymbol{O} \\ \vdots & & \ddots & \\ \hline \multicolumn{3}{c|}{\boldsymbol{\Lambda}^{(\ell)}_{kk} \boldsymbol{\Lambda}^{(\ell)}_{k(1\ldots k-1)}} & \boldsymbol{\Lambda}^{(\ell)}_{kk} \end{array} \right) \\ &= (\sigma^2)^{\frac{1}{2}} \left(\begin{array}{ccc|c} \boldsymbol{\Sigma}^{(\ell)^{-\frac{1}{2}}}_{11} & \boldsymbol{O} & \boldsymbol{O} & \\ -\boldsymbol{\Sigma}^{(\ell)^{-\frac{1}{2}}}_{22\cdot 1} \boldsymbol{\Sigma}^{(\ell)}_{21} \boldsymbol{\Sigma}^{(\ell)^{-1}}_{11} & \boldsymbol{\Sigma}^{(\ell)^{-\frac{1}{2}}}_{22\cdot 1} & \boldsymbol{O} & \boldsymbol{O} \\ \vdots & & \ddots & \\ \hline \multicolumn{3}{c|}{-\boldsymbol{\Sigma}^{(\ell)^{-\frac{1}{2}}}_{kk\cdot 1\ldots k-1} \boldsymbol{\Sigma}^{(\ell)}_{k(1\ldots k-1)} \boldsymbol{\Sigma}^{(\ell)^{-1}}_{(1\ldots k-1)(1\ldots k-1)}} & \boldsymbol{\Sigma}^{(\ell)^{-\frac{1}{2}}}_{kk\cdot 1\ldots k-1} \end{array} \right), \end{aligned}$$

respectively, where

$$\boldsymbol{\Sigma}^{(\ell)}_{(1\ldots j)(1\ldots j)} = \left(\begin{array}{c|c} \boldsymbol{\Sigma}^{(\ell)}_{(1\ldots j-1)(1\ldots j-1)} & \boldsymbol{\Sigma}^{(\ell)}_{(1\ldots j-1)j} \\ \hline \boldsymbol{\Sigma}^{(\ell)}_{j(1\ldots j-1)} & \boldsymbol{\Sigma}^{(\ell)}_{jj} \end{array}\right),\ j = 2, \ldots, k.$$

Thereafter, $\boldsymbol{\mu}^{(\ell)^*} = \mathbf{0}$ and $\boldsymbol{\Sigma}^{(\ell)^*} = \sigma^2 \boldsymbol{I}_p$. Therefore, considering

$$\boldsymbol{\mu}^{(\ell)} = \boldsymbol{\mu}^{(\ell)}_{(1\ldots k)}, \boldsymbol{\Sigma}^{(\ell)} = \boldsymbol{\Sigma}^{(\ell)}_{(1\ldots k)(1\ldots k)},$$

we can assume $\boldsymbol{\mu}^{(\ell)} = \mathbf{0}$ and $\boldsymbol{\Sigma}^{(\ell)} = \sigma^2 \boldsymbol{I}_p$ without loss of generality in deriving the null distribution of $-2\log\lambda^{(k)}_{(m)}$ and $-2\rho\log\lambda^{(k)}_{(m)}$ by transforming using $\boldsymbol{\Lambda}^{(\ell)}_{(1)}\boldsymbol{C}^{(\ell)}_{(1)}$. As $\boldsymbol{\Lambda}^{(\ell)}_{jj} = \boldsymbol{I}_{p_j}, \boldsymbol{\Lambda}^{(\ell)}_{j(1\ldots j-1)} = \boldsymbol{O}$ under H_0, the Wishart matrices after the affine transformation can be expressed as follows:

$$\begin{aligned} \boldsymbol{W}^{(\ell)^*}_{[k]11} &= \boldsymbol{W}^{(\ell)}_{[k]11}, \\ \boldsymbol{W}^{(\ell)^*}_{[k-j+1]jj} &= \boldsymbol{W}^{(\ell)}_{[k-j+1]jj}, \\ \boldsymbol{W}^{(\ell)^*}_{[k-j+1]jj\cdot 1\ldots j-1} &= \boldsymbol{W}^{(\ell)}_{[k-j+1]jj\cdot 1\ldots j-1}. \end{aligned}$$

Therefore, $\lambda^{(k)^*}_{(m)} = \lambda^{(k)}_{(m)}$.

From the above, the LR $\lambda^{(k)}_{(m)}$ is invariant under H_0 for the affine transformation, completing the proof of **Theorem 3**.

6 Complete data

To compare the approximation accuracy of the approximate upper percentile under monotone missing data and complete data, we discuss the sphericity test (1) under complete data for a multi-sample problem. For the ℓth population ($\ell = 1, \ldots, m$), let $\boldsymbol{x}^{(\ell)}_1, \ldots, \boldsymbol{x}^{(\ell)}_{N^{(\ell)}}$ be independently distributed as $N_p(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)})$, and $N^{(\ell)} > p$. Subsequently, we have the following theorem using the distribution in Box (1949).

Theorem 4

When H_0 holds and $\gamma^{(\ell)}_n = n^{(\ell)}/n \to \delta^{(\ell)}_n \in (0, 1]$ $(n^{(\ell)} \to \infty, n \to \infty)$, the distribution functions of $-2\log\lambda^*_{c(m)}$ and $-2\rho\log\lambda^*_{c(m)}$ can be expanded as

$$\Pr(-2\log\lambda^*_{c(m)} \le x) = G_f(x) + \frac{\beta_{(m)}}{n}\left\{G_{f+2}(x) - G_f(x)\right\}$$

$$+\frac{\gamma_{(m)}}{n^2}\{G_{f+4}(x)-G_f(x)\}+O(n^{-3}), \tag{11}$$

$$\Pr(-2\rho\log\lambda^*_{c(m)}\le x)=G_f(x)+\frac{\gamma^*_{(m)}}{v^2}\{G_{f+4}(x)-G_f(x)\}+O(v^{-3}), \tag{12}$$

respectively, where $\lambda^*_{c(m)}$ is the LR of the sphericity test (1) with replacing $N^{(\ell)}$ by $n^{(\ell)}$ and N by n under complete data for a multi-sample problem, and

$$\begin{aligned}
\lambda^*_{c(m)} &= \frac{\prod_{\ell=1}^{m}\left|\frac{1}{n^{(\ell)}}\boldsymbol{W}^{(\ell)}\right|^{\frac{n^{(\ell)}}{2}}}{\left(\frac{1}{np}\sum_{\ell=1}^{m}\mathrm{tr}\boldsymbol{W}^{(\ell)}\right)^{\frac{np}{2}}},\ \boldsymbol{W}^{(\ell)}=\sum_{j=1}^{N^{(\ell)}}(\boldsymbol{x}_j^{(\ell)}-\overline{\boldsymbol{x}}^{(\ell)})(\boldsymbol{x}_j^{(\ell)}-\overline{\boldsymbol{x}}^{(\ell)})',\ \overline{\boldsymbol{x}}^{(\ell)}=\frac{1}{N^{(\ell)}}\sum_{j=1}^{N^{(\ell)}}\boldsymbol{x}_j^{(\ell)},\\
n^{(\ell)} &= N^{(\ell)}-1, N=\sum_{\ell=1}^{m}N^{(\ell)}, n=\sum_{\ell=1}^{m}n^{(\ell)},\\
f &= \frac{1}{2}(mp^2+mp-2),\\
v &= \rho n,\\
\beta_{(m)} &= \frac{1}{24p}\left\{\left(\sum_{\ell=1}^{m}\frac{1}{\gamma_n^{(\ell)}}\right)p^2(2p^2+3p-1)-4\right\},\\
\gamma_{(m)} &= \frac{1}{48}\left(\sum_{\ell=1}^{m}\frac{1}{\gamma_n^{(\ell)2}}\right)(p-1)p(p+1)(p+2),\\
\rho &= 1-\frac{4}{(mp^2+mp-2)n}\beta_{(m)},\\
\gamma^*_{(m)} &= -\frac{2}{mp^2+mp-2}\beta^2_{(m)}+\gamma_{(m)}.
\end{aligned}$$

The LR $\lambda^*_{c(m)}$ satisfies **Theorem 3**. The limiting null distributions for $-2\log\lambda^*_{c(m)}$ and $-2\rho\log\lambda^*_{c(m)}$ are provided by Mendoza (1980). From Eqs.(11) and (12), when H_0 holds, the upper 100α percentiles of $-2\log\lambda^*_{c(m)}$ and $-2\rho\log\lambda^*_{c(m)}$ can be asymptotically expanded as follows:

$$\begin{aligned}
q_{c(m)}(\alpha) &= \chi^2_f(\alpha)+\frac{1}{n}\frac{2\beta_{(m)}}{f}\chi^2_f(\alpha)+\frac{1}{n^2}\frac{2\gamma_{(m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}+O(n^{-3}),\\
q^*_{c(m)}(\alpha) &= \chi^2_f(\alpha)+\frac{1}{v^2}\frac{2\gamma^*_{(m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}+O(v^{-3}),
\end{aligned}$$

respectively. Finally, using the same method as in Section 3.2, we can provide the approximate upper 100α percentiles of $-2\log\lambda^*_{c(m)}$: $q_{c_1}(\alpha), q_{c_2}(\alpha), q_{c_3}(\alpha)$, and approximate upper 100α percentiles of $-2\rho \log \lambda^*_{c(m)}$: $q_{c_1}(\alpha), q^\dagger_c(\alpha)$, where

$$
\begin{aligned}
q_{c_1}(\alpha) &= \chi^2_f(\alpha), && (13)\\
q_{c_2}(\alpha) &= \chi^2_f(\alpha) + \frac{1}{n}\frac{2\beta_{(m)}}{f}\chi^2_f(\alpha), && (14)\\
q_{c_3}(\alpha) &= \chi^2_f(\alpha) + \frac{1}{n}\frac{2\beta_{(m)}}{f}\chi^2_f(\alpha) + \frac{1}{n^2}\frac{2\gamma_{(m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}, && (15)\\
q^\dagger_c(\alpha) &= \chi^2_f(\alpha) + \frac{1}{\upsilon^2}\frac{2\gamma^*_{(m)}}{f(f+2)}\chi^2_f(\alpha)\{\chi^2_f(\alpha)+f+2\}. && (16)
\end{aligned}
$$

As in Section 4, we present estimators of σ^2 and evaluate them. The MLE of σ^2 under H_0 was provided by Mendoza (1980) as follows.

$$\widetilde{\sigma}^2 = \frac{1}{Np}\sum_{\ell=1}^{m} \mathrm{tr}\boldsymbol{W}^{(\ell)}.$$

The expectation of $\widetilde{\sigma}^2$ is

$$\mathrm{E}[\widetilde{\sigma}^2] = \left(1 - \frac{m}{N}\right)\sigma^2,$$

where $N - m > 0$. Therefore, we propose an unbiased estimator of σ^2 as follows:

$$\widetilde{\sigma}^2_{\mathrm{U}} = \left(1 - \frac{m}{N}\right)^{-1}\widetilde{\sigma}^2.$$

The MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ are

$$
\begin{aligned}
\mathrm{MSE}[\widetilde{\sigma}^2] &= \frac{2N + m^2 p - 2m}{N^2 p}\sigma^4,\\
\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}] &= \frac{2}{(N-m)p}\sigma^4,
\end{aligned}
$$

respectively. Considering the relationship between $\mathrm{MSE}[\widetilde{\sigma}^2]$ and $\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$ in terms of the magnitude, we obtain the following.

$$
\begin{cases}
\mathrm{MSE}[\widetilde{\sigma}^2] \le \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}] & \text{if } mp \le 4,\\
\mathrm{MSE}[\widetilde{\sigma}^2] \ge \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}] & \text{if } mp > 4,
\end{cases}
$$

where the equality relations are satisfied at $N \to \infty$. The evaluation of those inequalities is the same as that in Section 4.2.

7 Simulation studies

This section numerically evaluates the actual type I error rates for the approximate upper percentiles using Monte Carlo simulation (10^6 runs). Let

$$\begin{aligned}
\alpha_i &= \Pr\{-2\log\lambda_{(m)}^{(k)} > q_i(\alpha)\}, i = 1, 2, 3,\\
\alpha_{\chi^2} &= \Pr\{-2\rho\log\lambda_{(m)}^{(k)} > q_1(\alpha)\},\\
\alpha^\dagger &= \Pr\{-2\rho\log\lambda_{(m)}^{(k)} > q^\dagger(\alpha)\},\\
\alpha_{c_i} &= \Pr\{-2\log\lambda_{c(m)}^{*} > q_{c_i}(\alpha)\}, i = 1, 2, 3,\\
\alpha_{c\chi^2} &= \Pr\{-2\rho\log\lambda_{c(m)}^{*} > q_{c_1}(\alpha)\},\\
\alpha_c^\dagger &= \Pr\{-2\rho\log\lambda_{c(m)}^{*} > q_c^\dagger(\alpha)\}.
\end{aligned}$$

$q_1(\alpha), q_2(\alpha), q_3(\alpha), q^\dagger(\alpha), q_{c_1}(\alpha), q_{c_2}(\alpha), q_{c_3}(\alpha)$ and $q_c^\dagger(\alpha)$ are denoted by (6),(7),(8),(9),(13),(14),(15) and (16), respectively. In Tables 1-8, we provide the simulated upper 100α percentiles and approximate upper 100α percentiles of the LRT statistic and modified LRT statistic and actual type I error rates for $\alpha = 0.05$. Additionally, we provide the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}_{\mathrm{U}}^2$. We assume $\sigma^2 = 1$ because the LR is invariant under H_0 for the affine transformation and independent of σ^2 under H_0. In Tables 1-8, we set the population sizes as well as dimensions and sample sizes as follows.

- Table 1: Complete data for a two-sample problem

$$m = 2, p = 2, 8, N^{(\ell)} = 10, 20, 40, 50, 80, 100, 200, 400.$$

- Table 2: Complete data for a two-sample problem with missing parts removed from two-step monotone missing data for a two-sample problem

$$m = 2, (p_1, p_2) = (8, 0), (N_1^{(\ell)}, N_2^{(\ell)}) = (t, 0), t = 10, 20, 40, 50, 80, 100, 200, 400.$$

- Table 3: Two-step monotone missing data for a two-sample problem

$$m = 2,$$

$$
\begin{aligned}
(p_1, p_2) &= (1, 1),\\
(N_1^{(\ell)}, N_2^{(\ell)}) &= (t, t), t = 10, 20, 40, 50, 80, 100, 200, 400,\\
(p_1, p_2) &= (4, 4),\\
(N_1^{(\ell)}, N_2^{(\ell)}) &= \begin{cases} (t, t),\\ (t, 2t), t = 10, 20, 40, 50, 80, 100, 200, 400,\\ (t, t/2). \end{cases}
\end{aligned}
$$

- Table 4: Two-step monotone missing data for a two-sample problem

$$
\begin{aligned}
m = 2, (p_1, p_2) &= (2, 6), (6, 2),\\
(N_1^{(\ell)}, N_2^{(\ell)}) &= (t, t), t = 10, 20, 40, 50, 80, 100, 200, 400.
\end{aligned}
$$

- Table 5: Two-step monotone missing data for a three-sample problem

$$
\begin{aligned}
m = 3, (p_1, p_2) &= (4, 4),\\
(N_1^{(\ell)}, N_2^{(\ell)}) &= (t, t), t = 10, 20, 40, 50, 80, 100, 200, 400.
\end{aligned}
$$

- Table 6: Two-step monotone missing data for a five-sample problem

$$
\begin{aligned}
m = 5, (p_1, p_2) &= (4, 4),\\
(N_1^{(\ell)}, N_2^{(\ell)}) &= (t, t), t = 10, 20, 40, 50, 80, 100, 200, 400.
\end{aligned}
$$

- Table 7: Three-step monotone missing data for a three-sample problem

$$
\begin{aligned}
m = 3, (p_1, p_2, p_3) &= (4, 4, 4),\\
(N_1^{(\ell)}, N_2^{(\ell)}, N_3^{(\ell)}) &= (t, t, t), t = 20, 40, 50, 80, 100, 200, 400.
\end{aligned}
$$

- Table 8: Five-step monotone missing data for a two-sample problem

$$
\begin{aligned}
m = 2, (p_1, p_2, p_3, p_4, p_5) &= (4, 4, 4, 4, 4),\\
(N_1^{(\ell)}, N_2^{(\ell)}, N_3^{(\ell)}, N_4^{(\ell)}, N_5^{(\ell)}) &= (t, t, t, t, t), t = 30, 40, 50, 80, 100, 200, 400.
\end{aligned}
$$

Table 1 satisfies $N^{(1)} = N^{(2)}$ and Tables 2-8 satisfy $N_j^{(1)} = \cdots = N_j^{(m)}$ for $j = 1, \ldots, k$. The smaller value between $\mathrm{MSE}[\widetilde{\sigma}^2]$ and $\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$ in each row is in bold.

Tables 1-8 show that the simulated values are closer to the upper percentile of the χ^2 distribution when the sample size increases. In addition, under complete data, by comparing the actual type I error rates $\alpha_{c_1}, \alpha_{c_2}, \alpha_{c_3}, \alpha_{c\chi^2}, \alpha_c^\dagger$, the accuracy of the approximate upper 100α percentile $q_c^\dagger(\alpha)$ is the best. Under monotone missing data, by comparing the actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}, \alpha^\dagger$, the accuracy of the approximate upper 100α percentile $q^\dagger(\alpha)$ is ideal. Furthermore, comparing the approximation accuracy of the upper 100α percentile $q_c^\dagger(\alpha)$ in Table 1 when $p = 8$ and that of $q^\dagger(\alpha)$ in Table 2, no significant differences are observed. For the estimators of σ^2, $\mathrm{E}[\widetilde{\sigma}^2]$ approaches σ^2 as the sample size increases, and $\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$ is always σ^2. Moreover, only in Table 1 when $p = 2$ and in Table 3 when $(p_1, p_2) = (1, 1)$, we have $\mathrm{MSE}[\widetilde{\sigma}^2] < \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$; in all other cases in the Tables, $\mathrm{MSE}[\widetilde{\sigma}^2] > \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$. As the sample size increases, $\mathrm{MSE}[\widetilde{\sigma}^2] = \mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$.

Table 1: The simulated values for $-2\log\lambda^*_{c(2)}, -2\rho\log\lambda^*_{c(2)}$ and approximate values for $-2\log\lambda^*_{c(2)}, -2\rho\log\lambda^*_{c(2)}$ and actual type I error rates $\alpha_{c_1}, \alpha_{c_2}, \alpha_{c_3}, \alpha_{c\chi^2}$ and $\alpha^\dagger_c$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ for $\alpha = 0.05, m = 2, \sigma^2 = 1$

Sample size	Upper percentile					Type I error rate					Expectation		Mean squared error	
$N^{(\ell)}$	$q_{c_2}(\alpha)$	$q_{c_3}(\alpha)$	$-2\log\lambda^*_{c(2)}$	$q^\dagger_c(\alpha)$	$-2\rho\log\lambda^*_{c(2)}$	α_{c_1}	α_{c_2}	α_{c_3}	$\alpha_{c\chi^2}$	$\alpha^\dagger_c$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$
							$p = 2$							
10	12.12	12.26	12.26	11.09	11.10	.075	.053	**.050**	.051	**.050**	.900	1.000	**.0550**	.0557
20	11.57	11.60	11.60	11.07	11.08	.061	.051	**.050**	**.050**	**.050**	.950	1.000	**.0262**	.0263
40	11.31	11.32	11.32	11.07	11.08	.055	**.050**	**.050**	**.050**	**.050**	.975	1.000	**.0128**	**.0128**
50	11.26	11.27	11.27	11.07	11.08	.054	**.050**	**.050**	**.050**	**.050**	.980	1.000	**.0102**	**.0102**
80	11.19	11.19	11.19	11.07	11.07	.052	**.050**	**.050**	**.050**	**.050**	.987	1.000	**.0063**	**.0063**
100	11.17	11.17	11.17	11.07	11.08	.052	**.050**	**.050**	**.050**	**.050**	.990	1.000	**.0050**	**.0050**
200	11.12	11.12	11.11	11.07	11.06	.051	**.050**	**.050**	**.050**	**.050**	.995	1.000	**.0025**	**.0025**
400	11.09	11.09	11.08	11.07	11.06	**.050**	**.050**	**.050**	**.050**	**.050**	.997	1.000	**.0013**	**.0013**
							$p = 8$							
10	120.55	135.65	161.11	101.99	110.35	.915	.475	.242	.266	**.113**	.900	1.000	.0212	**.0139**
20	105.35	108.74	109.68	93.17	93.31	.300	.083	.056	.063	**.051**	.950	1.000	.0084	**.0066**
40	98.33	99.14	99.17	91.97	91.96	.126	.056	**.050**	.052	**.050**	.975	1.000	.0037	**.0032**
50	96.97	97.48	97.45	91.85	91.81	.105	.053	**.050**	.051	**.050**	.980	1.000	.0029	**.0026**
80	94.96	95.16	95.16	91.74	91.75	.080	.051	**.050**	.051	**.050**	.988	1.000	.0017	**.0016**
100	94.30	94.42	94.40	91.71	91.70	.073	.051	**.050**	**.050**	**.050**	.990	1.000	**.0013**	**.0013**
200	92.98	93.01	93.05	91.68	91.72	.061	.051	**.050**	**.050**	**.050**	.995	1.000	**.0006**	**.0006**
400	92.32	92.33	92.38	91.67	91.72	.055	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0003**	**.0003**

Note. The closest to α from among $\alpha_{c_1}, \alpha_{c_2}, \alpha_{c_3}, \alpha_{c\chi^2}$ and $\alpha^\dagger_c$ of each row is in bold.
$q_{c_1}(\alpha) = \chi^2_5(0.05) = 11.07$ for $p = 2$, $q_{c_1}(\alpha) = \chi^2_{71}(0.05) = 91.67$ for $p = 8$.

Table 2: The simulated values for $-2\log\lambda^{(2)}_{(2)}, -2\rho\log\lambda^{(2)}_{(2)}$ and approximate values for $-2\log\lambda^{(2)}_{(2)}, -2\rho\log\lambda^{(2)}_{(2)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ for $\alpha = 0.05, m = 2, \sigma^2 = 1$

Sample size		Upper percentile					Type I error rate					Expectation		Mean squared error	
$N^{(\ell)}_1$	$N^{(\ell)}_2$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda^{(2)}_{(2)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda^{(2)}_{(2)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$
							$(p_1, p_2) = (8, 0)$								
10	0	126.83	145.96	178.98	101.99	110.33	.972	.592	.300	.266	**.113**	.900	1.000	.0213	**.0139**
20	0	109.25	114.03	115.44	93.17	93.30	.414	.098	.059	.063	**.051**	.950	1.000	.0084	**.0066**
40	0	100.46	101.66	101.74	91.97	91.98	.164	.059	**.050**	.052	**.050**	.975	1.000	.0037	**.0032**
50	0	98.70	99.47	99.50	91.85	91.86	.131	.056	**.050**	.051	**.050**	.980	1.000	.0028	**.0025**
80	0	96.07	96.36	96.41	91.74	91.78	.093	.052	**.050**	.051	**.050**	.988	1.000	.0017	**.0016**
100	0	95.19	95.38	95.34	91.71	91.68	.082	.051	**.050**	**.050**	**.050**	.990	1.000	**.0013**	**.0013**
200	0	93.43	93.48	93.41	91.68	91.62	.064	**.050**	**.050**	**.050**	**.050**	.995	1.000	**.0006**	**.0006**
400	0	92.55	92.56	92.59	91.67	91.70	.057	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0003**	**.0003**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.
$q_1(\alpha) = \chi^2_{71}(0.05) = 91.67$.

Table 3: The simulated values for $-2\log\lambda_{(2)}^{(2)}, -2\rho\log\lambda_{(2)}^{(2)}$ and approximate values for $-2\log\lambda_{(2)}^{(2)}, -2\rho\log\lambda_{(2)}^{(2)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ for $\alpha = 0.05, m = 2, \sigma^2 = 1$

Sample size		Upper percentile					Type I error rate					Expectation		Mean squared error	
$N_1^{(\ell)}$	$N_2^{(\ell)}$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda_{(2)}^{(2)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda_{(2)}^{(2)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$
$(p_1, p_2) = (1, 1)$															
10	10	12.94	13.39	13.38	11.14	11.12	.103	.057	**.050**	.051	**.050**	.933	1.000	**.0355**	.0357
20	20	12.01	12.12	12.11	11.08	11.09	.072	.052	**.050**	**.050**	**.050**	.967	1.000	**.0172**	**.0172**
40	40	11.54	11.57	11.58	11.07	11.09	.060	.051	**.050**	**.050**	**.050**	.983	1.000	**.0085**	**.0085**
50	50	11.44	11.46	11.46	11.07	11.07	.058	**.050**	**.050**	**.050**	**.050**	.987	1.000	**.0068**	**.0068**
80	80	11.30	11.31	11.31	11.07	11.07	.055	**.050**	**.050**	**.050**	**.050**	.992	1.000	**.0042**	**.0042**
100	100	11.26	11.26	11.26	11.07	11.07	.054	**.050**	**.050**	**.050**	**.050**	.993	1.000	**.0033**	**.0033**
200	200	11.16	11.17	11.17	11.07	11.08	.052	**.050**	**.050**	**.050**	**.050**	.997	1.000	**.0017**	**.0017**
400	400	11.12	11.12	11.13	11.07	11.08	.051	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0008**	**.0008**
$(p_1, p_2) = (4, 4)$															
10	10	123.82	141.47	174.33	103.38	113.19	.958	.578	.304	.308	**.125**	.933	1.000	.0122	**.0089**
20	20	107.74	112.16	113.65	93.49	93.72	.377	.096	.059	.066	**.052**	.967	1.000	.0051	**.0043**
40	40	99.71	100.81	100.92	92.04	92.07	.151	.059	.051	.053	**.050**	.983	1.000	.0023	**.0021**
50	50	98.10	98.81	98.80	91.90	91.87	.123	.055	**.050**	.051	**.050**	.987	1.000	.0018	**.0017**
80	80	95.69	95.96	95.99	91.75	91.78	.089	.052	**.050**	.051	**.050**	.992	1.000	**.0011**	**.0011**
100	100	94.88	95.06	95.07	91.72	91.73	.079	.051	**.050**	**.050**	**.050**	.993	1.000	.0009	**.0008**
200	200	93.28	93.32	93.30	91.68	91.66	.063	**.050**	**.050**	**.050**	**.050**	.997	1.000	**.0004**	**.0004**
400	400	92.47	92.48	92.46	91.67	91.65	.056	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0002**	**.0002**
10	20	122.88	140.27	173.27	104.09	114.28	.955	.577	.307	.324	**.127**	.950	1.000	.0084	**.0066**
20	40	107.27	111.62	113.15	93.63	93.89	.365	.095	.060	.068	**.052**	.975	1.000	.0037	**.0032**
40	80	99.47	100.56	100.61	92.07	92.05	.147	.058	**.050**	.053	**.050**	.987	1.000	.0017	**.0016**
50	100	97.91	98.61	98.65	91.92	91.93	.120	.055	**.050**	.052	**.050**	.990	1.000	**.0013**	**.0013**
80	160	95.57	95.84	95.86	91.76	91.78	.087	.052	**.050**	.051	**.050**	.994	1.000	**.0008**	**.0008**
100	200	94.79	94.96	94.99	91.73	91.75	.078	.051	**.050**	.051	**.050**	.995	1.000	**.0006**	**.0006**
200	400	93.23	93.27	93.30	91.68	91.71	.063	.051	**.050**	**.050**	**.050**	.998	1.000	**.0003**	**.0003**
400	800	92.45	92.46	92.44	91.67	91.66	.056	**.050**	**.050**	**.050**	**.050**	.999	1.000	**.0002**	**.0002**
10	5	124.80	142.83	175.69	102.78	112.20	.963	.579	.302	.293	**.122**	.920	1.000	.0156	**.0109**
20	10	108.23	112.74	114.17	93.36	93.54	.387	.095	.059	.065	**.051**	.960	1.000	.0064	**.0052**
40	20	99.95	101.08	101.20	92.01	92.06	.155	.059	.051	.052	**.050**	.980	1.000	.0028	**.0025**
50	25	98.30	99.02	99.05	91.88	91.89	.125	.055	**.050**	.052	**.050**	.984	1.000	.0022	**.0020**
80	40	95.81	96.09	96.11	91.75	91.77	.090	.052	**.050**	.051	**.050**	.990	1.000	**.0013**	**.0013**
100	50	94.98	95.16	95.14	91.72	91.70	.080	.051	**.050**	**.050**	**.050**	.992	1.000	.0011	**.0010**
200	100	93.33	93.37	93.34	91.68	91.66	.063	**.050**	**.050**	**.050**	**.050**	.996	1.000	**.0005**	**.0005**
400	200	92.50	92.51	92.46	91.67	91.62	.056	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0003**	**.0003**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.

$q_1(\alpha) = \chi^2_5(0.05) = 11.07$ for $(p_1, p_2) = (1, 1)$.

$q_1(\alpha) = \chi^2_{71}(0.05) = 91.67$ for $(p_1, p_2) = (4, 4)$.

Table 4: The simulated values for $-2\log\lambda^{(2)}_{(2)}, -2\rho\log\lambda^{(2)}_{(2)}$ and approximate values for $-2\log\lambda^{(2)}_{(2)}, -2\rho\log\lambda^{(2)}_{(2)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ for $\alpha = 0.05, m = 2, \sigma^2 = 1$

Sample size		Upper percentile					Type I error rate					Expectation		Mean squared error	
$N_1^{(\ell)}$	$N_2^{(\ell)}$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda^{(2)}_{(2)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda^{(2)}_{(2)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$
$(p_1, p_2) = (2, 6)$															
10	10	126.27	145.21	178.12	102.55	110.90	.970	.591	.301	.276	**.113**	.920	1.000	.0156	**.0108**
20	20	108.97	113.70	115.18	93.27	93.45	.408	.098	.059	.064	**.051**	.960	1.000	.0064	**.0052**
40	40	100.32	101.50	101.59	91.99	92.01	.161	.059	.051	.053	**.050**	.980	1.000	.0028	**.0026**
50	50	98.59	99.35	99.36	91.87	91.86	.129	.056	**.050**	.051	**.050**	.984	1.000	.0022	**.0020**
80	80	95.99	96.29	96.35	91.74	91.80	.093	.052	**.050**	.051	**.050**	.990	1.000	**.0013**	**.0013**
100	100	95.13	95.32	95.34	91.72	91.74	.082	.051	**.050**	.051	**.050**	.992	1.000	.0011	**.0010**
200	200	93.40	93.45	93.46	91.68	91.70	.064	**.050**	**.050**	**.050**	**.050**	.996	1.000	**.0005**	**.0005**
400	400	92.54	92.55	92.56	91.67	91.69	.057	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0003**	**.0003**
$(p_1, p_2) = (6, 2)$															
10	10	118.48	132.17	161.52	101.34	114.27	.902	.493	.278	.314	**.156**	.943	1.000	.0100	**.0076**
20	20	105.08	108.50	109.79	93.33	93.73	.300	.086	.058	.066	**.053**	.971	1.000	.0043	**.0037**
40	40	98.37	99.23	99.30	92.02	92.04	.129	.057	**.050**	.053	**.050**	.986	1.000	.0020	**.0018**
50	50	97.03	97.58	97.56	91.89	91.85	.107	.054	**.050**	.051	**.050**	.989	1.000	.0015	**.0014**
80	80	95.02	95.24	95.30	91.75	91.82	.081	.052	**.050**	.051	**.050**	.993	1.000	**.0009**	**.0009**
100	100	94.35	94.49	94.47	91.72	91.70	.073	.051	**.050**	**.050**	**.050**	.994	1.000	**.0007**	**.0007**
200	200	93.01	93.05	93.06	91.68	91.70	.061	**.050**	**.050**	**.050**	**.050**	.997	1.000	**.0004**	**.0004**
400	400	92.34	92.35	92.32	91.67	91.65	.055	**.050**	**.050**	**.050**	**.050**	.999	1.000	**.0002**	**.0002**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.
$q_1(\alpha) = \chi^2_{71}(0.05) = 91.67$ for $(p_1, p_2) = (2, 6), (6, 2)$.

Table 5: The simulated values for $-2\log\lambda^{(2)}_{(3)}, -2\rho\log\lambda^{(2)}_{(3)}$ and approximate values for $-2\log\lambda^{(2)}_{(3)}, -2\rho\log\lambda^{(2)}_{(3)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ for $\alpha = 0.05, m = 3, \sigma^2 = 1$

Sample size		Upper percentile					Type I error rate					Expectation		Mean squared error	
$N_1^{(\ell)}$	$N_2^{(\ell)}$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda^{(2)}_{(3)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda^{(2)}_{(3)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$
$(p_1, p_2) = (4, 4)$															
10	10	178.31	203.15	249.46	148.68	162.31	.992	.710	.386	.386	**.144**	.933	1.000	.0096	**.0059**
20	20	155.23	161.44	163.56	134.71	134.99	.481	.107	.061	.070	**.052**	.967	1.000	.0038	**.0029**
40	40	143.69	145.24	145.35	132.67	132.66	.182	.060	.051	.053	**.050**	.983	1.000	.0016	**.0014**
50	50	141.38	142.37	142.38	132.47	132.43	.143	.056	**.050**	.052	**.050**	.987	1.000	.0013	**.0011**
80	80	137.92	138.30	138.31	132.26	132.27	.099	.052	**.050**	.051	**.050**	.992	1.000	.0008	**.0007**
100	100	136.76	137.01	137.05	132.22	132.26	.087	.052	**.050**	.051	**.050**	.993	1.000	**.0006**	**.0006**
200	200	134.45	134.51	134.47	132.16	132.13	.066	**.050**	**.050**	**.050**	**.050**	.997	1.000	**.0003**	**.0003**
400	400	133.30	133.31	133.29	132.15	132.13	.057	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0001**	**.0001**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.
$q_1(\alpha) = \chi^2_{107}(0.05) = 132.14$.

Table 6: The simulated values for $-2\log\lambda_{(5)}^{(2)}, -2\rho\log\lambda_{(5)}^{(2)}$ and approximate values for $-2\log\lambda_{(5)}^{(2)}, -2\rho\log\lambda_{(5)}^{(2)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}_{\mathrm{U}}^2$ for $\alpha = 0.05, m = 5, \sigma^2 = 1$

Sample size		Upper percentile					Type I error rate					Expectation		Mean squared error	
$N_1^{(\ell)}$	$N_2^{(\ell)}$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda_{(5)}^{(2)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda_{(5)}^{(2)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}_{\mathrm{U}}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}_{\mathrm{U}}^2]$
$(p_1, p_2) = (4, 4)$															
10	10	284.79	323.54	395.58	237.09	257.79	1.000	.867	.523	.517	**.176**	.933	1.000	.0076	**.0036**
20	20	248.00	257.69	261.22	215.24	215.72	.650	.128	.065	.075	**.052**	.967	1.000	.0027	**.0017**
40	40	229.61	232.03	232.23	212.04	212.01	.238	.063	.051	.054	**.050**	.983	1.000	.0011	**.0008**
50	50	225.93	227.48	227.57	211.72	211.72	.181	.058	**.050**	.052	**.050**	.987	1.000	.0008	**.0007**
80	80	220.41	221.02	221.02	211.40	211.39	.116	.053	**.050**	.051	**.050**	.992	1.000	.0005	**.0004**
100	100	218.57	218.96	218.95	211.33	211.32	.099	.052	**.050**	**.050**	**.050**	.993	1.000	.0004	**.0003**
200	200	214.90	214.99	215.04	211.25	211.29	.071	.051	**.050**	**.050**	**.050**	.997	1.000	**.0002**	**.0002**
400	400	213.06	213.08	213.11	211.22	211.26	.060	**.050**	**.050**	**.050**	**.050**	.998	1.000	**.0001**	**.0001**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.
$q_1(\alpha) = \chi_{179}^2(0.05) = 211.22$.

Table 7: The simulated values for $-2\log\lambda_{(3)}^{(3)}, -2\rho\log\lambda_{(3)}^{(3)}$ and approximate values for $-2\log\lambda_{(3)}^{(3)}, -2\rho\log\lambda_{(3)}^{(3)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}_{\mathrm{U}}^2$ for $\alpha = 0.05, m = 3, \sigma^2 = 1$

Sample size			Upper percentile					Type I error rate					Expectation		Mean squared error	
$N_1^{(\ell)}$	$N_2^{(\ell)}$	$N_3^{(\ell)}$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda_{(3)}^{(3)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda_{(3)}^{(3)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}_{\mathrm{U}}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}_{\mathrm{U}}^2]$
$(p_1, p_2, p_3) = (4, 4, 4)$																
20	20	20	326.73	347.45	363.80	282.05	286.72	.931	.310	.125	.163	**.071**	.975	1.000	.0020	**.0014**
40	40	40	298.17	303.35	304.47	272.03	272.22	.382	.080	.054	.062	**.051**	.987	1.000	.0008	**.0007**
50	50	50	292.46	295.77	296.33	271.08	271.21	.277	.067	.052	.057	**.051**	.990	1.000	.0007	**.0006**
80	80	80	283.89	285.18	285.33	270.15	270.21	.156	.056	.051	.053	**.050**	.994	1.000	.0004	**.0003**
100	100	100	281.03	281.86	281.88	269.94	269.94	.126	.054	**.050**	.051	**.050**	.995	1.000	**.0003**	**.0003**
200	200	200	275.32	275.53	275.58	269.69	269.74	.081	.051	**.050**	.051	**.050**	.998	1.000	**.0001**	**.0001**
400	400	400	272.46	272.52	272.54	269.63	269.65	.064	**.050**	**.050**	**.050**	**.050**	.999	1.000	**.0001**	**.0001**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.
$q_1(0.05) = \chi_{233}^2(0.05) = 269.61$.

Table 8: The simulated values for $-2\log\lambda_{(2)}^{(5)}, -2\rho\log\lambda_{(2)}^{(5)}$ and approximate values for $-2\log\lambda_{(2)}^{(5)}, -2\rho\log\lambda_{(2)}^{(5)}$ and actual type I error rates $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$, the expectations and MSEs of $\widetilde{\sigma}^2$ and $\widetilde{\sigma}^2_{\mathrm{U}}$ for $\alpha = 0.05, m = 2, \sigma^2 = 1$

Sample size					Upper percentile					Type I error rate					Expectation		Mean squared error	
$N_1^{(\ell)}$	$N_2^{(\ell)}$	$N_3^{(\ell)}$	$N_4^{(\ell)}$	$N_5^{(\ell)}$	$q_2(\alpha)$	$q_3(\alpha)$	$-2\log\lambda_{(2)}^{(5)}$	$q^\dagger(\alpha)$	$-2\rho\log\lambda_{(2)}^{(5)}$	α_1	α_2	α_3	α_{χ^2}	$\alpha^\dagger$	$\mathrm{E}[\widetilde{\sigma}^2]$	$\mathrm{E}[\widetilde{\sigma}^2_{\mathrm{U}}]$	$\mathrm{MSE}[\widetilde{\sigma}^2]$	$\mathrm{MSE}[\widetilde{\sigma}^2_{\mathrm{U}}]$
$(p_1, p_2, p_3, p_4, p_5) = (4, 4, 4, 4, 4)$																		
30	30	30	30	30	551.27	581.33	610.75	488.94	501.65	.979	.418	.172	.255	**.100**	.989	1.000	.0007	**.0006**
40	40	40	40	40	530.38	547.30	556.32	478.46	481.79	.800	.172	.080	.113	**.061**	.992	1.000	.0005	**.0004**
50	50	50	50	50	517.85	528.68	532.56	474.19	475.49	.598	.110	.062	.081	**.054**	.993	1.000	.0004	**.0003**
80	80	80	80	80	499.05	503.28	504.03	470.04	470.27	.295	.067	.052	.059	**.051**	.996	1.000	**.0002**	**.0002**
100	100	100	100	100	492.79	495.49	495.94	469.16	469.36	.218	.061	.051	.056	**.051**	.997	1.000	**.0002**	**.0002**
200	200	200	200	200	480.26	480.93	480.99	468.07	468.10	.110	.052	**.050**	.051	**.050**	.998	1.000	**.0001**	**.0001**
400	400	400	400	400	473.99	474.16	474.13	467.81	467.78	.075	**.050**	**.050**	**.050**	**.050**	.999	1.000	**.0000**	**.0000**

Note. The closest to α from among $\alpha_1, \alpha_2, \alpha_3, \alpha_{\chi^2}$ and $\alpha^\dagger$ of each row is in bold.
$q_1(0.05) = \chi^2_{419}(0.05) = 467.73$.

8 Numerical examples

This section provides instances of the test statistics $(-2\log\lambda_{(m)}^{(k)}, -2\rho\log\lambda_{(m)}^{(k)})$ and approximate upper percentiles $(q_1(\alpha), q_2(\alpha), q_3(\alpha), q^\dagger(\alpha))$ under two-step and three-step monotone missing data for a two-sample problem. We employ the part of the data provided in Flury (1997). Flury (1997) provides head dimentions of 200 young Swiss men and 59 young Swiss women, age 20.

8.1 Two-step case

We use the data that consist of the first 10 men and women obtained by Flury (1997), with the sixth variable removed and the fifth variable partially missing. The parameters are

$$m = 2, (p_1, p_2) = (4, 1), (N_1^{(1)}, N_2^{(1)}) = (6, 4), (N_1^{(2)}, N_2^{(2)}) = (7, 3).$$

Subsequently, $-2\log\lambda_{(2)}^{(2)}$, $-2\rho\log\lambda_{(2)}^{(2)}$ and the approximate upper 5 percentiles are, respectively,

$$-2\log\lambda_{(2)}^{(2)} = 62.43, -2\rho\log\lambda_{(2)}^{(2)} = 35.42,$$

$$q_1(0.05) = 42.56, q_2(0.05) = 60.97, q_3(0.05) = 75.09, q^\dagger(0.05) = 57.05.$$

Therefore, H_0 is rejected when $-2\log\lambda_{(2)}^{(2)}$ and $q_1(0.05)$ or $-2\log\lambda_{(2)}^{(2)}$ and $q_2(0.05)$ are compared, but H_0 is not rejected in all other cases.

8.2 Three-step case

We use the data consisting of the first 10 men and women obtained by Flury (1997), with the sixth variable removed and the fourth and fifth variables partially missing. The parameters are

$$m = 2, (p_1, p_2, p_3) = (3, 1, 1), (N_1^{(1)}, N_2^{(1)}, N_3^{(1)}) = (6, 3, 1), (N_1^{(2)}, N_2^{(2)}, N_3^{(2)}) = (7, 2, 1).$$

Subsequently, $-2\log\lambda_{(2)}^{(3)}$, $-2\rho\log\lambda_{(2)}^{(3)}$ and the approximate upper 5 percentiles are, respectively,

$$\begin{aligned} &-2\log\lambda_{(2)}^{(3)} = 69.07, -2\rho\log\lambda_{(2)}^{(3)} = 44.30, \\ &q_1(0.05) = 42.56, q_2(0.05) = 57.82, q_3(0.05) = 66.36, q^\dagger(0.05) = 47.55. \end{aligned}$$

Thus, H_0 is not rejected when $-2\rho\log\lambda_{(2)}^{(3)}$ and $q^\dagger(0.05)$ are compared, but H_0 is rejected in all other cases.

9 Conclusions

In this study, we discuss the sphericity test under monotone missing data for a multi-sample problem. Specifically, we asymptotically expanded the distribution functions and upper percentiles of the LRT statistic and modified LRT statistic under H_0. Furthermore, we provided approximate upper percentiles. The simulation results show that modifying the LRT statistic improves the approximation accuracy. Furthermore, the approximate upper 100α percentile $q^\dagger(\alpha)$ is effective even when the sample size for each step is small. In the estimation of σ^2, we derived the MLE $\widetilde{\sigma}^2$ of σ^2 under H_0 and proposed an unbiased estimator $\widetilde{\sigma}_{\mathrm{U}}^2$ of σ^2. It is ideal to use $\widetilde{\sigma}_{\mathrm{U}}^2$ rather than $\widetilde{\sigma}^2$ in most cases, especially when the sample size for each step is small. Finally, we provided instances of the LRT statistic and modified LRT statistic.

Acknowledgements

The second and third authors' research was partly supported by a Grant-in-Aid for Early-Career Scientists (22K13961) and a Grant-in-Aid for Scientific Research (C) (24K14863), respectively.

Appendix.A

Proof of Theorem 1

We consider a random variable Z $(0 < Z \leq 1)$ with moments.

$$\mathrm{E}[Z^h] = K \left[\frac{\prod_{g=1}^{r} y_g^{y_g}}{\prod_{\ell=1}^{m}\prod_{i=1}^{q_1} z_{1i}^{(\ell)^{z_{1i}^{(\ell)}}} \prod_{\ell=1}^{m}\prod_{i=1}^{q_2} z_{2i}^{(\ell)^{z_{2i}^{(\ell)}}}} \right]^h$$
$$\times \frac{\prod_{\ell=1}^{m}\prod_{i=1}^{q_1} \Gamma\left[z_{1i}^{(\ell)}(1+h)+\xi_{1i}\right] \prod_{\ell=1}^{m}\prod_{i=1}^{q_2} \Gamma\left[z_{2i}^{(\ell)}(1+h)+\xi_{2i}\right]}{\prod_{g=1}^{r} \Gamma\left[y_g(1+h)+\eta_g\right]}, \tag{17}$$

where

$$\sum_{g=1}^{r} y_g = \sum_{\ell=1}^{m}\sum_{i=1}^{q_1} z_{1i}^{(\ell)} + \sum_{\ell=1}^{m}\sum_{i=1}^{q_2} z_{2i}^{(\ell)} \tag{18}$$

and K is a constant such that $\mathrm{E}[Z^0] = 1$. Moreover, $y_g = a_g N, z_{1i}^{(\ell)} = b_{1i}^{(\ell)} N, z_{2i}^{(\ell)} = b_{2i}^{(\ell)} N$, where $a_g, b_{1i}^{(\ell)}$ and $b_{2i}^{(\ell)}$ are constants and N is the asymptotic variable. Subsequently, the cumulant generating function of $-2\rho \log Z$ $(0 < \rho \leq 1)$ can be denoted by

$$\Psi(t) = -\frac{1}{2} f \log(1-2it) + \sum_{u=1}^{s} \omega_u(\rho) \big\{(1-2it)^{-u} - 1\big\} + O(N^{-s-1}), \tag{19}$$

where

$$f = -2\left\{ m\sum_{i=1}^{q_1} \xi_{1i} + m\sum_{i=1}^{q_2} \xi_{2i} - \sum_{g=1}^{r} \eta_g - \frac{1}{2}\big(m(q_1+q_2)-r\big) \right\}, \tag{20}$$

$$\omega_u(\rho) = \frac{(-1)^{u+1}}{u(u+1)} \left(\sum_{\ell=1}^{m}\sum_{i=1}^{q_1} \frac{B_{u+1}(\beta_{1i}^{(\ell)}+\xi_{1i})}{(\rho z_{1i}^{(\ell)})^u} + \sum_{\ell=1}^{m}\sum_{i=1}^{q_2} \frac{B_{u+1}(\beta_{2i}^{(\ell)}+\xi_{2i})}{(\rho z_{2i})^u} \right.$$

$$
-\sum_{g=1}^{r}\frac{B_{u+1}(\varepsilon_g+\eta_g)}{(\rho y_g)^u}\Bigg). \tag{21}
$$

$B_{u+1}(a)$ is the Bernoulli polynomial of degree $u+1$ and $\beta_{1i}^{(\ell)}=(1-\rho)z_{1i}^{(\ell)}, \beta_{2i}^{(\ell)}=(1-\rho)z_{2i}^{(\ell)}, \varepsilon_g=(1-\rho)y_g$. By employing (18),(19),(20) and (21), the distribution functions of $-2\log Z$ and $-2\rho\log Z$ can be asymptotically expanded as

$$
\begin{aligned}
\Pr(-2\log Z\le x) &= G_f(x)+\omega_1(1)\big\{G_{f+2}(x)-G_f(x)\big\}\\
&\quad +\omega_2(1)\big\{G_{f+4}(x)-G_f(x)\big\}+O(N^{-3}), \qquad (22)\\
\Pr(-2\rho\log Z\le x) &= G_f(x)+\omega_2(\rho)\big\{G_{f+4}(x)-G_f(x)\big\}+O(N^{-3}), \qquad (23)
\end{aligned}
$$

respectively, where

$$
\begin{aligned}
\omega_1(1) &= \frac{1}{2}\Bigg\{\sum_{\ell=1}^{m}\sum_{i=1}^{q_1} z_{1i}^{(\ell)-1}\left(\xi_{1i}^2-\xi_{1i}+\frac{1}{6}\right)+\sum_{\ell=1}^{m}\sum_{i=1}^{q_2} z_{2i}^{(\ell)-1}\left(\xi_{2i}^2-\xi_{2i}+\frac{1}{6}\right)\\
&\quad -\sum_{g=1}^{r} y_g^{-1}\left(\eta_g^2-\eta_g+\frac{1}{6}\right)\Bigg\},\\
\omega_2(1) &= -\frac{1}{6}\Bigg\{\sum_{\ell=1}^{m}\sum_{i=1}^{q_1} z_{1i}^{(\ell)-2}\left(\xi_{1i}^3-\frac{3}{2}\xi_{1i}^2+\frac{1}{2}\xi_{1i}\right)+\sum_{\ell=1}^{m}\sum_{\ell=1}^{q_2} z_{2i}^{(\ell)-2}\left(\xi_{2i}^3-\frac{3}{2}\xi_{2i}^2+\frac{1}{2}\xi_{2i}\right)\\
&\quad -\sum_{g=1}^{r} y_g^{-2}\left(\eta_g^3-\frac{3}{2}\eta_g^2+\frac{1}{2}\eta_g\right)\Bigg\},\\
\rho &= 1-\frac{2}{f}\omega_1(1),\\
\omega_2(\rho) &= -\frac{1}{\rho^2}\left(\frac{1}{f}\omega_1(1)^2-\omega_2(1)\right).
\end{aligned}
$$

Subsequently, we discuss the following theorem, which we have derived.

Theorem 5

For $h=0,1,2,\ldots,$ the h-th null moment of the LR $\lambda_{(m)}$ is

$$
\mathrm{E}[\lambda_{(m)}^h]=\frac{(Np_1+N_1p_2)^{\frac{(Np_1+N_1p_2)h}{2}}}{\prod_{\ell=1}^{m}N^{(\ell)\frac{N^{(\ell)}p_1h}{2}}\prod_{\ell=1}^{m}N_1^{(\ell)\frac{N_1^{(\ell)}p_2h}{2}}}
$$

$$\times \prod_{\ell=1}^{m} \frac{\Gamma_{p_1}\left[\frac{1}{2}(N^{(\ell)}h + N^{(\ell)} - 1)\right]\Gamma_{p_2}\left[\frac{1}{2}(N_1^{(\ell)}h + N_1^{(\ell)} - p_1 - 1)\right]}{\Gamma_{p_1}\left[\frac{1}{2}(N^{(\ell)} - 1)\right]\Gamma_{p_2}\left[\frac{1}{2}(N_1^{(\ell)} - p_1 - 1)\right]}$$

$$\times \frac{\Gamma\left[\frac{1}{2}\{(N-m)p_1 + (N_1 - m)p_2\}\right]}{\Gamma\left[\frac{1}{2}\{(N-m)p_1 + (N_1 - m)p_2\} + \frac{1}{2}(Np_1 + N_1 p_2)h\right]},$$

where $\Gamma_{p_j}[a]$ *is the multivariate gamma function and*

$$\Gamma_{p_j}[a] = \pi^{\frac{p_j(p_j-1)}{4}} \prod_{i=1}^{p_j} \Gamma\left[a - \frac{1}{2}(i-1)\right]. \tag{24}$$

This theorem extends Chang and Richards (2010). Transforming the h-th null moment of $\lambda_{(m)}$ the same form as (17) using (24), we obtain the following expression.

$$\mathrm{E}[\lambda_{(m)}^h] = K \left[\frac{\left\{\frac{1}{2}(Np_1 + N_1 p_2)\right\}^{\frac{Np_1 + N_1 p_2}{2}}}{\prod_{\ell=1}^{m}\prod_{i=1}^{p_1}\left(\frac{1}{2}N^{(\ell)}\right)^{\frac{N^{(\ell)}}{2}} \prod_{\ell=1}^{m}\prod_{i=1}^{p_2}\left(\frac{1}{2}N_1^{(\ell)}\right)^{\frac{N_1^{(\ell)}}{2}}} \right]^h$$

$$\times \frac{\prod_{\ell=1}^{m}\prod_{i=1}^{p_1}\Gamma\left[\frac{1}{2}N^{(\ell)}(1+h) - \frac{1}{2}i\right]\prod_{\ell=1}^{m}\prod_{i=1}^{p_2}\Gamma\left[\frac{1}{2}N_1^{(\ell)}(1+h) - \frac{1}{2}(i+p_1)\right]}{\Gamma\left[\frac{1}{2}(Np_1 + N_1 p_2)(1+h) - \frac{1}{2}mp\right]}, \tag{25}$$

where

$$K = \frac{\Gamma\left[\frac{1}{2}\{(N-m)p_1 + (N_1 - m)p_2\}\right]}{\prod_{\ell=1}^{m}\prod_{i=1}^{p_1}\Gamma\left[\frac{1}{2}(N^{(\ell)} - i)\right]\prod_{\ell=1}^{m}\prod_{i=1}^{p_2}\Gamma\left[\frac{1}{2}(N_1^{(\ell)} - p_1 - i)\right]}.$$

By setting $Z = \lambda_{(m)}$ and comparing (25) with (17), we obtain the following equations.

$$\begin{aligned} r &= 1, q_1 = p_1, q_2 = p_2, \\ z_{1i}^{(\ell)} &= \frac{1}{2}N^{(\ell)}, \xi_{1i} = -\frac{1}{2}i, i = 1, \ldots, p_1, \\ z_{2i}^{(\ell)} &= \frac{1}{2}N_1^{(\ell)}, \xi_{2i} = -\frac{1}{2}(i + p_1), i = 1, \ldots, p_2, \end{aligned}$$

$$y_1 = \frac{1}{2}(Np_1 + N_1p_2), \eta_1 = -\frac{1}{2}mp,$$

where the above equations satisfy (18). By substituting the equations into (22) and (23), respectively, an asymptotic expansion of $-2\log\lambda_{(m)}$ and $-2\rho\log\lambda_{(m)}$ for the null distribution yield (2) and (3). Therefore, this completes the proof of **Theorem 1**.

Appendix.B

Proof of Theorem 2

We consider a random variable Z $(0 < Z \leq 1)$ with moments.

$$\mathrm{E}[Z^h] = K\left[\frac{\prod_{g=1}^{r} y_g^{y_g}}{\prod_{\ell=1}^{m}\prod_{j=1}^{k}\prod_{i=1}^{q_j} z_{ji}^{(\ell)\,z_{ji}^{(\ell)}}}\right]^h \frac{\prod_{\ell=1}^{m}\prod_{j=1}^{k}\prod_{i=1}^{q_j}\Gamma\left[z_{ji}^{(\ell)}(1+h)+\xi_{ji}\right]}{\prod_{g=1}^{r}\Gamma\left[y_g(1+h)+\eta_g\right]}, \tag{26}$$

where

$$\sum_{g=1}^{r} y_g = \sum_{\ell=1}^{m}\sum_{j=1}^{k}\sum_{i=1}^{q_j} z_{ji}^{(\ell)} \tag{27}$$

and K is a constant such that $\mathrm{E}[Z^0] = 1$. Moreover, $y_g = a_gN, z_{ji}^{(\ell)} = b_{ji}^{(\ell)}N$, where a_g and $b_{ji}^{(\ell)}$ are constants and N is the asymptotic variable. Subsequently, the distribution functions of $-2\log Z$ and $-2\rho\log Z$ $(0 < \rho \leq 1)$ can be expanded as

$$\begin{aligned}\Pr(-2\log Z \leq x) &= G_f(x) + \omega_1(1)\bigl\{G_{f+2}(x) - G_f(x)\bigr\}\\ &\quad + \omega_2(1)\bigl\{G_{f+4}(x) - G_f(x)\bigr\} + O(N^{-3}),\end{aligned} \tag{28}$$

$$\Pr(-2\rho\log Z \leq x) = G_f(x) + \omega_2(\rho)\bigl\{G_{f+4}(x) - G_f(x)\bigr\} + O(N^{-3}), \tag{29}$$

respectively, where

$$f = -2\left\{m\sum_{j=1}^{k}\sum_{i=1}^{q_j}\xi_{ji} - \sum_{g=1}^{r}\eta_g - \frac{1}{2}\left(m\sum_{j=1}^{k}q_j - r\right)\right\},$$

$$\omega_1(1) = \frac{1}{2}\left\{\sum_{\ell=1}^{m}\sum_{j=1}^{k}\sum_{i=1}^{q_j} z_{ji}^{(\ell)-1}\left(\xi_{ji}^2 - \xi_{ji} + \frac{1}{6}\right) - \sum_{g=1}^{r} y_g^{-1}\left(\eta_g^2 - \eta_g + \frac{1}{6}\right)\right\},$$

$$
\begin{aligned}
\omega_2(1) =& -\frac{1}{6}\left\{\sum_{\ell=1}^{m}\sum_{j=1}^{k}\sum_{i=1}^{q_j} z_{ji}^{(\ell)-2}\left(\xi_{ji}^3 - \frac{3}{2}\xi_{ji}^2 + \frac{1}{2}\xi_{ji}\right) - \sum_{g=1}^{r} y_g^{-2}\left(\eta_g^3 - \frac{3}{2}\eta_g^2 + \frac{1}{2}\eta_g\right)\right\},\\
\rho =& 1 - \frac{2}{f}\omega_1(1),\\
\omega_2(\rho) =& -\frac{1}{\rho^2}\left(\frac{1}{f}\omega_1(1)^2 - \omega_2(1)\right).
\end{aligned}
$$

Next, we discuss the following theorem that we have derived.

Theorem 6

For $h = 0, 1, 2, \ldots,$ the h-th null moment of the LR $\lambda_{(m)}^{(k)}$ is

$$
\begin{aligned}
\mathrm{E}[(\lambda_{(m)}^{(k)})^h] =& \frac{\left(\sum_{j=1}^{k} N_{(1\ldots k-j+1)}p_j\right)^{\frac{1}{2}\left(\sum_{j=1}^{k} N_{(1\ldots k-j+1)}p_j\right)h}}{\prod_{\ell=1}^{m}\prod_{j=1}^{k} N_{(1\ldots k-j+1)}^{(\ell)\,\frac{1}{2}N_{(1\ldots k-j+1)}^{(\ell)}p_j h}}\\
&\times \prod_{\ell=1}^{m}\prod_{j=1}^{k}\frac{\Gamma_{p_j}\left[\frac{1}{2}(N_{(1\ldots k-j+1)}^{(\ell)}h + N_{(1\ldots k-j+1)}^{(\ell)} - p_{(1\ldots j-1)} - 1)\right]}{\Gamma_{p_j}\left[\frac{1}{2}(N_{(1\ldots k-j+1)}^{(\ell)} - p_{(1\ldots j-1)} - 1)\right]}\\
&\times \frac{\Gamma\left[\frac{1}{2}\sum_{j=1}^{k}(N_{(1\ldots k-j+1)} - m)p_j\right]}{\Gamma\left[\frac{1}{2}\sum_{j=1}^{k}(N_{(1\ldots k-j+1)} - m)p_j + \frac{1}{2}\sum_{j=1}^{k} N_{(1\ldots k-j+1)}p_j h\right]},
\end{aligned}
$$

where we define $p_{(1\ldots j-1)} = 0$ when $j = 1$.

This theorem is an extension of **Theorem 5**. The h-th null moment of $\lambda_{(m)}^{(k)}$ can be transformed as follows.

$$
\mathrm{E}[(\lambda_{(m)}^{(k)})^h] = K\left[\frac{\left(\frac{1}{2}\sum_{j=1}^{k} N_{(1\ldots k-j+1)}p_j\right)^{\frac{1}{2}\sum_{j=1}^{k} N_{(1\ldots k-j+1)}p_j}}{\prod_{\ell=1}^{m}\prod_{j=1}^{k}\prod_{i=1}^{p_j}\left(\frac{1}{2}N_{(1\ldots k-j+1)}^{(\ell)}\right)^{\frac{N_{(1\ldots k-j+1)}^{(\ell)}}{2}}}\right]^h
$$

$$\times \frac{\displaystyle\prod_{\ell=1}^{m}\prod_{j=1}^{k}\prod_{i=1}^{p_j}\Gamma\left[\frac{1}{2}N_{(1\ldots k-j+1)}^{(\ell)}(1+h)-\frac{1}{2}(i+p_{(1\ldots j-1)})\right]}{\Gamma\left[\displaystyle\frac{1}{2}\left(\sum_{j=1}^{k}N_{(1\ldots k-j+1)}p_j\right)(1+h)-\frac{1}{2}mp\right]}, \tag{30}$$

where

$$K=\frac{\Gamma\left[\displaystyle\frac{1}{2}\sum_{j=1}^{k}(N_{(1\ldots k-j+1)}-m)p_j\right]}{\displaystyle\prod_{\ell=1}^{m}\prod_{j=1}^{k}\prod_{i=1}^{p_j}\Gamma\left[\frac{1}{2}(N_{(1\ldots k-j+1)}^{(\ell)}-p_{(1\ldots j-1)}-i)\right]}.$$

By comparing (30) with (26), we obtain the following equations.

For $j=1,\ldots,k,\ell=1,\ldots,m$,

$$r=1, q_j=p_j, z_{ji}^{(\ell)}=\frac{1}{2}N_{(1\ldots k-j+1)}^{(\ell)}, \xi_{ji}=-\frac{1}{2}(i+p_{(1\ldots j-1)}), i=1,\ldots,p_j,$$
$$y_1=\frac{1}{2}\sum_{j=1}^{k}N_{(1\ldots k-j+1)}p_j, \eta_1=-\frac{1}{2}mp,$$

where the above equations satisfy (27). By substituting these equations into (28) and (29), the proof of **Theorem 2** is complete.

Appendix.C

Subtracting $\mathrm{MSE}[\widetilde{\sigma}^2]$ from $\mathrm{MSE}[\widetilde{\sigma}_{\mathrm{U}}^2]$ yields the following result.

$$\mathrm{MSE}[\widetilde{\sigma}_{\mathrm{U}}^2]-\mathrm{MSE}[\widetilde{\sigma}^2]=\frac{mp\left\{mp(mp-2)-(mp-4)N\displaystyle\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right\}}{\left\{N\left(\displaystyle\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right)-mp\right\}\left(N\displaystyle\sum_{j=1}^{k}\gamma_{(1\ldots k-j+1)}p_j\right)^2}\sigma^4,$$

where $\mathrm{MSE}[\widetilde{\sigma}_{\mathrm{U}}^2]-\mathrm{MSE}[\widetilde{\sigma}^2]$ tends to zero as $N_{(1\ldots k-j+1)}$ and N tend to infinity. Therefore, we obtain (10) and establish the equality relations.

References

[1] Anderson, T. W. (2003). *An introduction to multivariate statistical analysis, third ed.,* John Wiley & Sons Inc., New York.

[2] Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, **36**, 317–346.

[3] Chang, W. -Y. and Richards, D. St. P. (2010). Finite-sample inference with monotone incomplete multivariate normal data, II. *Journal of Multivariate Analysis*, **101**, 603–620.

[4] Flury, B. (1997). *A first course in multivariate statistics*, Springer Texts in Statistics, Springer-Verlag, New York.

[5] Kanda, T. and Fujikoshi, Y. (1998). Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data. *American Journal of Mathematical and Management Sciences*, **18**, 161–190.

[6] Mendoza, J. L. (1980). A significance test for multisample sphericity. *Psychometrika*, **45**, No.4, 495–498.

[7] Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*, John Wiley & Sons Inc., New York.

[8] Sato, T., Yagi, A., and Seo, T. (2025). Sphericity test on variance-covariance matrix with monotone missing data. *Journal of Statistical Theory and Practice*, **19**, 16.

[9] Tsukada, S. (2014). Equivalence testing of mean vector and covariance matrix for multi-populations under a two-step monotone incomplete sample. *Journal of Multivariate Analysis*, **132**, 183–196.

[10] Tsukada, S. (2024). Hypothesis testing for mean vector and covariance matrix of multi-populations under a two-step monotone incomplete sample in large sample and dimension. *Journal of Multivariate Analysis*, **202**, 105290.