



Statistical hypothesis test for some coefficients of the linear discriminant function with two-step monotone missing data

岡田, 悠希

(Degree)

博士 (理学)

(Date of Degree)

2025-03-25

(Date of Publication)

2026-03-01

(Resource Type)

doctoral thesis

(Report Number)

甲第9138号

(URL)

<https://hdl.handle.net/20.500.14094/0100496419>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



博 士 論 文

Statistical hypothesis test for some coefficients of
the linear discriminant function
with two-step monotone missing data
(2-ステップ単調欠測データで構成される
線形判別関数の係数に対する統計的仮説検定)

令和 7 年 1 月

神戸大学大学院理学研究科

岡田 悠希
(Yuki OKADA)

Acknowledgments

I want to express my supervisor Professor Nobumichi Shutoh, who persistently guided me until the very end, when the results of my research were not readily available. I would also like to reatly indebted to Professors Satoshi Aoki, Katusi Fukuyama, and Takashi Taniguchi for their careful readings. In addition, I am greatly indebted to Professor Masashi Hyodo and Mr. Naoki Matsuchi for their insightful advice and constructive comments. Lastly, I wish to express my deepest appreciation to my family in Oita and Osaka for their steadfast support throughout the course of this research.

Contents

1	Introduction	1
2	MLEs with two-step monotone missing data and LRT statistic	3
3	Asymptotic expansion and Bartlett correction of LRT statistic	14
4	Simulation studies	21
5	Conclusion	25
A	Lemmas and corollary	26

1 Introduction

Incomplete data is a major issue in practical data analysis. In the datasets observed in many research areas, e.g., social science, medicine, and epidemiology, we encounter incomplete data because missing observations exist due to reasons such as participant refusal to answer or omission from the survey. Especially in the medical research area, it is difficult to obtain sufficiently large sample sizes due to the invasive measurements. In this situation, variable selection is also one of the crucial problems in terms of reducing the cost of making the estimation. In existing studies, Anderson and Olkin [2] derived maximum likelihood estimators (MLEs) of mean vector $\boldsymbol{\mu}$ and covariance matrix Σ based on monotone missing data drawn from $\Pi : N_p(\boldsymbol{\mu}, \Sigma)$. Kanda and Fujikoshi [7] gave some distributional properties of MLEs. By making use of the results derived by [2, 7], statistical inferences based on incomplete data are considered by several authors. For example, Krishnamoorthy and Pannala [8] derived simplified Hotelling's T^2 statistics for the mean vector with two-step monotone missing data in one sample problem and extended the result to the general monotone case. Yagi and Seo [13] proposed Bartlett correction for the statistic with general monotone missing data. Also, Chang and Richards [3] gave stochastic representation for exact distribution of $\hat{\boldsymbol{\mu}}$, MLE of $\boldsymbol{\mu}$ in finite-sample, and generalization of Hotelling's T^2 statistic. Moreover, they derived asymptotic distribution of probability inequalities for T^2 statistic under complete and incomplete samples. For two sample problems, Yu et al. [14] discussed the approximate distributions for the simplified Hotelling's T^2 -type test statistic. In discriminant analysis, Shutoh et al. [12] focused on two-group discriminant analysis and derived an asymptotic approximation for the probability of misclassification based on linear discriminant function (LDF) with two-step monotone missing data. Also, Shutoh [10] extended the result derived by [12] to general monotone missing data. Moreover, Shutoh [11] gave asymptotic expansions for the distribution of linear discriminant function up to the first-order term with respect to N^{-1} , where N is the total sample size. For the variable selection of discriminant analysis, the criteria based on error probability misclassification (EPMC) rate and Akaike's information criteria (AIC) are discussed (see [1, 6], etc.). Fujikoshi [4] showed that criteria based on the estimated error rate of misclassification and AIC for linear discriminant analysis (LDA) are asymptotically equivalent for two groups with multivariate normality. However, the variable selection problem in discriminant analysis based on missing data is not sufficiently discussed. In this paper, we consider a redundancy test for variables

in LDA for two groups with two-step monotone missing data and suggest the corrected testing procedure via Bartlett correction. Henceforth, we assume that two-step monotone missing data drawn from two groups $\Pi^{(g)} : N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$, ($g \in \{1, 2\}$) that is of the form

$$\begin{pmatrix} \mathbf{x}_{11}^{(g)} & \mathbf{x}_{12}^{(g)} & \cdots & \mathbf{x}_{1N_1^{(g)}}^{(g)} & \mathbf{x}_{1N_1^{(g)}+1}^{(g)} & \cdots & \mathbf{x}_{1N^{(g)}}^{(g)} \\ \mathbf{x}_{21}^{(g)} & \mathbf{x}_{22}^{(g)} & \cdots & \mathbf{x}_{2N_1^{(g)}}^{(g)} & \mathbf{x}_{2N_1^{(g)}+1}^{(g)} & \cdots & \mathbf{x}_{2N^{(g)}}^{(g)} \\ \mathbf{x}_{31}^{(g)} & \mathbf{x}_{32}^{(g)} & \cdots & \mathbf{x}_{3N_1^{(g)}}^{(g)} & * & \cdots & * \end{pmatrix},$$

$$\mathbf{x}_{(12)j}^{(g)} = \left(\mathbf{x}_{1j}^{(g)'} \quad \mathbf{x}_{2j}^{(g)'} \right)', \quad \mathbf{x}_{1j}^{(g)} = \left(x_{1j}^{(g)} \quad \cdots \quad x_{p_1j}^{(g)} \right)',$$

$$\mathbf{x}_{2j}^{(g)} = \left(x_{p_1+1j}^{(g)} \quad \cdots \quad x_{p_1+p_2j}^{(g)} \right)', \quad \mathbf{x}_{3j}^{(g)} = \left(x_{p_1+p_2+1j}^{(g)} \quad \cdots \quad x_{pj}^{(g)} \right)',$$

where $j \in \{1, 2, \dots, N^{(g)}\}$, $N^{(g)} = N_1^{(g)} + N_2^{(g)}$ and $p = p_1 + p_2 + p_3$, and the symbol “*” indicates missing observations. We have $N_1^{(g)}$ observations with p dimensions and incomplete data for $N_2^{(g)}$ observations with $(p_1 + p_2)$ -dimensions. Then, $\mathbf{x}_1^{(g)}, \dots, \mathbf{x}_{N_1^{(g)}}^{(g)}$ are independently and identically distributed as $N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$ and $\mathbf{x}_{(12)N_1^{(g)}+1}^{(g)}, \dots, \mathbf{x}_{(12)N^{(g)}}^{(g)}$ are independently and identically distributed as $N_{p_1+p_2}(\boldsymbol{\mu}_{(12)}^{(g)}, \Sigma_{(12)})$, where

$$\boldsymbol{\mu}^{(g)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_2^{(g)} \\ \boldsymbol{\mu}_3^{(g)} \end{pmatrix}, \quad \boldsymbol{\mu}_{(12)}^{(g)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_2^{(g)} \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}, \quad \Sigma_{(12)} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

In these settings, for a set of the coefficients of linear discriminant function, we consider the redundancy hypothesis testing as follows:

$$H_0 : \boldsymbol{\eta}_2^{(1)} = \boldsymbol{\eta}_2^{(2)} \text{ and } \boldsymbol{\eta}_3^{(1)} = \boldsymbol{\eta}_3^{(2)}, \quad H_1 : \text{not } H_0$$

where

$$\begin{aligned} \boldsymbol{\eta}_2^{(g)} &= \boldsymbol{\mu}_2^{(g)} - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1^{(g)}, \quad \boldsymbol{\eta}_3^{(g)} = \boldsymbol{\mu}_3^{(g)} - \Sigma_{3(12)}\Sigma_{(12)}^{-1}\boldsymbol{\mu}_{(12)}^{(g)}, \\ \boldsymbol{\mu}_i^{(g)} &= C_i\boldsymbol{\mu}^{(g)}, \quad \Sigma_{ii'} = C_i\Sigma C_{i'}, \quad ((i, i') \in \{1, 2, 3\}^2) \\ C_1 &= (I_{p_1} \quad O_{p_1 \times p_2} \quad O_{p_1 \times p_3}), \quad C_2 = (O_{p_2 \times p_1} \quad I_{p_2} \quad O_{p_2 \times p_3}), \\ C_3 &= (O_{p_3 \times p_1} \quad O_{p_3 \times p_2} \quad I_{p_3}), \quad \Sigma_{3(12)} = (\Sigma_{31} \quad \Sigma_{32}) = \Sigma'_{(12)3}, \quad \boldsymbol{\mu}_{(12)}^{(g)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_2^{(g)} \end{pmatrix}. \end{aligned}$$

Under H_0 , linear discriminant function W_0 with parameters is

$$\begin{aligned} W_0 &= \left(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)} \right)' \Sigma^{-1} \left\{ \mathbf{x} - \frac{1}{2} \left(\boldsymbol{\mu}^{(1)} + \boldsymbol{\mu}^{(2)} \right) \right\} \\ &= \left(\boldsymbol{\mu}_1^{(1)} - \boldsymbol{\mu}_1^{(2)} \right)' \Sigma_{11}^{-1} \left\{ \mathbf{x}_1 - \frac{1}{2} \left(\boldsymbol{\mu}_1^{(1)} + \boldsymbol{\mu}_1^{(2)} \right) \right\} \end{aligned}$$

then $\mathbf{x}$ is a p -dimensional sample vector to be discriminated and $\mathbf{x}_1 = C_1 \mathbf{x}$. Therefore, satisfying H_0 implies that $(p_2 + p_3)$ -dimensions for the coefficients of $\mathbf{x}$ are redundant for discriminant analysis.

In this paper, we derive the likelihood ratio test (LRT) statistic for H_0 in Chapter 2. However, when $N^{(g)}$ is not so large, the type I error of the LRT exceeds the nominal significant level α . Therefore, we proposed the corrected testing procedure for H_0 as a main result in Chapter 3. We perform simulation studies to evaluate our proposed testing procedure in Chapter 4. Finally, we consider this paper and address future problems in Chapter 5. The auxiliary results to derive the main result are stated in Appendix A.

2 MLEs with two-step monotone missing data and LRT statistic

In this chapter, we obtain the MLEs using the decomposition of the density into conditional densities presented in [7]. The joint density function of the observations

$\mathbf{x}_1^{(g)}, \mathbf{x}_2^{(g)}, \dots, \mathbf{x}_{N_1^{(g)}}^{(g)}, \mathbf{x}_{(12)N_1^{(g)}+1}^{(g)}, \dots, \mathbf{x}_{(12)N^{(g)}}^{(g)}$ can be written as

$$\prod_{j=1}^{N_1^{(g)}} f_1 \left(\mathbf{x}_j^{(g)} \right) \times \prod_{j=N_1^{(g)}+1}^{N^{(g)}} f_2 \left(\mathbf{x}_{(12)j}^{(g)} \right)$$

where

$$\begin{aligned} f_1 \left(\mathbf{x}_j^{(g)} \right) &= (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left(\mathbf{x}_j^{(g)} - \boldsymbol{\mu}^{(g)} \right)' \Sigma^{-1} \left(\mathbf{x}_j^{(g)} - \boldsymbol{\mu}^{(g)} \right) \right\}, \\ f_2 \left(\mathbf{x}_{(12)j}^{(g)} \right) &= (2\pi)^{-\frac{p_1+p_2}{2}} |\Sigma_{(12)}|^{-\frac{1}{2}} \\ &\quad \times \exp \left\{ -\frac{1}{2} \left(\mathbf{x}_{(12)j}^{(g)} - \boldsymbol{\mu}_{(12)}^{(g)} \right)' \Sigma_{(12)}^{-1} \left(\mathbf{x}_{(12)j}^{(g)} - \boldsymbol{\mu}_{(12)}^{(g)} \right) \right\}. \end{aligned}$$

The log-likelihood function can be given as follows:

$$\begin{aligned} \ell(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \Sigma) &= Const. - \frac{1}{2} \sum_{g=1}^2 \left(N_1^{(g)} \log |\Sigma| + N_2^{(g)} \log |\Sigma_{(12)}| \right) \\ &\quad - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_j^{(g)} - \boldsymbol{\mu}^{(g)} \right)' \Sigma^{-1} \left(\mathbf{x}_j^{(g)} - \boldsymbol{\mu}^{(g)} \right) \end{aligned}$$

$$-\frac{1}{2} \sum_{g=1}^2 \sum_{j=N_1^{(g)}+1}^{N^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \boldsymbol{\mu}_{(12)}^{(g)} \right)' \Sigma_{(12)}^{-1} \left(\mathbf{x}_{(12)j}^{(g)} - \boldsymbol{\mu}_{(12)}^{(g)} \right).$$

We multiply observation vectors $\mathbf{x}_j^{(g)}$ by a transform matrix on the left side

$$P = \begin{pmatrix} P_{(12)} & O_{(p_1+p_2) \times p_3} \\ -\Sigma_{3(12)} \Sigma_{(12)}^{-1} & I_{p_3} \end{pmatrix}, \quad P_{(12)} = \begin{pmatrix} I_{p_1} & O_{p_1 \times p_2} \\ -\Sigma_{21} \Sigma_{11}^{-1} & I_{p_2} \end{pmatrix}.$$

The transformed observation vectors are

$$\mathbf{y}_j^{(g)} = P \mathbf{x}_j^{(g)} = \begin{pmatrix} \mathbf{x}_{1j}^{(g)} \\ \mathbf{x}_{2j}^{(g)} - \Psi_{21} \mathbf{x}_{1j}^{(g)} \\ \mathbf{x}_{3j}^{(g)} - \Psi_{3(12)} \mathbf{x}_{(12)j}^{(g)} \end{pmatrix} \sim N_p \left(\boldsymbol{\eta}^{(g)}, \Psi \right),$$

$$\mathbf{y}_{(12)j}^{(g)} = P_{(12)} \mathbf{x}_{(12)j}^{(g)} = \begin{pmatrix} \mathbf{x}_{1j}^{(g)} \\ \mathbf{x}_{2j}^{(g)} - \Psi_{21} \mathbf{x}_{1j}^{(g)} \end{pmatrix} \sim N_{p_1+p_2} \left(\boldsymbol{\eta}_{(12)}^{(g)}, \Psi_{(12)} \right),$$

where

$$\boldsymbol{\eta}^{(g)} = P \boldsymbol{\mu}^{(g)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_2^{(g)} - \Psi_{21} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_3^{(g)} - \Psi_{3(12)} \boldsymbol{\mu}_{(12)}^{(g)} \end{pmatrix}, \quad \Psi = \begin{pmatrix} \Psi_{(12)} & \Psi_{(12)3} \\ \Psi_{3(12)} & \Psi_{33} \end{pmatrix},$$

$$\Psi_{(12)} = \begin{pmatrix} \Sigma_{11} & \Sigma_{11}^{-1} \Sigma_{12} \\ \Sigma_{21} \Sigma_{11}^{-1} & \Sigma_{22 \cdot 1} \end{pmatrix}, \quad \Sigma_{22 \cdot 1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12},$$

$$\Psi_{3(12)} = \begin{pmatrix} \Psi_{31} & \Psi_{32} \end{pmatrix} = \Psi'_{(12)3} = \Sigma_{3(12)} \Sigma_{(12)}^{-1}, \quad \Psi_{33} = \Sigma_{33} - \Sigma_{3(12)} \Sigma_{(12)}^{-1} \Sigma_{(12)3}.$$

Also, the transformed log-likelihood function from $(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \Sigma)$ to $(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)$ can be written as

$$\begin{aligned} \ell \left(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi \right) &= \text{Const.} - \frac{1}{2} \sum_{g=1}^2 \left(N^{(g)} \log |\Psi_{11}| + N^{(g)} \log |\Psi_{22}| + N_1^{(g)} \log |\Psi_{33}| \right) \\ &\quad - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left\{ \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right)' \Psi_{11}^{-1} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right) \right\} \\ &\quad - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left\{ \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right)' \Psi_{22}^{-1} \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right) \right\} \\ &\quad - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left\{ \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right)' \Psi_{33}^{-1} \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right) \right\}. \end{aligned}$$

It should be noted that $(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \Sigma)$ and $(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)$ are one-to-one correspondence. We give the following theorems to construct the LRT statistic using a similar derivation [7].

Theorem 2.1. Suppose that two-step monotone missing data

$\mathbf{x}_1^{(g)}, \dots, \mathbf{x}_{N_1^{(g)}}^{(g)}, \mathbf{x}_{(12)N_1^{(g)}+1}^{(g)}, \dots, \mathbf{x}_{(12)N^{(g)}}^{(g)}$ is observed from $\Pi^{(g)} : N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$ ($g \in \{1, 2\}$). Then, the MLEs denoted as $(\hat{\boldsymbol{\eta}}^{(1)}, \hat{\boldsymbol{\eta}}^{(2)}, \hat{\Psi})$ are given by

$$\begin{aligned}\hat{\boldsymbol{\eta}}^{(g)} &= \begin{pmatrix} \hat{\boldsymbol{\eta}}_1^{(g)} \\ \hat{\boldsymbol{\eta}}_2^{(g)} \\ \hat{\boldsymbol{\eta}}_3^{(g)} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{x}}_1^{(g)} \\ \bar{\mathbf{x}}_2^{(g)} - \hat{\Psi}_{21}\bar{\mathbf{x}}_1^{(g)} \\ \bar{\mathbf{x}}_3^{(g,1)} - \hat{\Psi}_{3(12)}\bar{\mathbf{x}}_{(12)}^{(g,1)} \end{pmatrix}, \\ \hat{\Psi}_{11} &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_1^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_1^{(g)} \right)', \\ \hat{\Psi}_{21} &= \hat{\Psi}_{12}' = \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \bar{\mathbf{x}}_2^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1^{(g)} \right)' \right\} \hat{\Psi}_{11}^{-1}, \\ \hat{\Psi}_{22} &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \hat{\Psi}_{21}\mathbf{x}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_2^{(g)} \right) \left(\mathbf{x}_{2j}^{(g)} - \hat{\Psi}_{21}\mathbf{x}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_2^{(g)} \right)', \\ \hat{\Psi}_{3(12)} &= \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(g,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,1)} \right)' \right\} \\ &\quad \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,1)} \right)' \right\}^{-1}, \\ \hat{\Psi}_{33} &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \hat{\Psi}_{3(12)}\mathbf{x}_{(12)j}^{(g)} - \hat{\boldsymbol{\eta}}_3^{(g)} \right) \\ &\quad \left(\mathbf{x}_{3j}^{(g)} - \hat{\Psi}_{3(12)}\mathbf{x}_{(12)j}^{(g)} - \hat{\boldsymbol{\eta}}_3^{(g)} \right)',\end{aligned}$$

where

$$\bar{\mathbf{x}}_{(12)}^{(g)} = \frac{1}{N^{(g)}} \left(N_1^{(g)} \bar{\mathbf{x}}_{(12)}^{(g,1)} + N_2^{(g)} \bar{\mathbf{x}}_{(12)}^{(g,2)} \right), \quad \bar{\mathbf{x}}_k^{(g)} = B_k \bar{\mathbf{x}}_{(12)}^{(g)} \quad (k \in \{1, 2\}),$$

$$\bar{\mathbf{x}}^{(g,1)} = \frac{1}{N_1^{(g)}} \sum_{j=1}^{N_1^{(g)}} \mathbf{x}_j^{(g)}, \quad \bar{\mathbf{x}}_i^{(g,1)} = C_i \bar{\mathbf{x}}^{(g,1)}, \quad \bar{\mathbf{x}}_{(12)}^{(g,1)} = C_{(12)} \bar{\mathbf{x}}^{(g,1)},$$

$$\bar{\mathbf{x}}_{(12)}^{(g,2)} = \frac{1}{N_2^{(g)}} \sum_{j=N_1^{(g)}+1}^{N^{(g)}} \mathbf{x}_{(12)j}^{(g)}, \quad \bar{\mathbf{x}}_k^{(g,2)} = B_k \bar{\mathbf{x}}_{(12)}^{(g,2)},$$

$$B_1 = \begin{pmatrix} I_{p_1} & O_{p_1 \times p_2} \end{pmatrix}, \quad B_2 = \begin{pmatrix} O_{p_2 \times p_1} & I_{p_2} \end{pmatrix}, \quad C_{(12)} = \begin{pmatrix} C_1' & C_2' \end{pmatrix}'.$$

Proof. Using Lemma A.1, we obtain

$$\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\eta}_1^{(g)}} = \sum_{j=1}^{N^{(g)}} \boldsymbol{\Psi}_{11}^{-1} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right).$$

The MLEs of $\boldsymbol{\eta}_1^{(1)}$ and $\boldsymbol{\eta}_1^{(2)}$ can be obtained by solving $\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\eta}_1^{(g)}} = \mathbf{0}_{p_1}$:

$$\hat{\boldsymbol{\eta}}_1^{(g)} = \frac{1}{N^{(g)}} \sum_{j=1}^{N^{(g)}} \mathbf{y}_{1j}^{(g)} = \frac{1}{N^{(g)}} \left(\sum_{j=1}^{N_1^{(g)}} \mathbf{x}_{1j}^{(g)} + \sum_{j=N_1^{(g)}+1}^{N^{(g)}} \mathbf{x}_{1j}^{(g)} \right) = \bar{\mathbf{x}}_1^{(g)}.$$

Furthermore, it holds that

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\Psi}_{11}^{-1}} &= \frac{1}{2} (N^{(1)} + N^{(2)}) (2\boldsymbol{\Psi}_{11} - \text{diag} \boldsymbol{\Psi}_{11}) \\ &\quad - \frac{1}{2} \sum_{g=1}^2 \left\{ \sum_{j=1}^{N^{(g)}} 2 \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right) \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right)' \right. \\ &\quad \left. - \text{diag} \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right) \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right)' \right\}. \end{aligned} \quad (2.1)$$

Putting $\hat{\boldsymbol{\eta}}_1^{(g)}$ into (2.1), by solving $\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\Psi}_{11}^{-1}} = \mathbf{0}_{p_1 \times p_1}$, we obtain

$$\begin{aligned} \hat{\boldsymbol{\Psi}}_{11} &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_1^{(g)} \right) \left(\mathbf{y}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_1^{(g)} \right)' \\ &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_1^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \hat{\boldsymbol{\eta}}_1^{(g)} \right)'. \end{aligned}$$

We find the MLEs of $\boldsymbol{\eta}_2^{(g)}$ and $\boldsymbol{\Psi}_{21}$ by solving the partial differential equation

$$\begin{cases} \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\eta}_2^{(g)}} = \mathbf{0}_{p_2} \\ \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\Psi}_{21}} = \mathbf{0}_{p_2 \times p_1}. \end{cases}$$

Using Lemma A.1,

$$\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \boldsymbol{\Psi})}{\partial \boldsymbol{\eta}_2^{(g)}} = \sum_{j=1}^{N^{(g)}} \boldsymbol{\Psi}_{22}^{-1} \left(\mathbf{y}_2^{(g)} - \boldsymbol{\eta}_2^{(g)} \right)$$

Thus,

$$\begin{aligned}
\widehat{\boldsymbol{\eta}}_2^{(g)} &= \frac{1}{N^{(g)}} \sum_{j=1}^{N^{(g)}} \mathbf{y}_{2j}^{(g)} \\
&= \frac{1}{N^{(g)}} \sum_{j=1}^{N^{(g)}} \mathbf{x}_{2j}^{(g)} - \widehat{\Psi}_{21} \left(\frac{1}{N^{(g)}} \sum_{j=1}^{N^{(g)}} \mathbf{x}_{1j}^{(g)} \right) \\
&= \overline{\mathbf{x}}_2^{(g)} - \widehat{\Psi}_{21} \overline{\mathbf{x}}_1^{(g)}.
\end{aligned}$$

Putting $\widehat{\boldsymbol{\eta}}_2^{(g)}$ into $\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{21}} = O_{p_2 \times p_1}$,

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{21}} &= \Psi_{22}^{-1} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left\{ \left(\mathbf{x}_{2j}^{(g)} - \overline{\mathbf{x}}_2^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \overline{\mathbf{x}}_1^{(g)} \right)' \right. \\
&\quad \left. - \Psi_{21} \left(\mathbf{x}_{1j}^{(g)} - \overline{\mathbf{x}}_1^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \overline{\mathbf{x}}_1^{(g)} \right)' \right\}.
\end{aligned}$$

therefore,

$$\begin{aligned}
\widehat{\Psi}_{21} &= \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \overline{\mathbf{x}}_2^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \overline{\mathbf{x}}_1^{(g)} \right)' \right\} \\
&\quad \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{1j}^{(g)} - \overline{\mathbf{x}}_1^{(g)} \right) \left(\mathbf{x}_{1j}^{(g)} - \overline{\mathbf{x}}_1^{(g)} \right)' \right\}^{-1}.
\end{aligned}$$

From Lemma A.1, it holds that

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{22}} &= \frac{1}{2} N (2\Psi_{22} - \text{diag} \Psi_{22}) \\
&\quad - \frac{1}{2} \sum_{g=1}^2 \left\{ 2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right) \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right)' \right. \\
&\quad \left. - \text{diag} \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right) \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right)' \right\}.
\end{aligned}$$

Thus, we obtain by substituting $\widehat{\boldsymbol{\eta}}_2^{(g)}$ and $\widehat{\Psi}_{21}$,

$$\begin{aligned}
\widehat{\Psi}_{22} &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right) \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2^{(g)} \right)' \\
&= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \widehat{\Psi}_{21} \mathbf{x}_{1j}^{(g)} \right) \left(\mathbf{x}_{2j}^{(g)} - \widehat{\Psi}_{21} \mathbf{x}_{1j}^{(g)} \right)'.
\end{aligned}$$

In the same way, solving the partial differential equations

$$\begin{cases} \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \boldsymbol{\eta}_3^{(g)}} = \mathbf{0}_{p_3} \\ \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{3(12)}} = O_{p_3 \times (p_1 + p_2)} \end{cases}$$

we obtain

$$\hat{\boldsymbol{\eta}}_3^{(g)} = \frac{1}{N_1^{(g)}} \sum_{j=1}^{N_1^{(g)}} \mathbf{y}_{3j}^{(g)} = \bar{\mathbf{x}}_3^{(g,1)} - \hat{\Psi}_{3(12)} \bar{\mathbf{x}}_{(12)}^{(g,1)}.$$

Substituting $\hat{\boldsymbol{\eta}}_3^{(g)}$ into $\frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{3(12)}}$ and using Lemma A.1, it holds that

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{3(12)}} &= -\frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \Psi_{33}^{-1} \left\{ -2 \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(g)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g)} \right) \right. \\ &\quad \left. + 2 \Psi_{3(12)} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g)} \right)' \right\} \\ &= \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \Psi_{33}^{-1} \left\{ \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(g)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g)} \right)' \right. \\ &\quad \left. - \Psi_{3(12)} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g)} \right)' \right\}. \end{aligned}$$

Thus, we obtain

$$\begin{aligned} \hat{\Psi}_{3(12)} &= \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(g,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,1)} \right)' \right\} \\ &\quad \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,1)} \right)' \right\}^{-1}. \end{aligned}$$

Also, it holds that

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \Psi)}{\partial \Psi_{33}^{-1}} &= \frac{1}{2} N_1 (2 \Psi_{33} - \text{diag} \Psi_{33}) \\ &\quad - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left\{ 2 \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right) \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right)' \right. \\ &\quad \left. - 2 \text{diag} \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right) \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right)' \right\}. \end{aligned}$$

therefore,

$$\begin{aligned}\widehat{\Psi}_{33} &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right) \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3^{(g)} \right)' \\ &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \widehat{\Psi}_{3(12)} \mathbf{x}_{(12)j}^{(g)} - \widehat{\boldsymbol{\eta}}_3^{(g)} \right) \left(\mathbf{x}_{3j}^{(g)} - \widehat{\Psi}_{3(12)} \mathbf{x}_{(12)j}^{(g)} - \widehat{\boldsymbol{\eta}}_3^{(g)} \right)'.\end{aligned}$$

□

Next, we derive MLEs under $H_0 : \boldsymbol{\eta}_2^{(1)} = \boldsymbol{\eta}_2^{(2)} = \boldsymbol{\eta}_2, \boldsymbol{\eta}_3^{(1)} = \boldsymbol{\eta}_3^{(2)} = \boldsymbol{\eta}_3$ in order to obtain LRT statistic.

Theorem 2.2. Suppose the two-step monotone missing data

$\mathbf{x}_1^{(g)}, \dots, \mathbf{x}_{N_1^{(g)}}^{(g)}, \mathbf{x}_{1N_1^{(g)}+1}^{(g)}, \dots, \mathbf{x}_{1N^{(g)}}^{(g)}$ is observed from $\Pi^{(g)} : N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$ ($g \in \{1, 2\}$).

Then the MLEs under the null hypothesis H_0 denoted as $\widetilde{\boldsymbol{\eta}}^{(1)}, \widetilde{\boldsymbol{\eta}}^{(2)}, \widetilde{\Psi}$ are given by

$$\begin{aligned}\widetilde{\boldsymbol{\eta}}^{(g)} &= \begin{pmatrix} \widetilde{\boldsymbol{\eta}}_1^{(g)} \\ \widetilde{\boldsymbol{\eta}}_2 \\ \widetilde{\boldsymbol{\eta}}_3 \end{pmatrix} = \begin{pmatrix} \widehat{\boldsymbol{\eta}}_1^{(g)}, \\ \bar{\mathbf{x}}_2 - \widetilde{\Psi}_{21} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_3^{(\cdot,1)} - \widetilde{\Psi}_{3(12)} \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \end{pmatrix}, \quad \widetilde{\Psi}_{11} = \widehat{\Psi}_{11}, \\ \widetilde{\Psi}_{21} &= \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \bar{\mathbf{x}}_2 \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right)' \right\} \\ &\quad \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right)' \right\}^{-1}, \\ \widetilde{\Psi}_{22} &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \widetilde{\Psi}_{21} \mathbf{x}_{1j}^{(g)} - \widetilde{\boldsymbol{\eta}}_2 \right) \left(\mathbf{x}_{2j}^{(g)} - \widetilde{\Psi}_{21} \mathbf{x}_{1j}^{(g)} - \widetilde{\boldsymbol{\eta}}_2 \right)', \\ \widetilde{\Psi}_{3(12)} &= \widetilde{\Psi}'_{(12)3} = \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(\cdot,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right)' \right\} \\ &\quad \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right)' \right\}^{-1}, \\ \widetilde{\Psi}_{33} &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \widetilde{\Psi}_{3(12)} \mathbf{x}_{(12)j}^{(g)} - \widetilde{\boldsymbol{\eta}}_3 \right) \left(\mathbf{x}_{3j}^{(g)} - \widetilde{\Psi}_{3(12)} \mathbf{x}_{(12)j}^{(g)} - \widetilde{\boldsymbol{\eta}}_3 \right)',\end{aligned}$$

where

$$\bar{\mathbf{x}}_{(12)} = \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \mathbf{x}_{(12)j}^{(g)}, \quad \bar{\mathbf{x}}_k = B_k \bar{\mathbf{x}}_{(12)}, \quad N_1 = \sum_{g=1}^2 N_1^{(g)}, \quad N = \sum_{g=1}^2 \sum_{\ell=1}^2 N_\ell^{(g)},$$

$$\bar{\mathbf{x}}^{(\cdot,1)} = \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \mathbf{x}_j^{(g)}, \quad \bar{\mathbf{x}}_i^{(\cdot,1)} = C_i \bar{\mathbf{x}}^{(\cdot,1)}, \quad \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} = C_{(12)} \bar{\mathbf{x}}^{(\cdot,1)}.$$

Proof. Under $H_0 : \boldsymbol{\eta}_2^{(1)} = \boldsymbol{\eta}_2^{(2)} = \boldsymbol{\eta}_2, \boldsymbol{\eta}_3^{(1)} = \boldsymbol{\eta}_3^{(2)} = \boldsymbol{\eta}_3$, the transformed log-likelihood function can be written as

$$\begin{aligned} \ell_{H_0} = & \text{Const.} - \frac{1}{2} \sum_{g=1}^2 \left(N^{(g)} \log |\Psi_{11}| + N^{(g)} \log |\Psi_{22}| + N_1^{(g)} \log |\Psi_{33}| \right) \\ & - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left\{ \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right)' \Psi_{11}^{-1} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right) \right\} \\ & - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left\{ \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2 \right)' \Psi_{22}^{-1} \left(\mathbf{y}_{2j}^{(g)} - \boldsymbol{\eta}_2 \right) \right\} \\ & - \frac{1}{2} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left\{ \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3 \right)' \Psi_{33}^{-1} \left(\mathbf{y}_{3j}^{(g)} - \boldsymbol{\eta}_3 \right) \right\}. \end{aligned}$$

we have $\tilde{\boldsymbol{\eta}}_1^{(g)}, \tilde{\Psi}_{11}$ by solving the partial differential equation

$$\begin{cases} \frac{\partial \ell_{H_0}}{\partial \boldsymbol{\eta}_1^{(g)}} = \mathbf{0}_{p_1} \\ \frac{\partial \ell_{H_0}}{\partial \Psi_{11}} = O_{p_1 \times p_1}. \end{cases}$$

By Lemma A.1, we obtain

$$\begin{aligned} \frac{\partial \ell_{H_0}}{\partial \boldsymbol{\eta}_1^{(g)}} &= \sum_{j=1}^{N^{(g)}} \Psi_{11}^{-1} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right), \\ \frac{\partial \ell_{H_0}}{\partial \Psi_{11}^{-1}} &= \frac{1}{2} N (2\Psi_{11} - \text{diag} \Psi_{11}) \\ &\quad - \frac{1}{2} \left\{ 2 \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right) \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right)' \right. \\ &\quad \left. - \text{diag} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right) \left(\mathbf{y}_{1j}^{(g)} - \boldsymbol{\eta}_1^{(g)} \right)' \right\} \end{aligned}$$

therefore,

$$\tilde{\boldsymbol{\eta}}_1^{(g)} = \frac{1}{N^{(g)}} \sum_{j=1}^{N^{(g)}} \mathbf{y}_{1j}^{(g)} = \hat{\boldsymbol{\eta}}_1^{(g)},$$

$$\tilde{\Psi}_{11} = \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{1j}^{(g)} - \tilde{\boldsymbol{\eta}}_1^{(g)} \right) \left(\mathbf{y}_{1j}^{(g)} - \tilde{\boldsymbol{\eta}}_1^{(g)} \right)' = \hat{\Psi}_{11}.$$

The MLEs of $\boldsymbol{\eta}_2$, Ψ_{21} , Ψ_{22} can be given in a similar manner to proof of $\tilde{\boldsymbol{\eta}}_1^{(g)}$, $\tilde{\Psi}_{11}$ respectively. We obtain $\tilde{\boldsymbol{\eta}}_2$, $\tilde{\Psi}_{21}$, $\tilde{\Psi}_{22}$ solving the partial differential equation.

$$\begin{cases} \frac{\partial \ell_{H_0}}{\partial \boldsymbol{\eta}_2} = \mathbf{0}_{p_2} \\ \frac{\partial \ell_{H_0}}{\partial \Psi_{21}} = O_{p_2 \times p_1} \\ \frac{\partial \ell_{H_0}}{\partial \Psi_{22}^{-1}} = O_{p_2 \times p_2}. \end{cases}$$

It follows from Lemma A.1,

$$\begin{aligned} \frac{\partial \ell_{H_0}}{\partial \boldsymbol{\eta}_2} &= - \sum_{j=1}^{N^{(1)}} \Psi_{22}^{-1} \left(\mathbf{y}_{2j}^{(1)} - \boldsymbol{\eta}_2 \right) - \sum_{j=1}^{N^{(2)}} \Psi_{22}^{-1} \left(\mathbf{y}_{2j}^{(2)} - \boldsymbol{\eta}_2 \right) \\ &= -\Psi_{22}^{-1} \left(\sum_{j=1}^{N^{(1)}} \mathbf{y}_{2j}^{(1)} - N^{(1)} \boldsymbol{\eta}_2 + \sum_{j=1}^{N^{(2)}} \mathbf{y}_{2j}^{(2)} - N^{(2)} \boldsymbol{\eta}_2 \right) \\ &= -\Psi_{22}^{-1} \left(\sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \mathbf{y}_{2j}^{(g)} - N \boldsymbol{\eta}_2 \right). \end{aligned}$$

Then,

$$\begin{aligned} \tilde{\boldsymbol{\eta}}_2 &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \mathbf{y}_{2j}^{(g)} \\ &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \mathbf{x}_{2j}^{(g)} - \tilde{\Psi}_{21} \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \mathbf{x}_{1j}^{(g)} \\ &= \bar{\mathbf{x}}_2 - \tilde{\Psi}_{21} \bar{\mathbf{x}}_1. \end{aligned}$$

Substituting $\tilde{\boldsymbol{\eta}}_2$ into $\frac{\partial \ell_{H_0}}{\partial \Psi_{21}}$,

$$\frac{\partial \ell_{H_0}}{\partial \Psi_{21}} = \Psi_{22}^{-1} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left\{ \left(\mathbf{x}_{2j}^{(g)} - \bar{\mathbf{x}}_2 \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right)' - \tilde{\Psi}_{21} \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right)' \right\}.$$

Thus, we have

$$\tilde{\Psi}_{21} = \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \bar{\mathbf{x}}_2 \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right)' \right\} \left\{ \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right) \left(\mathbf{x}_{1j}^{(g)} - \bar{\mathbf{x}}_1 \right)' \right\}^{-1}.$$

Also, putting $\tilde{\boldsymbol{\eta}}_2$ and $\tilde{\Psi}_{21}$ into $\frac{\partial \ell_{H_0}}{\partial \Psi_{22}^{-1}}$,

$$\begin{aligned} \frac{\partial \ell_{H_0}}{\partial \Psi_{22}^{-1}} &= \frac{1}{2} N (2\Psi_{22} - \text{diag} \Psi_{22}) - \frac{1}{2} \left\{ 2 \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{2j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right) \left(\mathbf{y}_{2j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right)' \right. \\ &\quad \left. - \text{diag} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{2j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right) \left(\mathbf{y}_{2j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right)' \right\}, \end{aligned}$$

we obtain

$$\begin{aligned} \tilde{\Psi}_{22} &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{y}_{2j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right) \left(\mathbf{y}_{2j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right)' \\ &= \frac{1}{N} \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{2j}^{(g)} - \tilde{\Psi}_{21} \mathbf{x}_{1j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right) \left(\mathbf{x}_{2j}^{(g)} - \tilde{\Psi}_{21} \mathbf{x}_{1j}^{(g)} - \tilde{\boldsymbol{\eta}}_2 \right)'. \end{aligned}$$

In similar way, we obtain the MLEs of $\boldsymbol{\eta}_3$, $\Psi_{3(12)}$, Ψ_{33} by solving the partial differential equation

$$\begin{cases} \frac{\partial \ell_{H_0}}{\partial \boldsymbol{\eta}_3} = \mathbf{0}_{p_3} \\ \frac{\partial \ell_{H_0}}{\partial \Psi_{3(12)}} = O_{p_3 \times (p_1 + p_2)} \\ \frac{\partial \ell_{H_0}}{\partial \Psi_{33}^{-1}} = O_{p_3 \times p_3}, \end{cases}$$

respectively. It follows that

$$\begin{aligned} \frac{\partial \ell_{H_0}}{\partial \boldsymbol{\eta}_3} &= -\frac{1}{2} \left\{ 2 \sum_{j=1}^{N_1^{(1)}} \Psi_{33} \left(\mathbf{y}_{3j}^{(1)} - \boldsymbol{\eta}_3 \right) + 2 \sum_{j=1}^{N_1^{(2)}} \Psi_{33} \left(\mathbf{y}_{3j}^{(2)} - \boldsymbol{\eta}_3 \right) \right\} \\ &= -\Psi_{33}^{-1} \left(\sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \mathbf{y}_{3j}^{(g)} - N_1 \boldsymbol{\eta}_3 \right), \end{aligned}$$

and we obtain

$$\begin{aligned} \tilde{\boldsymbol{\eta}}_3 &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \mathbf{y}_{3j}^{(g)} \\ &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \mathbf{x}_{3j}^{(g)} - \tilde{\Psi}_{3(12)} \left(\frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \mathbf{x}_{(12)j}^{(g)} \right) \\ &= \bar{\mathbf{x}}_3^{(\cdot, 1)} - \tilde{\Psi}_{3(12)} \bar{\mathbf{x}}_{(12)}^{(\cdot, 1)}. \end{aligned}$$

Putting $\tilde{\boldsymbol{\eta}}_3$ into $\frac{\partial \ell_{H_0}}{\partial \Psi_{3(12)}}$, it follows that

$$\begin{aligned} \frac{\partial \ell_{H_0}}{\partial \Psi_{3(12)}} &= \Psi_{33}^{-1} \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(\cdot,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right)' \right\} \\ &\quad - \Psi_{33}^{-1} \Psi_{3(12)} \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right)' \right\}. \end{aligned}$$

We obtain

$$\begin{aligned} \tilde{\Psi}_{3(12)} &= \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \bar{\mathbf{x}}_3^{(\cdot,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right)' \right\} \\ &\quad \left\{ \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(\cdot,1)} \right)' \right\}^{-1}. \end{aligned}$$

Substituting $\tilde{\boldsymbol{\eta}}_3$ and $\tilde{\Psi}_{3(12)}$ into $\frac{\partial \ell_{H_0}}{\partial \Psi_{33}^{-1}}$,

$$\begin{aligned} \frac{\partial \ell_{H_0}}{\partial \Psi_{33}^{-1}} &= -\frac{1}{2} N_1 (2\Psi_{33} - \text{diag} \Psi_{33}) - \frac{1}{2} \left\{ 2 \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{y}_{3j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right) \left(\mathbf{y}_{3j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right)' \right. \\ &\quad \left. - \text{diag} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{y}_{3j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right) \left(\mathbf{y}_{3j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right)' \right\}. \end{aligned}$$

So, we have

$$\begin{aligned} \tilde{\Psi}_{33} &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{y}_{3j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right) \left(\mathbf{y}_{3j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right)' \\ &= \frac{1}{N_1} \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_{3j}^{(g)} - \tilde{\Psi}_{3(12)} \mathbf{x}_{(12)j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right) \left(\mathbf{x}_{3j}^{(g)} - \tilde{\Psi}_{3(12)} \mathbf{x}_{(12)j}^{(g)} - \tilde{\boldsymbol{\eta}}_3 \right)'. \end{aligned}$$

□

Using MLEs, the LRT statistic for the testing H_0 can be constructed as

$$-2 \log \lambda = NQ_1 + N_1 Q_2$$

where

$$Q_1 = \log \left\{ 1 + m^{(1)} m^{(2)} \left(\bar{\mathbf{x}}_{(12)}^{(1)} - \bar{\mathbf{x}}_{(12)}^{(2)} \right)' \left(\frac{1}{N} \hat{S} \right)^{-1} \left(\bar{\mathbf{x}}_{(12)}^{(1)} - \bar{\mathbf{x}}_{(12)}^{(2)} \right) \right\}$$

$$\begin{aligned}
& -\log \left\{ 1 + m^{(1)}m^{(2)} \left(\bar{\mathbf{x}}_1^{(1)} - \bar{\mathbf{x}}_1^{(2)} \right)' \left(\frac{1}{N} \hat{S}_{11} \right)^{-1} \left(\bar{\mathbf{x}}_1^{(1)} - \bar{\mathbf{x}}_1^{(2)} \right) \right\}, \\
Q_2 = & \log \left\{ 1 + \frac{m_1^{(1)}m_1^{(2)}}{m_1^2} \left(\bar{\mathbf{x}}^{(1,1)} - \bar{\mathbf{x}}^{(2,1)} \right)' \left(\frac{1}{N_1} S^{(\cdot,1)} \right)^{-1} \left(\bar{\mathbf{x}}^{(1,1)} - \bar{\mathbf{x}}^{(2,1)} \right) \right\} \\
& -\log \left\{ 1 + \frac{m_1^{(1)}m_1^{(2)}}{m_1^2} \left(\bar{\mathbf{x}}_{(12)}^{(1,1)} - \bar{\mathbf{x}}_{(12)}^{(2,1)} \right)' \left(\frac{1}{N_1} S_{(12)}^{(\cdot,1)} \right)^{-1} \left(\bar{\mathbf{x}}_{(12)}^{(1,1)} - \bar{\mathbf{x}}_{(12)}^{(2,1)} \right) \right\},
\end{aligned}$$

$$\begin{aligned}
\hat{S} &= \sum_{g=1}^2 \sum_{j=1}^{N^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}^{(g)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}^{(g)} \right)', \quad \hat{S}_{11} = B_1 \hat{S} B_1', \\
S^{(\cdot,1)} &= \sum_{g=1}^2 \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_j^{(g)} - \bar{\mathbf{x}}^{(g,1)} \right) \left(\mathbf{x}_j^{(g)} - \bar{\mathbf{x}}^{(g,1)} \right)', \quad S_{(12)}^{(\cdot,1)} = C_{(12)} S^{(\cdot,1)} C_{(12)}', \\
m_\ell^{(g)} &= \frac{N_\ell^{(g)}}{N} \quad (\ell \in \{1, 2\}), \quad m_1 = \frac{N_1}{N}, \quad m^{(g)} = \frac{N^{(g)}}{N}.
\end{aligned}$$

Under H_0 , the LRT statistic is asymptotically chi-squared distribution with $(p_2 + p_3)$ degrees of freedom as $N \rightarrow \infty$ and $m_\ell^{(g)} \rightarrow \gamma_\ell^{(g)} \in (0, 1)$. Therefore, an approximate testing procedure for H_0 can be considered: H_0 is rejected if $-2 \log \lambda > \chi_{p_2+p_3}^2(\alpha)$ holds, otherwise, H_0 is retained, where α denotes the nominal significance level of testing H_0 and $\chi_{p_2+p_3}^2(\alpha)$ denotes the upper 100 $\alpha\%$ point of chi-squared distribution with $(p_2 + p_3)$ degrees of freedom.

3 Asymptotic expansion and Bartlett correction of LRT statistic

In this section, we derive the asymptotic expansion of the expectation of the LRT statistic $-2 \log \lambda = NQ_1 + N_1Q_2$ up to the first order with respect to N^{-1} under H_0 to suggest Bartlett corrected terms. For this purpose, we define $\Sigma^{-\frac{1}{2}}$ and $\Sigma_{(12)}^{-\frac{1}{2}}$ as

$$\Sigma^{-\frac{1}{2}} = \begin{pmatrix} \Sigma_{11}^{-\frac{1}{2}} & O & O \\ O & \Sigma_{22 \cdot 1}^{-\frac{1}{2}} & O \\ O & O & \Sigma_{33 \cdot 12}^{-\frac{1}{2}} \end{pmatrix} P, \quad \Sigma_{(12)}^{-\frac{1}{2}} = \begin{pmatrix} \Sigma_{11}^{-\frac{1}{2}} & O \\ O & \Sigma_{22 \cdot 1}^{-\frac{1}{2}} \end{pmatrix} P_{(12)},$$

respectively. Furthermore, we also define

$$\begin{aligned}
\mathbf{z}^{(g,1)} &= \sqrt{N_1^{(g)}} \Sigma^{-\frac{1}{2}} \left(\bar{\mathbf{x}}^{(g,1)} - \boldsymbol{\mu}^{(g)} \right), \quad \mathbf{z}_i^{(g,1)} = C_i \mathbf{z}^{(g,1)}, \\
\mathbf{z}_{(12)}^{(g,2)} &= \sqrt{N_2^{(g)}} \Sigma_{(12)}^{-\frac{1}{2}} \left(\bar{\mathbf{x}}_{(12)}^{(g,2)} - \boldsymbol{\mu}_{(12)}^{(g)} \right), \quad \mathbf{z}_k^{(g,\ell)} = B_k \mathbf{z}_{(12)}^{(g,2)},
\end{aligned}$$

$$n_\ell^{(g)} = N_\ell^{(g)} - 1 \ (\ell \in \{1, 2\}), \ n^{(g)} = \sum_{\ell=1}^2 n_\ell^{(g)}, \ n = \sum_{g=1}^2 \sum_{\ell=1}^2 n_\ell^{(g)},$$

$$S^{(g,1)} = \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_j^{(g)} - \bar{\mathbf{x}}^{(g,1)} \right) \left(\mathbf{x}_j^{(g)} - \bar{\mathbf{x}}^{(g,1)} \right)', \ S_{(12)}^{(g,1)} = C_{(12)} S^{(g,1)} C'_{(12)},$$

$$S_{(12)}^{(g,2)} = \sum_{j=N_1^{(g)}+1}^{N^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,2)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g,2)} \right)',$$

$$T^{(g,1)} = \sqrt{n_1^{(g)}} \Sigma^{-\frac{1}{2}} \left(\frac{1}{n_1^{(g)}} S^{(g,1)} - \Sigma \right) \Sigma^{-\frac{1}{2}'},$$

$$T_{(12)}^{(g,\ell)} = \sqrt{n_\ell^{(g)}} \Sigma_{(12)}^{-\frac{1}{2}} \left(\frac{1}{n_\ell^{(g)}} S_{(12)}^{(g,\ell)} - \Sigma_{(12)} \right) \Sigma_{(12)}^{-\frac{1}{2}'},$$

$$T = \sum_{g=1}^2 \sum_{\ell=1}^2 \sqrt{m_\ell^{(g)}} T_{(12)}^{(g,\ell)}, \ T_{11} = B_1 T B_1',$$

$$L = \sum_{g=1}^2 \sum_{\ell=1}^2 \left(4\sqrt{m_\ell^{(g)}} - \frac{1}{2} \frac{1}{\sqrt{m_\ell^{(g)}}} \right) T_{(12)}^{(g,\ell)}, \ L_{11} = B_1 L B_1',$$

$$U^{(g)} = \frac{m_1^{(g)} m_2^{(g)}}{m^{(g)}} \left(\frac{1}{\sqrt{m_1^{(g)}}} \mathbf{z}_{(12)}^{(g,1)} - \frac{1}{\sqrt{m_2^{(g)}}} \mathbf{z}_{(12)}^{(g,2)} \right) \\ \left(\frac{1}{\sqrt{m_1^{(g)}}} \mathbf{z}_{(12)}^{(g,1)} - \frac{1}{\sqrt{m_2^{(g)}}} \mathbf{z}_{(12)}^{(g,2)} \right)',$$

$$U = \sum_{g=1}^2 U^{(g)}, \ U_{11} = B_1 U B_1',$$

$$\mathbf{u}_{(12)} = \sqrt{\frac{m^{(2)}}{m^{(1)}}} \sum_{\ell=1}^2 \sqrt{m_\ell^{(1)}} \mathbf{z}_{(12)}^{(1,\ell)} - \sqrt{\frac{m^{(1)}}{m^{(2)}}} \sum_{\ell=1}^2 \sqrt{m_\ell^{(2)}} \mathbf{z}_{(12)}^{(2,\ell)},$$

$$\mathbf{b} = \sqrt{m^{(1)} m^{(2)}} \Sigma^{-\frac{1}{2}} \left(\boldsymbol{\mu}_{(12)}^{(1)} - \boldsymbol{\mu}_{(12)}^{(2)} \right), \ \mathbf{b}_1 = B_1 \mathbf{b},$$

$$M = \frac{1}{m_1} \sum_{g=1}^2 \sqrt{m_1^{(g)}} T^{(g,1)}, \ M_{(12)} = C_{(12)} M C'_{(12)},$$

$$L^* = \frac{1}{2m_1} \sum_{g=1}^2 \frac{1}{\sqrt{m_1^{(g)}}} T^{(g,1)}, \ L_{(12)}^* = C_{(12)} L^* C'_{(12)},$$

$$\mathbf{v} = \sqrt{\frac{m_1^{(2)}}{m_1}} \mathbf{z}^{(1,1)} - \sqrt{\frac{m_1^{(1)}}{m_1}} \mathbf{z}^{(2,1)}, \ \mathbf{v}_{(12)} = C_{(12)} \mathbf{v},$$

$$\mathbf{d} = \sqrt{\frac{m_1^{(1)} m_1^{(2)}}{m_1}} \Sigma^{-\frac{1}{2}} \left(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)} \right), \quad \mathbf{d}_{(12)} = C_{(12)} \mathbf{d}.$$

By using the expressions addressed in the above, we have a stochastic expansion of Q_1 under H_0 as

$$\begin{aligned} Q_1 = & \frac{1}{N} (1 + a_0)^{-1} a_2 + \frac{1}{N\sqrt{N}} \left\{ (1 + a_0)^{-1} a_3 - (1 + a_0)^{-2} a_1 a_2 \right\} \\ & + \frac{1}{N^2} \left\{ (1 + a_0)^{-1} a_4 - (1 + a_0)^{-2} a_1 a_3 - \frac{1}{2} (1 + a_0)^{-2} a_2^2 \right. \\ & \quad \left. + (1 + a_0)^{-3} a_1^2 a_2 \right\} \\ & - \left[\frac{1}{N} (1 + b_0)^{-1} b_2 + \frac{1}{N\sqrt{N}} \left\{ (1 + b_0)^{-1} b_3 - (1 + b_0)^{-2} b_1 b_2 \right\} \right. \\ & \quad \left. + \frac{1}{N^2} \left\{ (1 + b_0)^{-1} b_4 - (1 + b_0)^{-2} b_1 b_3 - \frac{1}{2} (1 + b_0)^{-2} b_2^2 \right. \right. \\ & \quad \left. \left. + (1 + b_0)^{-3} b_1^2 b_2 \right\} \right] + o_p(N^{-2}), \end{aligned}$$

where

$$\begin{aligned} a_0 &= \mathbf{b}' \mathbf{b}, \\ a_1 &= 2\mathbf{b}' \mathbf{u}_{(12)} - \mathbf{b}' T \mathbf{b}, \\ a_2 &= \mathbf{u}_{(12)}' \mathbf{u}_{(12)} - 2\mathbf{b}' T \mathbf{u}_{(12)} + \mathbf{b}' (T^2 - U) \mathbf{b}, \\ a_3 &= -\mathbf{u}_{(12)}' T \mathbf{u}_{(12)} + 2\mathbf{b}' (T^2 - U) \mathbf{u}_{(12)} + \mathbf{b}' (-L + TU + UT - T^3) \mathbf{b}, \\ a_4 &= \mathbf{u}_{(12)}' (T^2 - U) \mathbf{u}_{(12)} + 2\mathbf{b}' (-L + TU + UT - T^3) \mathbf{u}_{(12)} \\ & \quad + \mathbf{b}' (-4U + T^2 + TL + LT - TUT - UT^2 - T^2 U + T^4 + U^2) \mathbf{b}, \\ b_0 &= \mathbf{b}_1' \mathbf{b}_1, \\ b_1 &= 2\mathbf{b}_1' \mathbf{u}_1 - \mathbf{b}_1' T_{11} \mathbf{b}_1, \\ b_2 &= \mathbf{u}_1' \mathbf{u}_1 - 2\mathbf{b}_1' T_{11} \mathbf{u}_1 + \mathbf{b}_1' (T_{11}^2 - U_{11}) \mathbf{b}_1, \\ b_3 &= -\mathbf{u}_1' T_{11} \mathbf{u}_1 + 2\mathbf{b}_1' (T_{11}^2 - U_{11}) \mathbf{u}_1 + \mathbf{b}_1' (-L_{11} + T_{11} U_{11} + U_{11} T_{11} - T_{11}^3) \mathbf{b}_1, \\ b_4 &= \mathbf{u}_1' (T_{11}^2 - U_{11}) \mathbf{u}_1 + 2\mathbf{b}_1' (-L_{11} + T_{11} U_{11} + U_{11} T_{11} - T_{11}^3) \mathbf{u}_1 \\ & \quad + \mathbf{b}_1' (-4U_{11} + T_{11}^2 + T_{11} L_{11} + L_{11} T_{11} - T_{11} U_{11} T_{11} - U_{11} T_{11}^2 \\ & \quad - T_{11}^2 U_{11} + T_{11}^4 + U_{11}^2) \mathbf{b}_1. \end{aligned}$$

We also have a following stochastic expansion of Q_2 :

$$Q_2 = \frac{1}{N} \{ (1 + c_0)^{-1} c_2 \} + \frac{1}{N\sqrt{N}} \{ (1 + c_0)^{-1} c_3 - (1 + c_0)^{-2} c_1 c_2 \}$$

$$\begin{aligned}
& + \frac{1}{N^2} \left\{ (1+c_0)^{-1}c_4 - (1+c_0)^{-2}c_1c_3 - \frac{1}{2}(1+c_0)^{-2}c_2^2 \right. \\
& + (1+c_0)^{-3}c_1^2c_2 \left. \right\} \\
& - \left[\frac{1}{N} \{ (1+d_0)^{-1}d_2 \} + \frac{1}{N\sqrt{N}} \{ (1+d_0)^{-1}d_3 - (1+d_0)^{-2}d_1d_2 \} \right. \\
& + \frac{1}{N^2} \left\{ (1+d_0)^{-1}d_4 - (1+d_0)^{-2}d_1d_3 - \frac{1}{2}(1+d_0)^{-2}d_2^2 \right. \\
& + (1+d_0)^{-3}d_1^2d_2 \left. \right\} \left. \right] + o_p(N^{-2}),
\end{aligned}$$

where

$$\begin{aligned}
c_0 &= \frac{1}{m_1} \mathbf{d}' \mathbf{d}, \\
c_1 &= \frac{2}{m_1} \mathbf{d}' \mathbf{v} - \frac{1}{m_1} \mathbf{d}' M \mathbf{d}, \\
c_2 &= \frac{1}{m_1} \mathbf{v}' \mathbf{v} + \frac{2}{m_1^2} \mathbf{d}' \mathbf{d} - \frac{2}{m_1} \mathbf{d}' M \mathbf{v} + \frac{1}{m_1} \mathbf{d}' M^2 \mathbf{d}, \\
c_3 &= -\frac{1}{m_1} \mathbf{v}' M \mathbf{v} + \frac{4}{m_1^2} \mathbf{d}' \mathbf{v} + \frac{1}{m_1} \mathbf{d}' L^* \mathbf{d} + \frac{2}{m_1} \mathbf{v}' M^2 \mathbf{d} - \frac{4}{m_1^2} \mathbf{d}' M \mathbf{d} \\
& \quad - \frac{1}{m_1} \mathbf{d}' M^3 \mathbf{d}, \\
c_4 &= \frac{2}{m_1^2} \mathbf{v}' \mathbf{v} + \frac{2}{m_1} \mathbf{v}' L^* \mathbf{d} + \frac{1}{m_1} \mathbf{v}' M^2 \mathbf{v} + \frac{4}{m_1^3} \mathbf{d}' \mathbf{d} - \frac{8}{m_1^2} \mathbf{v}' M \mathbf{d} \\
& \quad - \frac{1}{m_1} \mathbf{d}' (M L^* + L^* M) \mathbf{d} - \frac{2}{m_1} \mathbf{v}' M^3 \mathbf{d} + \frac{6}{m_1^2} \mathbf{d}' M^2 \mathbf{d} + \frac{1}{m_1} \mathbf{d}' M^4 \mathbf{d}, \\
d_0 &= \frac{1}{m_1} \mathbf{d}'_{(12)} \mathbf{d}_{(12)}, \\
d_1 &= \frac{2}{m_1} \mathbf{d}'_{(12)} \mathbf{v}_{(12)} - \frac{1}{m_1} \mathbf{d}'_{(12)} M_{(12)} \mathbf{d}_{(12)}, \\
d_2 &= \frac{1}{m_1} \mathbf{v}'_{(12)} \mathbf{v}_{(12)} - \frac{2}{m_1} \mathbf{d}'_{(12)} M_{(12)} \mathbf{v}_{(12)} + \frac{2}{m_1^2} \mathbf{d}'_{(12)} \mathbf{d}_{(12)} \\
& \quad + \frac{1}{m_1} \mathbf{d}'_{(12)} M_{(12)}^2 \mathbf{d}_{(12)}, \\
d_3 &= -\frac{1}{m_1} \mathbf{v}'_{(12)} M_{(12)} \mathbf{v}_{(12)} + \frac{4}{m_1^2} \mathbf{d}'_{(12)} \mathbf{v}_{(12)} + \frac{1}{m_1} \mathbf{d}'_{(12)} L_{(12)}^* \mathbf{d}_{(12)} \\
& \quad + \frac{2}{m_1} \mathbf{v}'_{(12)} M_{(12)}^2 \mathbf{d}_{(12)} - \frac{4}{m_1^2} \mathbf{d}'_{(12)} M_{(12)} \mathbf{d}_{(12)} - \frac{1}{m_1} \mathbf{d}'_{(12)} M_{(12)}^3 \mathbf{d}_{(12)}, \\
d_4 &= \frac{2}{m_1^2} \mathbf{v}'_{(12)} \mathbf{v}_{(12)} + \frac{2}{m_1} \mathbf{v}'_{(12)} L_{(12)}^* \mathbf{d}_{(12)} + \frac{1}{m_1} \mathbf{v}_{(12)} M_{(12)}^2 \mathbf{v}_{(12)} + \frac{4}{m_1^3} \mathbf{d}'_{(12)} \mathbf{d}_{(12)} \\
& \quad - \frac{8}{m_1^2} \mathbf{v}'_{(12)} M_{(12)} \mathbf{d}_{(12)} - \frac{1}{m_1} \mathbf{d}'_{(12)} \left(M_{(12)} L_{(12)}^* + L_{(12)}^* M_{(12)} \right) \mathbf{d}_{(12)}
\end{aligned}$$

$$-\frac{2}{m_1} \mathbf{v}_{(12)}' M_{(12)}^3 \mathbf{d}_{(12)} + \frac{6}{m_1^2} \mathbf{d}_{(12)}' M_{(12)}^2 \mathbf{d}_{(12)} + \frac{1}{m_1} \mathbf{d}_{(12)}' M_{(12)}^4 \mathbf{d}_{(12)}.$$

By evaluating the expectations of each term, we can complete the following main theorem in this paper.

Theorem 3.1. Suppose that two-step monotone missing data $\mathbf{x}_1^{(g)}, \dots, \mathbf{x}_{N_1^{(g)}}^{(g)}, \mathbf{x}_{(12)N_1^{(g)}+1}^{(g)}, \dots, \mathbf{x}_{(12)N^{(g)}}^{(g)}$ is observed from $\Pi^{(g)} : N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$ ($g \in \{1, 2\}$). Then, under H_0 , the expectation of LRT statistic is given by

$$E[-2 \log \lambda] = p_2 \left(1 + \frac{1}{N} \sum_{i=1}^3 A_i \right) + p_3 \left(1 + \frac{1}{N} \sum_{j=4}^6 A_j \right) + o(N^{-1}),$$

where

$$\begin{aligned} A_1 &= (1 + a_0)^{-1} \{2p_1 + p_2 - 1 + (2p_1 + p_2 + 3)a_0\}, \\ A_2 &= -(1 + a_0)^{-2} \left\{ \frac{1}{2}(2p_1 + p_2 + 2) + \frac{3}{2}(2p_1 + p_2 + 2)a_0 + 5a_0 + 2a_0^2 \right\}, \\ A_3 &= (1 + a_0)^{-3}(4a_0 + 2a_0^3), \\ A_4 &= \frac{1}{m_1}(1 + c_0)^{-1} \left\{ p_3 + 2(p_1 + p_2) + 3 + (p_3 + 2(p_1 + p_2) + 6)c_0 \right\}, \\ A_5 &= \frac{1}{m_1}(1 + c_0)^{-2} \left\{ -\frac{1}{2}(p_3 + 2p_1 + 2p_2 + 2) - (p_3 + 2p_1 + 2p_2 + 5)c_0 \right. \\ &\quad \left. - \frac{1}{2}(p_3 + 2p_1 + 2p_2 + 12)c_0^2 \right\}, \\ A_6 &= \frac{1}{m_1}(1 + c_0)^{-3}(4c_0 + 4c_0^2 + 2c_0^3). \end{aligned}$$

Proof. $\mathbf{u}_{(12)}$ and U depend only on standardized sample mean vector $\mathbf{z}_{(12)}^{(g, \ell)}$ and T and L depend only on sample covariance matrix $(n_\ell^{(g)})^{-1} S^{(g, \ell)}$. Therefore we note that $(\mathbf{u}_{(12)}, U)$ and (T, L) are mutually independent. Because $\mathbf{b} = (\mathbf{b}_1' \mathbf{0}_{p_2}')'$ under H_0 , it holds that $a_0 = b_0 = m^{(1)}m^{(2)}\Delta_{11}^2$ and $a_1 = b_1 = 2\mathbf{b}_1\mathbf{u}_1 - m^{(1)}m^{(2)}\mathbf{b}_1'T_{11}\mathbf{b}_1$. Then, by Lemma A.5,

$$\begin{aligned} E[a_2 - b_2] &= p_1 + p_2 + (p_1 + p_2 + 1)\mathbf{b}'\mathbf{b} - 2\mathbf{b}'\mathbf{b} \\ &\quad - \{p_1 + (p_1 + 1)\mathbf{b}_1'\mathbf{b}_1 - 2\mathbf{b}_1'\mathbf{b}_1\} \\ &= (1 + a_0)p_2, \\ E[a_3 - b_3] &= -\frac{1}{\sqrt{N}} \{(p_1 + p_2)^2 + 3(p_1 + p_2) + 4\}\mathbf{b}'\mathbf{b} \\ &\quad + \frac{1}{\sqrt{N}} \{p_1^2 + 3p_1 + 4\}\mathbf{b}_1'\mathbf{b}_1 + o(N^{-\frac{1}{2}}) \end{aligned} \tag{3.1}$$

$$= -\frac{1}{\sqrt{N}}(p_2^2 + 2p_1p_2 + 3p_2)a_0 + o(N^{-\frac{1}{2}}), \quad (3.2)$$

$$\begin{aligned} E[a_1a_2 - b_1b_2] &= -\frac{1}{\sqrt{N}} \{2(p_1 + p_2) + 6\} (\mathbf{b}'\mathbf{b})^2 \\ &\quad + \frac{1}{\sqrt{N}}(2p_1 + 6)(\mathbf{b}'_1\mathbf{b}_1)^2 + o(N^{-\frac{1}{2}}) \\ &= -\frac{1}{\sqrt{N}}2p_2a_0^2 + o(N^{-\frac{1}{2}}), \end{aligned} \quad (3.3)$$

$$\begin{aligned} E[a_4 - b_4] &= (p_1 + p_2)^2 - (p_1 + p_2) + \{2(p_1 + p_2)^2 + 6(p_1 + p_2) + 2\}\mathbf{b}'\mathbf{b} \\ &\quad - \{p_1^2 - p_1 + (2p_1^2 + 6p_1 + 2)\mathbf{b}'_1\mathbf{b}_1\} \\ &= p_2^2 + 2p_1p_2 - p_2 + 2(p_2^2 + 2p_1p_2 + 3p_2)a_0, \end{aligned} \quad (3.4)$$

$$\begin{aligned} E[a_1a_3 - b_1b_3] &= \{5(p_1 + p_2) + 1\}\mathbf{b}'\mathbf{b} + 4(p_1 + p_2 + 1)(\mathbf{b}'\mathbf{b})^2 \\ &\quad - \{(5p_1 + 1)\mathbf{b}'_1\mathbf{b}_1 + 4(p_1 + 1)(\mathbf{b}'_1\mathbf{b}_1)^2\} \\ &= 5p_2a_0 + 4p_2a_0^2, \end{aligned} \quad (3.5)$$

$$\begin{aligned} E[a_2^2 - b_2^2] &= (p_1 + p_2)^2 + 2(p_1 + p_2) + 3(p_1 + p_2)^2\mathbf{b}'\mathbf{b} + 2(p_1 + p_2)\mathbf{b}'\mathbf{b} \\ &\quad + 4\mathbf{b}'\mathbf{b} + 11(\mathbf{b}'\mathbf{b})^2 \\ &\quad - \{p_1^2 + 2p_1 + 3p_1^2\mathbf{b}'_1\mathbf{b}_1 + 2p_1\mathbf{b}'_1\mathbf{b}_1 + 4\mathbf{b}'_1\mathbf{b}_1 + 11(\mathbf{b}'_1\mathbf{b}_1)^2\} \\ &= p_2^2 + 2p_1p_2 + 2p_2 + 3(p_2^2 + 2p_1p_2)a_0 + 2p_2a_0, \end{aligned} \quad (3.6)$$

$$\begin{aligned} E[a_1^2a_2 - b_1^2b_2] &= 4(p_1 + p_2 + 2)\mathbf{b}'\mathbf{b} + 8(\mathbf{b}'\mathbf{b})^2 + \{2(p_1 + p_2) + 5\}(\mathbf{b}'\mathbf{b})^3 \\ &\quad - \{4(p_1 + 2)\mathbf{b}'_1\mathbf{b}_1 + 8(\mathbf{b}'_1\mathbf{b}_1)^2 + (2p_1 + 5)(\mathbf{b}'_1\mathbf{b}_1)^3\} \\ &= 4p_2\mathbf{b}'_1\mathbf{b}_1 + 2p_2(\mathbf{b}'_1\mathbf{b}_1)^3. \end{aligned} \quad (3.7)$$

Combining the expectations from (3.1) to (3.7), we obtain an asymptotic expansion of $NE[Q_1]$:

$$NE[Q_1] = p_2 \left(1 + \frac{1}{N} \sum_{i=1}^3 A_i \right) + o(N^{-1}).$$

For N_1Q_2 , $\mathbf{v}$ depends on standardized sample mean vector $\mathbf{z}^{(g,1)}$, and M and L^* depend on sample covariance matrix $(n_1^{(g)})^{-1}S^{(g,1)}$. Therefore, we note that $\mathbf{v}$ and (M, L^*) are mutually independent. Because $\mathbf{d} = (\mathbf{d}'_1 \ \mathbf{0}'_{p_2} \ \mathbf{0}'_{p_3})'$ and $\mathbf{d}_{(12)} = (\mathbf{d}'_1 \ \mathbf{0}'_{p_2})'$ under H_0 , it holds that $c_0 = d_0$ and $c_1 = d_1$. Then, by Corollary A.1, we have

$$\begin{aligned} E[c_2 - d_2] &= \frac{1}{m_1}(1 + c_0)p_3, \\ E[c_3 - d_3] &= -\frac{1}{\sqrt{N}}\frac{1}{m_1^3}(p^2 + 3p + 4)\mathbf{d}'\mathbf{d} \\ &\quad + \frac{1}{\sqrt{N}}\frac{1}{m_1^3}\{(p_1 + p_2)^2 + 3(p_1 + p_2) + 4\}\mathbf{d}'_{(12)}\mathbf{d}_{(12)} + o(N^{-1}) \end{aligned} \quad (3.8)$$

$$= -\frac{1}{\sqrt{N}} \frac{1}{m_1^2} \{p_3^2 + 2p_3(p_1 + p_2) + 3p_3\} c_0 + o(N^{-1}), \quad (3.9)$$

$$\begin{aligned} E[c_1 c_2 - d_1 d_2] &= -\frac{1}{\sqrt{N}} \frac{1}{m_1^4} (2p + 6) (\mathbf{d}' \mathbf{d})^2 \\ &\quad + \frac{1}{\sqrt{N}} \frac{1}{m_1^4} \{2(p_1 + p_2) + 6\} (\mathbf{d}'_{(12)} \mathbf{d}_{(12)})^2 + o(N^{-1}) \\ &= -\frac{1}{\sqrt{N}} \frac{1}{m_1^2} 2p_3 c_0^2 + o(N^{-1}), \end{aligned} \quad (3.10)$$

$$\begin{aligned} E[c_4 - d_4] &= \frac{1}{m_1^2} \{p_3^2 + 2p_3(p_1 + p_2) + 3p_3\} \\ &\quad + \frac{1}{m_1^2} \{2p_3^2 + 4p_3(p_1 + p_2) + 9p_3\} c_0, \end{aligned} \quad (3.11)$$

$$E[c_1 c_3 - d_1 d_3] = \frac{4}{m_1^2} p_3 c_0^2, \quad (3.12)$$

$$\begin{aligned} E[c_2^2 - d_2^2] &= \frac{1}{m_1^2} \{p_3^2 + 2p_3(p_1 + p_2) + 2p_3\} \\ &\quad + \frac{2}{m_1^2} \{p_3^2 + 2p_3(p_1 + p_2) + 5p_3\} c_0 \\ &\quad + \frac{1}{m_1^2} \{p_3^2 + 2p_3(p_1 + p_2) + 8p_3\} c_0^2, \end{aligned} \quad (3.13)$$

$$E[c_1^2 c_2 - d_1^2 d_2] = \frac{4}{m_1^2} p_3 c_0 + \frac{4}{m_1^2} p_3 c_0^2 + \frac{2}{m_1^2} p_3 c_0^3. \quad (3.14)$$

Combining the expectations from (3.8) to (3.14), we obtain an asymptotic expansion of $N_1 Q_2$:

$$N_1 E[Q_2] = N m_1 E[Q_2] = p_3 \left(1 + \frac{1}{N} \sum_{j=4}^6 A_j \right) + o(N^{-1}).$$

□

In conclusion, by Theorem 3.1, we propose the corrected test statistic:

$$-2 \log \lambda^* = \rho_1 N Q_1 + \rho_2 N_1 Q_2,$$

where, ρ_1 and ρ_2 are correction terms as follows:

$$\rho_1 = \left(1 + \frac{1}{N} \sum_{i=1}^3 A_i \right)^{-1}, \quad \rho_2 = \left(1 + \frac{1}{N} \sum_{j=4}^6 A_j \right)^{-1}.$$

However, $-2 \log \lambda^*$ depends on unknown quantities a_0 and c_0 . In practical use, we use $-2 \log \hat{\lambda}^*$ constructed by substituted consistent estimators $\hat{a}_0$ and $\hat{c}_0$ into a_0 and c_0 , respectively. Therefore, we finally proposed the testing procedure : if $-2 \log \hat{\lambda}^* > \chi_{p_2+p_3}^2(\alpha)$, then H_0 is rejected, otherwise H_0 is not rejected.

4 Simulation studies

In this chapter, we investigate the sizes of the LRT statistic and the corrected LRT in Chapter 3 by Monte Carlo simulation for the significance level $\alpha = 0.05$. Also, in practical situations, we should estimate $a_0 = m^{(1)}m^{(2)}\Delta_{11}^2$ and $c_0 = m_1^{(1)}m_1^{(2)}m_1^{-2}\Delta_{11}^2$. Then, the common parameters between a_0 and c_0 is Δ_{11}^2 . So, the plug-in corrected LRT statistic using the asymptotic unbiased estimator of Δ_{11}^2 derived by [12] is noted as $-2\log\hat{\lambda}^*$. We generate two-step monotone missing data for two groups under several settings of (p_1, p_2, p_3) and $(N_1^{(1)}, N_2^{(1)}, N_1^{(2)}, N_2^{(2)})$. We compute the sizes of the LRT and the corrected LRT with 100,000 runs under these situations.

In the following tables, we present the simulation results for each case:

Case 1 : $\Pi^{(1)} : N_p(1.0\boldsymbol{\mu}, I_p), \Pi^{(2)} : N_p(\mathbf{0}_p, I_p), (p_1, p_2, p_3) = (1, 1, 1),$

Case 2 : $\Pi^{(1)} : N_p(1.0\boldsymbol{\mu}, I_p), \Pi^{(2)} : N_p(\mathbf{0}_p, I_p), (p_1, p_2, p_3) = (2, 2, 2),$

Case 3 : $\Pi^{(1)} : N_p(1.0\boldsymbol{\mu}, I_p), \Pi^{(2)} : N_p(\mathbf{0}_p, I_p), (p_1, p_2, p_3) = (3, 3, 3),$

Case 4 : $\Pi^{(1)} : N_p(1.5\boldsymbol{\mu}, I_p), \Pi^{(2)} : N_p(\mathbf{0}_p, I_p), (p_1, p_2, p_3) = (3, 3, 3),$

where $\boldsymbol{\mu} = (\mathbf{1}_{p_1}' \mathbf{0}_{(p_2+p_3)}')'$.

Table 1 The simulated sizes of the test based on $-2 \log \lambda$ and $-2 \log \hat{\lambda}^*$ for Case 1

$N_1^{(1)}$	$N_2^{(1)}$	$N_1^{(2)}$	$N_2^{(2)}$	$-2 \log \lambda$	$-2 \log \hat{\lambda}^*$
20	20	20	20	0.063	0.053
30	30	30	30	0.058	0.051
40	40	40	40	0.057	0.052
50	50	50	50	0.055	0.051
60	60	60	60	0.053	0.049
70	70	70	70	0.055	0.052
20	20	10	10	0.069	0.055
20	20	30	30	0.061	0.053
20	20	40	40	0.058	0.051
20	20	50	50	0.056	0.050
20	20	60	60	0.057	0.052
20	20	70	70	0.056	0.052
10	10	20	20	0.069	0.054
30	30	20	20	0.059	0.051
40	40	20	20	0.058	0.052
50	50	20	20	0.058	0.052
60	60	20	20	0.057	0.051
70	70	20	20	0.056	0.052
10	20	10	20	0.079	0.056
30	20	30	20	0.059	0.052
40	20	40	20	0.056	0.050
50	20	50	20	0.055	0.051
60	20	60	20	0.053	0.050
70	20	70	20	0.054	0.051

Table 2 The simulated sizes of the test based on $-2 \log \lambda$ and $-2 \log \hat{\lambda}^*$ for Case 2

$N_1^{(1)}$	$N_2^{(1)}$	$N_1^{(2)}$	$N_2^{(2)}$	$-2 \log \lambda$	$-2 \log \hat{\lambda}^*$
20	20	20	20	0.082	0.054
30	30	30	30	0.068	0.050
40	40	40	40	0.064	0.051
50	50	50	50	0.059	0.050
60	60	60	60	0.058	0.050
70	70	70	70	0.058	0.051
20	20	10	10	0.095	0.057
20	20	30	30	0.073	0.052
20	20	40	40	0.068	0.051
20	20	50	50	0.066	0.052
20	20	60	60	0.064	0.051
20	20	70	70	0.060	0.049
10	10	20	20	0.095	0.063
30	30	20	20	0.073	0.054
40	40	20	20	0.069	0.054
50	50	20	20	0.066	0.054
60	60	20	20	0.062	0.053
70	70	20	20	0.063	0.056
10	20	10	20	0.126	0.078
30	20	30	20	0.070	0.051
40	20	40	20	0.065	0.050
50	20	50	20	0.060	0.048
60	20	60	20	0.059	0.048
70	20	70	20	0.057	0.048

Table 3 The simulated sizes of the test based on $-2 \log \lambda$ and $-2 \log \hat{\lambda}^*$ for Case 3

$N_1^{(1)}$	$N_2^{(1)}$	$N_1^{(2)}$	$N_2^{(2)}$	$-2 \log \lambda$	$-2 \log \hat{\lambda}^*$
20	20	20	20	0.106	0.054
30	30	30	30	0.083	0.051
40	40	40	40	0.073	0.051
50	50	50	50	0.069	0.050
60	60	60	60	0.064	0.049
70	70	70	70	0.062	0.049
20	20	10	10	0.137	0.062
20	20	30	30	0.091	0.052
20	20	40	40	0.083	0.052
20	20	50	50	0.076	0.050
20	20	60	60	0.072	0.051
20	20	70	70	0.070	0.052
10	10	20	20	0.137	0.063
30	30	20	20	0.091	0.053
40	40	20	20	0.083	0.052
50	50	20	20	0.077	0.053
60	60	20	20	0.072	0.051
70	70	20	20	0.069	0.050
10	20	10	20	0.217	0.080
30	20	30	20	0.083	0.050
40	20	40	20	0.074	0.050
50	20	50	20	0.069	0.050
60	20	60	20	0.065	0.050
70	20	70	20	0.063	0.049

Table 4 The simulated sizes of the test based on $-2\log \lambda$ and $-2\log \hat{\lambda}^*$ for Case 4

$N_1^{(1)}$	$N_2^{(1)}$	$N_1^{(2)}$	$N_2^{(2)}$	$-2\log \lambda$	$-2\log \hat{\lambda}^*$
20	20	20	20	0.108	0.055
30	30	30	30	0.081	0.051
40	40	40	40	0.073	0.051
50	50	50	50	0.067	0.049
60	60	60	60	0.062	0.048
70	70	70	70	0.062	0.050
20	20	10	10	0.139	0.064
20	20	30	30	0.091	0.053
20	20	40	40	0.083	0.052
20	20	50	50	0.077	0.051
20	20	60	60	0.073	0.051
20	20	70	70	0.069	0.051
10	10	20	20	0.135	0.061
30	30	20	20	0.092	0.052
40	40	20	20	0.082	0.051
50	50	20	20	0.078	0.052
60	60	20	20	0.071	0.050
70	70	20	20	0.069	0.050
10	20	10	20	0.217	0.079
30	20	30	20	0.084	0.051
40	20	40	20	0.076	0.051
50	20	50	20	0.071	0.051
60	20	60	20	0.066	0.049
70	20	70	20	0.064	0.049

5 Conclusion

In this paper, we constructed the LRT statistic of redundancy test for some variables in linear discriminant analysis based on two-step missing data in Chapter 2. In Chapter 3, we derived Bartlett correction terms for controlling the sizes of test more exactly when $N^{(g)}$ is not so large. According to the results of Chapter 4, we observed that

the corrected test statistics outperformed the original LRT statistic for all conducted cases. Furthermore, we also give the formula of $E[WAWBWCW]$ for completing the main theorem in this paper, where $W \sim W_p(n, \Sigma)$, A , B , and C are $p \times p$ constant matrices. Finally, we address some of the future problems. By extending the main result, we may derive the asymptotic power function of the LRT for the problem considered in this paper. Moreover, by using a more flexible variable selection, we may reduce the bias of AIC under settings similar to this paper.

A Lemmas and corollary

In this chapter, we show the lemmas for completing the proofs of Theorem 2.1, 2.2 and 3.1.

Lemma A.1. Suppose that $\mathbf{a}$ and $\mathbf{b}$ are n -dimensional vectors, $A = (a_{ij})$ and $B = (b_{ij})$ are $n \times n$ square matrices, and $A = A'$. It holds that:

- (i) $\frac{\partial}{\partial A} \text{tr}(AB) = B + B' - \text{diag}(B)$
- (ii) $\frac{\partial}{\partial A} \log |A| = 2A^{-1} - \text{diag}(A^{-1})$
- (iii) $\frac{\partial}{\partial \mathbf{b}} (\mathbf{a} - \mathbf{b})' A (\mathbf{a} - \mathbf{b}) = -2A (\mathbf{a} - \mathbf{b})$

Proof. First, we prove (i). Since

$$AB = \left(\sum_{k=1}^n a_{ik} b_{kj} \right),$$

we have

$$\text{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ji} b_{ij}.$$

Thus,

$$\frac{\partial}{\partial A} \text{tr}(AB) = \begin{pmatrix} b_{11} & \cdots & b_{n1} + b_{1n} \\ \vdots & \ddots & \vdots \\ b_{1n} + b_{n1} & \cdots & b_{nn} \end{pmatrix} = B + B' - \text{diag}(B).$$

Next, we prove (ii). Using cofactor expansion, we have:

$$|A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{i=1}^n a_{ij} \Delta_{ij},$$

where $\Delta_{ij} = (-1)^{i+j}|A_{(i,j)}|$, and $A_{(i,j)}$ is the matrix obtained by removing the i -th row and j -th column from A . Using $a_{ij} = a_{ji}$, we get:

$$\frac{\partial \log |A|}{\partial a_{ij}} = \begin{cases} \frac{1}{|A|}(\Delta_{ij} + \Delta_{ji}) & (i \neq j), \\ \frac{1}{|A|}\Delta_{ij} & (i = j). \end{cases}$$

Since $\Delta_{ij} = \Delta_{ji}$ when A is symmetric, the cofactor matrix $\tilde{A}$, whose elements are $\tilde{a}_{ij} = (-1)^{i+j}|A_{(i,j)}| = \Delta_{ij}$, is also symmetric. Thus, we obtain:

$$\frac{\partial \log |A|}{\partial A} = \begin{pmatrix} \frac{\Delta_{11}}{|A|} & \cdots & \frac{\Delta_{1n} + \Delta_{n1}}{|A|} \\ \vdots & \ddots & \vdots \\ \frac{\Delta_{n1} + \Delta_{1n}}{|A|} & \cdots & \frac{\Delta_{nn}}{|A|} \end{pmatrix}.$$

This simplifies to:

$$\frac{\partial \log |A|}{\partial A} = \frac{1}{|A|}(2\tilde{A} - \text{diag}(\tilde{A})) = \frac{1}{|A|}(2|A|A^{-1} - |A|\text{diag}(A^{-1})) = 2A^{-1} - \text{diag}(A^{-1}).$$

Finally, we suppose that $\mathbf{a} = (a_1, \dots, a_n)'$ and $\mathbf{b} = (b_1, \dots, b_n)'$ to prove (iii). It holds that

$$\begin{aligned} (\mathbf{a} - \mathbf{b})'A(\mathbf{a} - \mathbf{b}) &= \mathbf{a}'A\mathbf{a} - \mathbf{a}'A\mathbf{b} - \mathbf{b}'A\mathbf{a} + \mathbf{b}'A\mathbf{b} \\ &= \sum_i \sum_j a_i a_{ij} a_j - \sum_i \sum_j a_i a_{ij} b_j - \sum_i \sum_j b_i a_{ij} a_j + \sum_i \sum_j b_i a_{ij} b_j. \end{aligned}$$

Therefor, we obtain

$$\begin{aligned} \frac{\partial}{\partial \mathbf{b}}(\mathbf{a} - \mathbf{b})'A(\mathbf{a} - \mathbf{b}) &= -(\mathbf{a}'A)' - A\mathbf{a} + 2A\mathbf{b} \\ &= -2A(\mathbf{a} - \mathbf{b}). \end{aligned}$$

Thus, the proof is complete. □

Lemma A.2. Let A be an $r \times r$ square matrix, Ξ an $r \times q$ matrix, and B a $q \times q$ square matrix, where $r \leq q$.

Then,

$$\frac{\partial}{\partial \Xi} \text{tr}(A\Xi B\Xi') = A\Xi B + A'\Xi B'$$

holds.

Proof. Let

$$c = a_{i\ell}\xi_{\ell k}b_{kj}\xi'_{ji} = a_{i\ell}b_{kj}\xi_{\ell k}\xi_{ij}.$$

Then,

$$\frac{\partial c}{\partial \xi_{mn}} = \begin{cases} a_{m\ell}b_{kn}\xi_{\ell k} & (i = m, j = n, \ell \neq m, k \neq n), \\ a_{im}b_{nj}\xi_{ij} & (i \neq m, j \neq n, \ell = m, k = n), \\ 2a_{mm}b_{nn}\xi_{mn} & (i = \ell = m, j = k = n), \\ 0 & (\text{otherwise}). \end{cases}$$

Thus,

$$\begin{aligned} \frac{\partial}{\partial \xi_{mn}} \sum_{i=1}^r \sum_{j=1}^q \sum_{k=1}^q \sum_{\ell=1}^r a_{i\ell}\xi_{\ell k}b_{kj}\xi'_{ji} &= \sum_{\ell \neq m, k \neq n} a_{m\ell}\xi_{\ell k}b_{kn} + \sum_{i \neq m, j \neq n} a'_{mi}\xi_{ij}b'_{jn} \\ &\quad + 2a_{mm}b_{nn}\xi_{mn} \\ &= \sum_{\ell, k} a_{m\ell}\xi_{\ell k}b_{kn} + \sum_{m, n} a'_{mi}\xi_{ij}b'_{jn}. \end{aligned}$$

Therefore, it shows that

$$\frac{\partial}{\partial \Xi} \text{tr} \left(A \Xi B \Xi' \right) = A \Xi B + A' \Xi B'.$$

□

Lemma A.3. Suppose that $W \sim W_p(n, \Sigma)$ and A , B , and C are $p \times p$ constant matrices. Then the expectations of $WAWBW$ and $WAWBWCW$ can be obtained as

$$\begin{aligned} E[W] &= n\Sigma, \\ E[WAW] &= n^2\Sigma A \Sigma + n\Sigma A' \Sigma + n\text{tr}(\Sigma A)\Sigma, \\ E[\text{tr}(AW)W] &= n^2\text{tr}(A\Sigma)\Sigma + n\Sigma A \Sigma + n\Sigma A' \Sigma, \\ E[WAWBW] &= n^3\Sigma A \Sigma B \Sigma \\ &\quad + n^2 \left\{ \Sigma A' \Sigma B \Sigma + \Sigma B \Sigma \text{tr}(A' \Sigma) + \Sigma B' \Sigma A' \Sigma \right. \\ &\quad \left. + \Sigma \text{tr}(A \Sigma B \Sigma) + \Sigma A \Sigma B' \Sigma + \Sigma A \Sigma \text{tr}(B' \Sigma) \right\} \\ &\quad + n \left\{ \Sigma A' \Sigma B' \Sigma + \Sigma A' \Sigma \text{tr}(B' \Sigma) + \Sigma B \Sigma A \Sigma \right. \\ &\quad \left. + \Sigma B \Sigma A' \Sigma + \Sigma B' \Sigma \text{tr}(A' \Sigma) + \Sigma B' \Sigma A \Sigma \right. \\ &\quad \left. + \Sigma \text{tr}(A' \Sigma) \text{tr}(B' \Sigma) + \Sigma \text{tr}(A \Sigma B' \Sigma) \right\}, \\ E[WAWBWCW] &= n^4\Sigma A \Sigma B \Sigma C \Sigma + n^3R_3(\Sigma) + n^2R_2(\Sigma) + nR_1(\Sigma), \end{aligned}$$

respectively, where

$$\begin{aligned}
R_3(\Sigma) = & \Sigma A \Sigma B \Sigma C' \Sigma + \Sigma A \Sigma B \text{tr}(C \Sigma) + \Sigma A \Sigma B' \Sigma C \Sigma + \Sigma A \Sigma C \Sigma \text{tr}(B \Sigma) \\
& + \Sigma A \Sigma C' \Sigma B' \Sigma + \Sigma A \Sigma \text{tr}(B \Sigma C \Sigma) + \Sigma A' \Sigma B \Sigma C \Sigma + \Sigma B \Sigma C \Sigma \text{tr}(A \Sigma) \\
& + \Sigma B' \Sigma A' \Sigma C \Sigma + \Sigma C \Sigma \text{tr}(B \Sigma A \Sigma) + \Sigma C' \Sigma B' \Sigma A' \Sigma \\
& + \Sigma \text{tr}(A \Sigma B \Sigma C \Sigma),
\end{aligned}$$

$$\begin{aligned}
R_2(\Sigma) = & \Sigma A \Sigma B' \Sigma C' \Sigma + \Sigma A \Sigma B' \Sigma \text{tr}(C \Sigma) + \Sigma A \Sigma C \Sigma B \Sigma + \Sigma A \Sigma C \Sigma B' \Sigma \\
& + \Sigma A \Sigma C' \Sigma \text{tr}(B \Sigma) + \Sigma A \Sigma C' \Sigma B \Sigma + \Sigma A \Sigma \text{tr}(B \Sigma) \text{tr}(C \Sigma) \\
& + \Sigma A \Sigma \text{tr}(B \Sigma C' \Sigma) + \Sigma A' \Sigma B \Sigma C' \Sigma + \Sigma A' \Sigma B \Sigma \text{tr}(C \Sigma) \\
& + \Sigma A' \Sigma B' \Sigma C \Sigma + \Sigma A' \Sigma C \Sigma \text{tr}(B \Sigma) + \Sigma A' \Sigma C' \Sigma B' \Sigma + \Sigma A' \Sigma \text{tr}(B \Sigma C \Sigma) \\
& + \Sigma B \Sigma C' \Sigma \text{tr}(A \Sigma) + \Sigma B \Sigma \text{tr}(A \Sigma) \text{tr}(C \Sigma) \\
& + \Sigma B \Sigma A \Sigma C \Sigma + \Sigma B \Sigma A' \Sigma C \Sigma + \Sigma B \Sigma C \Sigma A \Sigma + \Sigma B \Sigma C \Sigma A' \Sigma \\
& + \Sigma B' \Sigma C \Sigma \text{tr}(A \Sigma) + \Sigma B' \Sigma A \Sigma C \Sigma + \Sigma B' \Sigma \text{tr}(A \Sigma C' \Sigma) \\
& + \Sigma B' \Sigma C' \Sigma A \Sigma + \Sigma B' \Sigma A' \Sigma C' \Sigma + \Sigma B' \Sigma A' \Sigma \text{tr}(C \Sigma) \\
& + \Sigma C \Sigma \text{tr}(A \Sigma) \text{tr}(B \Sigma) + \Sigma C \Sigma \text{tr}(A \Sigma B' \Sigma) + \Sigma C \Sigma A' \Sigma B' \Sigma \\
& + \Sigma C \Sigma B \Sigma A \Sigma + \Sigma C \Sigma A \Sigma B \Sigma + \Sigma C \Sigma B' \Sigma A' \Sigma + \Sigma C' \Sigma B' \Sigma \text{tr}(A \Sigma) \\
& + \Sigma C' \Sigma B' \Sigma A \Sigma + \Sigma C' \Sigma \text{tr}(A \Sigma B \Sigma) + \Sigma C' \Sigma A \Sigma B \Sigma + \Sigma C' \Sigma B \Sigma A' \Sigma \\
& + \Sigma C' \Sigma A' \Sigma \text{tr}(B \Sigma) + \Sigma \text{tr}(A \Sigma) \text{tr}(B \Sigma C \Sigma) + \Sigma \text{tr}(A \Sigma C' \Sigma B' \Sigma) \\
& + \Sigma \text{tr}(A \Sigma B \Sigma) \text{tr}(C \Sigma) + \Sigma \text{tr}(A \Sigma B \Sigma C' \Sigma) + \Sigma \text{tr}(A \Sigma B' \Sigma C \Sigma) \\
& + \Sigma \text{tr}(A \Sigma C \Sigma) \text{tr}(B \Sigma),
\end{aligned}$$

and

$$\begin{aligned}
R_1(\Sigma) = & \Sigma A' \Sigma B' \Sigma C' \Sigma + \Sigma A' \Sigma B' \Sigma \text{tr}(C \Sigma) + \Sigma A' \Sigma C \Sigma B \Sigma + \Sigma A' \Sigma C \Sigma B' \Sigma \\
& + \Sigma A' \Sigma C' \Sigma \text{tr}(B \Sigma) + \Sigma A' \Sigma C' \Sigma B \Sigma + \Sigma A' \Sigma \text{tr}(B \Sigma) \text{tr}(C \Sigma) + \Sigma A' \Sigma \text{tr}(B \Sigma C' \Sigma) \\
& + \Sigma B \Sigma A \Sigma C' \Sigma + \Sigma B \Sigma A \Sigma \text{tr}(C \Sigma) + \Sigma B \Sigma \text{tr}(A \Sigma C' \Sigma) + \Sigma B \Sigma C' \Sigma A \Sigma \\
& + \Sigma B \Sigma A' \Sigma C' \Sigma + \Sigma B \Sigma \text{tr}(A \Sigma C \Sigma) + \Sigma B \Sigma A' \Sigma \text{tr}(C \Sigma) + \Sigma B \Sigma C' \Sigma A' \Sigma \\
& + \Sigma B' \Sigma C' \Sigma \text{tr}(A \Sigma) + \Sigma B' \Sigma \text{tr}(A \Sigma) \text{tr}(C \Sigma) + \Sigma B' \Sigma A \Sigma C' \Sigma + \Sigma B' \Sigma A \Sigma \text{tr}(C \Sigma) \\
& + \Sigma B' \Sigma \text{tr}(A \Sigma C \Sigma) + \Sigma B' \Sigma C \Sigma A \Sigma + \Sigma B' \Sigma C' \Sigma A' \Sigma + \Sigma B' \Sigma C \Sigma A' \Sigma \\
& + \Sigma C \Sigma B \Sigma \text{tr}(A \Sigma) + \Sigma C \Sigma B' \Sigma \text{tr}(A \Sigma) + \Sigma C \Sigma A' \Sigma B \Sigma + \Sigma C \Sigma B' \Sigma A \Sigma \\
& + \Sigma C \Sigma A \Sigma B' \Sigma + \Sigma C \Sigma A \Sigma \text{tr}(B \Sigma) + \Sigma C \Sigma B \Sigma A' \Sigma + \Sigma C \Sigma A' \Sigma \text{tr}(B \Sigma) \\
& + \Sigma C' \Sigma \text{tr}(A \Sigma) \text{tr}(B \Sigma) + \Sigma C' \Sigma B \Sigma \text{tr}(A \Sigma) + \Sigma C' \Sigma \text{tr}(A \Sigma B' \Sigma) + \Sigma C' \Sigma A' \Sigma B \Sigma \\
& + \Sigma C' \Sigma A' \Sigma B' \Sigma + \Sigma C' \Sigma B \Sigma A \Sigma + \Sigma C' \Sigma A \Sigma B' \Sigma + \Sigma C' \Sigma A \Sigma \text{tr}(B \Sigma) \\
& + \Sigma \text{tr}(A \Sigma) \text{tr}(B \Sigma) \text{tr}(C \Sigma) + \Sigma \text{tr}(A \Sigma) \text{tr}(B \Sigma C' \Sigma) + \Sigma \text{tr}(A \Sigma B' \Sigma) \text{tr}(C \Sigma) \\
& + \Sigma \text{tr}(A \Sigma C \Sigma B' \Sigma) + \Sigma \text{tr}(A \Sigma C \Sigma B \Sigma) + \Sigma \text{tr}(A \Sigma C' \Sigma B \Sigma) + \Sigma \text{tr}(A \Sigma B' \Sigma C' \Sigma) \\
& + \Sigma \text{tr}(A \Sigma C' \Sigma) \text{tr}(B \Sigma).
\end{aligned}$$

Proof. We primarily complete the last two equations. The proofs of the other equations are in Gupta and Nagar [5]. At first, we consider $E[w_{ab}w_{cd}w_{ef}]$ to derive $E[WAWBW]$, where $W = (w_{ij})$. It should be noted that $w_{ij} \stackrel{d}{=} \sum_{r=1}^n y_{ir}y_{jr}$ where $\mathbf{y}_r = (y_{1r}, \dots, y_{pr})' \stackrel{i.i.d.}{\sim} N_p(\mathbf{0}, \Sigma)$, $r \in \{1, \dots, n\}$. By using the expectation formulas, we have

$$\begin{aligned}
E[w_{ab}w_{cd}w_{ef}] &= \sum_{r_1, r_2, r_3}^n E[y_{ar_1}y_{br_1}y_{cr_2}y_{dr_2}y_{er_3}y_{fr_3}] \\
&= \sum_r^n E[y_{ar}y_{br}y_{cr}y_{dr}y_{er}y_{fr}] \\
&\quad + \sum_{r_1 \neq r_3}^n E[y_{ar_1}y_{br_1}y_{cr_1}y_{dr_1}] E[y_{er_3}y_{fr_3}] \\
&\quad + \sum_{r_1 \neq r_2}^n E[y_{ar_1}y_{br_1}y_{er_1}y_{fr_1}] E[y_{cr_2}y_{dr_2}] \\
&\quad + \sum_{r_1 \neq r_2}^n E[y_{ar_1}y_{br_1}] E[y_{cr_2}y_{dr_2}y_{er_2}y_{fr_2}] \\
&\quad + \sum_{r_1 \neq r_2 \neq r_3 \neq r_1}^n E[y_{ar_1}y_{br_1}] E[y_{cr_2}y_{dr_2}] E[y_{er_3}y_{fr_3}] \\
&= n(\sigma_{ac}\sigma_{be}\sigma_{df} + \sigma_{ac}\sigma_{bf}\sigma_{de} + \sigma_{ad}\sigma_{be}\sigma_{cf} + \sigma_{ad}\sigma_{bf}\sigma_{ce} + \sigma_{ae}\sigma_{bc}\sigma_{df} \\
&\quad + \sigma_{ae}\sigma_{bd}\sigma_{cf} + \sigma_{af}\sigma_{bc}\sigma_{de} + \sigma_{af}\sigma_{bd}\sigma_{ce}) \\
&\quad + n^2 \left\{ (\sigma_{ac}\sigma_{bd} + \sigma_{ad}\sigma_{bc}) \sigma_{ef} + (\sigma_{ae}\sigma_{bf} + \sigma_{af}\sigma_{be}) \sigma_{cd} \right. \\
&\quad \left. + \sigma_{ab}(\sigma_{ce}\sigma_{df} + \sigma_{cf}\sigma_{de}) \right\} \\
&\quad + n^3 \sigma_{ab}\sigma_{cd}\sigma_{ef},
\end{aligned}$$

where $\Sigma = (\sigma_{ij})$. The expectations of (i, j) th element of $WAWBW$ is

$$\begin{aligned}
E[(WAWBW)_{ij}] &= \sum_{r_1, r_2, r_3, r_4}^p E[w_{ir_1}a_{r_1r_2}w_{r_2r_3}b_{r_3r_4}w_{r_4j}] \\
&= \sum_{r_1, r_2, r_3, r_4}^p a_{r_1r_2}b_{r_3r_4}E[w_{ir_1}w_{r_2r_3}w_{r_4j}],
\end{aligned}$$

which completes the result of $E[WAWBW]$, $A = (a_{ij})$, and $B = (b_{ij})$. By using $E[WAWBWCW]$ can be completed similarly to $E[WAWBW]$: because it holds that

$$E[w_{ab}w_{cd}w_{ef}w_{gh}]$$

$$\begin{aligned}
&= n^4 \sigma_{ab} \sigma_{cd} \sigma_{ef} \sigma_{gh} + n^3 (\sigma_{ab} \sigma_{cd} \sigma_{eg} \sigma_{fh} + \sigma_{ab} \sigma_{cd} \sigma_{eh} \sigma_{fg} + \sigma_{ab} \sigma_{ce} \sigma_{df} \sigma_{gh} \\
&\quad + \sigma_{ab} \sigma_{cf} \sigma_{de} \sigma_{gh} + \sigma_{ab} \sigma_{cg} \sigma_{dh} \sigma_{ef} + \sigma_{ab} \sigma_{ch} \sigma_{dg} \sigma_{ef} + \sigma_{ac} \sigma_{bd} \sigma_{ef} \sigma_{gh} \\
&\quad + \sigma_{ad} \sigma_{bc} \sigma_{ef} \sigma_{gh} + \sigma_{ae} \sigma_{bf} \sigma_{cd} \sigma_{gh} + \sigma_{af} \sigma_{be} \sigma_{cd} \sigma_{gh} + \sigma_{ag} \sigma_{bh} \sigma_{cd} \sigma_{ef} \\
&\quad + \sigma_{ah} \sigma_{bg} \sigma_{cd} \sigma_{ef}) \\
&\quad + n^2 (\sigma_{ab} \sigma_{ce} \sigma_{dg} \sigma_{fh} + \sigma_{ab} \sigma_{ce} \sigma_{dh} \sigma_{fg} + \sigma_{ab} \sigma_{cf} \sigma_{dg} \sigma_{eh} + \sigma_{ab} \sigma_{cf} \sigma_{dh} \sigma_{eg} \\
&\quad + \sigma_{ab} \sigma_{cg} \sigma_{de} \sigma_{fh} + \sigma_{ab} \sigma_{cg} \sigma_{df} \sigma_{eh} + \sigma_{ab} \sigma_{ch} \sigma_{de} \sigma_{fg} + \sigma_{ab} \sigma_{ch} \sigma_{df} \sigma_{eg} \\
&\quad + \sigma_{ac} \sigma_{bd} \sigma_{eg} \sigma_{fh} + \sigma_{ac} \sigma_{bd} \sigma_{eh} \sigma_{fg} + \sigma_{ac} \sigma_{be} \sigma_{df} \sigma_{gh} + \sigma_{ac} \sigma_{bf} \sigma_{de} \sigma_{gh} \\
&\quad + \sigma_{ac} \sigma_{bg} \sigma_{dh} \sigma_{ef} + \sigma_{ac} \sigma_{bh} \sigma_{dg} \sigma_{ef} + \sigma_{ad} \sigma_{bc} \sigma_{eg} \sigma_{fh} + \sigma_{ad} \sigma_{bc} \sigma_{eh} \sigma_{fg} \\
&\quad + \sigma_{ad} \sigma_{be} \sigma_{cf} \sigma_{gh} + \sigma_{ad} \sigma_{bf} \sigma_{ce} \sigma_{gh} + \sigma_{ad} \sigma_{bg} \sigma_{ch} \sigma_{ef} + \sigma_{ad} \sigma_{bh} \sigma_{cg} \sigma_{ef} \\
&\quad + \sigma_{ae} \sigma_{bc} \sigma_{df} \sigma_{gh} + \sigma_{ae} \sigma_{bd} \sigma_{cf} \sigma_{gh} + \sigma_{ae} \sigma_{bf} \sigma_{cg} \sigma_{dh} + \sigma_{ae} \sigma_{bf} \sigma_{ch} \sigma_{dg} \\
&\quad + \sigma_{ae} \sigma_{bg} \sigma_{cd} \sigma_{fh} + \sigma_{ae} \sigma_{bh} \sigma_{cd} \sigma_{fg} + \sigma_{af} \sigma_{bc} \sigma_{de} \sigma_{gh} + \sigma_{af} \sigma_{bd} \sigma_{ce} \sigma_{gh} \\
&\quad + \sigma_{af} \sigma_{be} \sigma_{cg} \sigma_{dh} + \sigma_{af} \sigma_{be} \sigma_{ch} \sigma_{dg} + \sigma_{af} \sigma_{bg} \sigma_{cd} \sigma_{eh} + \sigma_{af} \sigma_{bh} \sigma_{cd} \sigma_{eg} \\
&\quad + \sigma_{ag} \sigma_{bc} \sigma_{dh} \sigma_{ef} + \sigma_{ag} \sigma_{bd} \sigma_{ch} \sigma_{ef} + \sigma_{ag} \sigma_{be} \sigma_{cd} \sigma_{fh} + \sigma_{ag} \sigma_{bf} \sigma_{cd} \sigma_{eh} \\
&\quad + \sigma_{ag} \sigma_{bh} \sigma_{ce} \sigma_{df} + \sigma_{ag} \sigma_{bh} \sigma_{cf} \sigma_{de} + \sigma_{ah} \sigma_{bc} \sigma_{dg} \sigma_{ef} + \sigma_{ah} \sigma_{bd} \sigma_{cg} \sigma_{ef} \\
&\quad + \sigma_{ah} \sigma_{be} \sigma_{cd} \sigma_{fg} + \sigma_{ah} \sigma_{bf} \sigma_{cd} \sigma_{eg} + \sigma_{ah} \sigma_{bg} \sigma_{ce} \sigma_{df} + \sigma_{ah} \sigma_{bg} \sigma_{cf} \sigma_{de}) \\
&\quad + n (\sigma_{ac} \sigma_{be} \sigma_{dg} \sigma_{fh} + \sigma_{ac} \sigma_{be} \sigma_{dh} \sigma_{fg} + \sigma_{ac} \sigma_{bf} \sigma_{dg} \sigma_{eh} + \sigma_{ac} \sigma_{bf} \sigma_{dh} \sigma_{eg} \\
&\quad + \sigma_{ac} \sigma_{bg} \sigma_{de} \sigma_{fh} + \sigma_{ac} \sigma_{bg} \sigma_{df} \sigma_{eh} + \sigma_{ac} \sigma_{bh} \sigma_{de} \sigma_{fg} + \sigma_{ac} \sigma_{bh} \sigma_{df} \sigma_{eg} \\
&\quad + \sigma_{ad} \sigma_{be} \sigma_{cg} \sigma_{fh} + \sigma_{ad} \sigma_{be} \sigma_{ch} \sigma_{fg} + \sigma_{ad} \sigma_{bf} \sigma_{cg} \sigma_{eh} + \sigma_{ad} \sigma_{bf} \sigma_{ch} \sigma_{eg} \\
&\quad + \sigma_{ad} \sigma_{bg} \sigma_{ce} \sigma_{fh} + \sigma_{ad} \sigma_{bg} \sigma_{cf} \sigma_{eh} + \sigma_{ad} \sigma_{bh} \sigma_{ce} \sigma_{fg} + \sigma_{ad} \sigma_{bh} \sigma_{cf} \sigma_{eg} \\
&\quad + \sigma_{ae} \sigma_{bc} \sigma_{dg} \sigma_{fh} + \sigma_{ae} \sigma_{bc} \sigma_{dh} \sigma_{fg} + \sigma_{ae} \sigma_{bd} \sigma_{cg} \sigma_{fh} + \sigma_{ae} \sigma_{bd} \sigma_{ch} \sigma_{fg} \\
&\quad + \sigma_{ae} \sigma_{bg} \sigma_{cf} \sigma_{dh} + \sigma_{ae} \sigma_{bg} \sigma_{ch} \sigma_{df} + \sigma_{ae} \sigma_{bh} \sigma_{cf} \sigma_{dg} + \sigma_{ae} \sigma_{bh} \sigma_{cg} \sigma_{df} \\
&\quad + \sigma_{af} \sigma_{bc} \sigma_{dg} \sigma_{eh} + \sigma_{af} \sigma_{bc} \sigma_{dh} \sigma_{eg} + \sigma_{af} \sigma_{bd} \sigma_{cg} \sigma_{eh} + \sigma_{af} \sigma_{bd} \sigma_{ch} \sigma_{eg} \\
&\quad + \sigma_{af} \sigma_{bg} \sigma_{ce} \sigma_{dh} + \sigma_{af} \sigma_{bg} \sigma_{ch} \sigma_{de} + \sigma_{af} \sigma_{bh} \sigma_{ce} \sigma_{dg} + \sigma_{af} \sigma_{bh} \sigma_{cg} \sigma_{de} \\
&\quad + \sigma_{ag} \sigma_{bc} \sigma_{de} \sigma_{fh} + \sigma_{ag} \sigma_{bc} \sigma_{df} \sigma_{eh} + \sigma_{ag} \sigma_{bd} \sigma_{ce} \sigma_{fh} + \sigma_{ag} \sigma_{bd} \sigma_{cf} \sigma_{eh} \\
&\quad + \sigma_{ag} \sigma_{be} \sigma_{cf} \sigma_{dh} + \sigma_{ag} \sigma_{be} \sigma_{ch} \sigma_{df} + \sigma_{ag} \sigma_{bf} \sigma_{ce} \sigma_{dh} + \sigma_{ag} \sigma_{bf} \sigma_{ch} \sigma_{de} \\
&\quad + \sigma_{ah} \sigma_{bc} \sigma_{de} \sigma_{fg} + \sigma_{ah} \sigma_{bc} \sigma_{df} \sigma_{eg} + \sigma_{ah} \sigma_{bd} \sigma_{ce} \sigma_{fg} + \sigma_{ah} \sigma_{bd} \sigma_{cf} \sigma_{eg} \\
&\quad + \sigma_{ah} \sigma_{be} \sigma_{cf} \sigma_{dg} + \sigma_{ah} \sigma_{be} \sigma_{cg} \sigma_{df} + \sigma_{ah} \sigma_{bf} \sigma_{ce} \sigma_{dg} + \sigma_{ah} \sigma_{bf} \sigma_{cg} \sigma_{de}),
\end{aligned}$$

and $E[(WAWBWCW)_{ij}] = \sum_{r_1, r_2, \dots, r_6} a_{r_1 r_2} b_{r_3 r_4} c_{r_5 r_6} E[w_{ir_1} w_{r_2 r_3} w_{r_4 r_5} w_{r_6 j}]$, we can obtain $E[WAWBWCW]$. $\square$

Lemma A.4. Suppose that $W \sim W_p(n, \Sigma)$, $T_M = \sqrt{n} (n^{-1}W - \Sigma)$, A , B , and C are $p \times p$ dimensional constant matrices, respectively. Then, $E[T_M A T_M]$, $E[\text{tr}(A T_M) T_M]$, $E[T_M A T_M B T_M]$ and $E[T_M A T_M B T_M C T_M]$ can be obtained as

$$\begin{aligned}
E[T_M A T_M] &= \Sigma A' \Sigma + \text{tr}(A \Sigma) \Sigma, \\
E[T_M A T_M B T_M] &= \frac{1}{\sqrt{n}} \left\{ \Sigma A' \Sigma B' \Sigma + \Sigma A' \Sigma \text{tr}(B' \Sigma) + \Sigma B \Sigma A \Sigma \right.
\end{aligned}$$

$$\begin{aligned}
& + \Sigma B \Sigma' A' \Sigma + \Sigma B' \Sigma \text{tr}(A' \Sigma) + \Sigma B' \Sigma' A \Sigma \\
& + \Sigma \text{tr}(A' \Sigma) \text{tr}(B' \Sigma) + \Sigma \text{tr}(A \Sigma B' \Sigma) \Big\},
\end{aligned}$$

$$E[\text{tr}(AT_M)T_M] = \Sigma A \Sigma + \Sigma A' \Sigma,$$

$$\begin{aligned}
E[T_M AT_M B T_M C T_M] &= \Sigma A' \Sigma B \Sigma C' \Sigma + \Sigma A' \Sigma B \Sigma \text{tr}(C \Sigma) + \Sigma B \Sigma C' \Sigma \text{tr}(A \Sigma) \\
&+ \Sigma B \Sigma \text{tr}(A \Sigma) \text{tr}(C \Sigma) + \Sigma B' \Sigma \text{tr}(A \Sigma C' \Sigma) \\
&+ \Sigma B' \Sigma C' \Sigma A \Sigma + \Sigma C \Sigma A' \Sigma B' \Sigma \\
&+ \Sigma C \Sigma B \Sigma A \Sigma + \Sigma C' \Sigma B \Sigma A' \Sigma \\
&+ \Sigma C' \Sigma A' \Sigma \text{tr}(B \Sigma) + \Sigma \text{tr}(A \Sigma B' \Sigma C \Sigma) \\
&+ \Sigma \text{tr}(A \Sigma C \Sigma) \text{tr}(B \Sigma) + o(1),
\end{aligned}$$

respectively.

Proof. Substituting $W = \sqrt{n}T_M + n\Sigma$ into Lemma A.3 immediately proves this lemma. $\square$

Lemma A.5. Suppose that two-step monotone missing data

$\mathbf{x}_1^{(g)}, \dots, \mathbf{x}_{N_1^{(g)}}^{(g)}, \mathbf{x}_{(12)N_1^{(g)}+1}^{(g)}, \dots, \mathbf{x}_{(12)N^{(g)}}^{(g)}$ is observed from $\Pi^{(g)} : N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$ ($g \in \{1, 2\}$). Then, we denote that $T = \sum_{g=1}^2 \sum_{\ell=1}^2 \sqrt{m_\ell^{(g)}} T_{(12)}^{(g, \ell)}$ ($\ell \in \{1, 2\}$)

where

$$\begin{aligned}
T_{(12)}^{(g, \ell)} &= \sqrt{n_\ell^{(g)}} \Sigma_{(12)}^{-\frac{1}{2}} \left(\frac{1}{n_\ell^{(g)}} S_{(12)}^{(g, \ell)} - \Sigma_{(12)} \right) \Sigma_{(12)}^{-\frac{1}{2}'} \\
S^{(g, 1)} &= \sum_{j=1}^{N_1^{(g)}} \left(\mathbf{x}_j^{(g)} - \bar{\mathbf{x}}^{(g, 1)} \right) \left(\mathbf{x}_j^{(g)} - \bar{\mathbf{x}}^{(g, 1)} \right)', \quad S_{(12)}^{(g, 1)} = C_{(12)} S^{(g, 1)} C_{(12)}', \\
S^{(g, 2)} &= \sum_{j=N_1^{(g)}+1}^{N^{(g)}} \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g, 2)} \right) \left(\mathbf{x}_{(12)j}^{(g)} - \bar{\mathbf{x}}_{(12)}^{(g, 2)} \right)', \\
\bar{\mathbf{x}}^{(g, 1)} &= \frac{1}{N_1^{(g)}} \sum_{j=1}^{N_1^{(g)}} \mathbf{x}_j^{(g)}, \quad \bar{\mathbf{x}}_{(12)}^{(g, 1)} = C_{(12)} \bar{\mathbf{x}}^{(g, 1)}, \quad \bar{\mathbf{x}}_{(12)}^{(g, 2)} = \frac{1}{N_2^{(g)}} \sum_{j=N_1^{(g)}+1}^{N^{(g)}} \mathbf{x}_{(12)j}^{(g)}, \\
N_2^{(g)} &= N^{(g)} - N_1^{(g)}, \quad N_1 = N_1^{(1)} + N_1^{(2)}, \quad N = \sum_{g=1}^2 \sum_{\ell=1}^2 N_\ell^{(g)}, \\
m_\ell^{(g)} &= \frac{N_\ell^{(g)}}{N}, \quad m_1 = \frac{N_1}{N}, \quad m^{(g)} = \frac{N^{(g)}}{N}, \quad n_\ell^{(g)} = N_\ell^{(g)} - 1,
\end{aligned}$$

$\Sigma^{-\frac{1}{2}}$ and $\Sigma_{(12)}^{-\frac{1}{2}}$ are square root inverse matrix of Σ and $\Sigma_{(12)}$, A , B , and C are $(p_1 + p_2) \times$

$(p_1 + p_2)$ constant symmetric matrices, respectively. Then, the following expectations can be obtained:

$$\begin{aligned}
E[T] &= O_{(p_1+p_2) \times (p_1+p_2)}, \\
E[TAT] &= A' + \text{tr}(A)I_{p_1+p_2}, \\
E[TATBT] &= \frac{1}{\sqrt{N}} \left\{ A'B' + A'\text{tr}(B') + BA + BA' + B'\text{tr}(A') \right. \\
&\quad \left. + B'A + I_{p_1+p_2}\text{tr}(A')\text{tr}(B') + I_{p_1+p_2}\text{tr}(AB') \right\} + o(N^{-1/2}), \\
E[TATBTCT] &= A'BC' + A'B\text{tr}(C) + BC'\text{tr}(A) \\
&\quad + B\text{tr}(A)\text{tr}(C) + B'\text{tr}(AC') + B'C'A + CA'B' + CBA \\
&\quad + C'BA' + C'A'\text{tr}(B) + I_{p_1+p_2}\text{tr}(AB'C) \\
&\quad + I_{p_1+p_2}\text{tr}(AC)\text{tr}(B) + o(1), \\
E[TL] &= 2(p_1 + p_2 + 1)I_{p_1+p_2}.
\end{aligned}$$

Proof. $T_{(12)}^{(g,\ell)}$ and $T_{(12)}^{(g',\ell')}$ are symmetric matrices. Therefore $T_{(12)}^{(g,\ell)}$ and $T_{(12)}^{(g',\ell')}$ are commute. Because $T_{(12)}^{(g,\ell)}$ and $T_{(12)}^{(g',\ell')}$ are mutually independent, and it holds that $E[T_{(12)}^{(g,\ell)}] = O$, by using Lemma A.4.

$$\begin{aligned}
E[TAT] &= \sum_{g=1}^2 \sum_{\ell=1}^2 m_{\ell}^{(g)} E \left[T_{(12)}^{(g,\ell)} A T_{(12)}^{(g,\ell)} \right] \\
&= \left(\sum_{g=1}^2 \sum_{\ell=1}^2 m_{\ell}^{(g)} \right) (A' + \text{tr}(A)I_{p_1+p_2}) \\
&= A' + \text{tr}(A)I_{p_1+p_2}
\end{aligned}$$

holds. Furthermore, we can evaluate

$$\begin{aligned}
E[TATBT] &= \sum_{g=1}^2 \sum_{\ell=1}^2 m_{\ell}^{(g)} \sqrt{m_{\ell}^{(g)}} E \left[T_{(12)}^{(g,\ell)} A T_{(12)}^{(g,\ell)} B T_{(12)}^{(g,\ell)} \right] \\
&= \sum_{g=1}^2 \sum_{\ell=1}^2 m_{\ell}^{(g)} \sqrt{m_{\ell}^{(g)}} \frac{1}{\sqrt{n_{\ell}^{(g)}}} \left\{ A'B' + A'\text{tr}(B') + BA \right. \\
&\quad \left. + BA' + B'\text{tr}(A') + B'A + I_{p_1+p_2}\text{tr}(A')\text{tr}(B') \right. \\
&\quad \left. + I_{p_1+p_2}\text{tr}(AB') \right\} \\
&= \frac{1}{\sqrt{N}} \left\{ A'B' + A'\text{tr}(B') + BA + BA' + B'\text{tr}(A') + B'A \right. \\
&\quad \left. + I_{p_1+p_2}\text{tr}(A')\text{tr}(B') + I_{p_1+p_2}\text{tr}(AB') \right\} + o(N^{-1/2}),
\end{aligned}$$

$$\begin{aligned}
& E[TATBTCT] \\
&= \sum_{g=1}^2 \sum_{\ell=1}^2 m_{\ell}^{(g)2} E[T_{(12)}^{(g,\ell)} AT_{(12)}^{(g,\ell)} BT_{(12)}^{(g,\ell)} CT_{(12)}^{(g,\ell)}] \\
&+ \sum_{(g,\ell)} \sum_{(g',\ell')} m_{\ell}^{(g)} m_{\ell'}^{(g')} E \left[T_{(12)}^{(g,\ell)} AT_{(12)}^{(g,\ell)} BT_{(12)}^{(g',\ell')} CT_{(12)}^{(g',\ell')} \right] \\
&+ \sum_{(g,\ell)} \sum_{(g',\ell')} m_{\ell}^{(g)} m_{\ell'}^{(g')} E \left[T_{(12)}^{(g,\ell)} AT_{(12)}^{(g',\ell')} BT_{(12)}^{(g',\ell')} CT_{(12)}^{(g,\ell)} \right] \\
&+ \sum_{(g,\ell)} \sum_{(g',\ell')} m_{\ell}^{(g)} m_{\ell'}^{(g')} E \left[T_{(12)}^{(g,\ell)} AT_{(12)}^{(g',\ell')} BT_{(12)}^{(g,\ell)} CT_{(12)}^{(g',\ell')} \right] \\
&= \left(\sum_{(g,\ell)} m_{\ell}^{(g)2} \right) \left\{ A'BC' + A'B\text{tr}(C) + BC'\text{tr}(A) \right. \\
&+ B\text{tr}(A)\text{tr}(C) + B'\text{tr}(AC') + B'C'A' + CA'B' + CBA \\
&+ C'BA' + C'A'\text{tr}(B) + I_{p_1+p_2}\text{tr}(AB'C) + I_{p_1+p_2}\text{tr}(AC)\text{tr}(B) \left. \right\} \\
&+ \sum_{(g,\ell) \neq (g',\ell')} m_{\ell}^{(g)} m_{\ell'}^{(g')} \left\{ A'BC' + A'B\text{tr}(C) + BC'\text{tr}(A) + B\text{tr}(A)\text{tr}(C) \right\} \\
&+ \sum_{(g,\ell) \neq (g',\ell')} m_{\ell}^{(g)} m_{\ell'}^{(g')} \left\{ B'\text{tr}(AC') + B'C'A + CA'B' + CBA \right\} \\
&+ \sum_{(g,\ell) \neq (g',\ell')} m_{\ell}^{(g)} m_{\ell'}^{(g')} \left\{ C'BA' + \text{tr}(B')C'A' \right. \\
&+ \text{tr}(AB'C)I_{p_1+p_2} + \text{tr}(AC)\text{tr}(B)I_{p_1+p_2} \left. \right\} \\
&= A'BC' + A'B\text{tr}(C) + BC'\text{tr}(A) + B\text{tr}(A)\text{tr}(C) \\
&+ B'\text{tr}(AC') + B'C'A + CA'B' + CBA + C'BA' \\
&+ C'A'\text{tr}(B) + \text{tr}(AB'C)I_{p_1+p_2} + \text{tr}(AC)\text{tr}(B)I_{p_1+p_2},
\end{aligned}$$

$$\begin{aligned}
TL &= \left(\sum_{g=1}^2 \sum_{\ell=1}^2 \sqrt{m_{\ell}^{(g)}} T_{(12)}^{(g,\ell)} \right) \left\{ \sum_{g=1}^2 \sum_{\ell=1}^2 \left(4\sqrt{m_{\ell}^{(g)}} - \frac{1}{2\sqrt{m_{\ell}^{(g)}}} \right) T_{(12)}^{(g,\ell)} \right\} \\
&= \sum_{g=1}^2 \sum_{\ell=1}^2 \left(4m_{\ell}^{(g)} - \frac{1}{2} \right) \left(T_{(12)}^{(g,\ell)} \right)^2 \\
&+ \sum_{(g,\ell)} \sum_{(g',\ell')} \sqrt{m_{\ell}^{(g)}} \left(4\sqrt{m_{\ell'}^{(g')}} - \frac{1}{2\sqrt{m_{\ell'}^{(g')}}} \right) T_{(12)}^{(g,\ell)} T_{(12)}^{(g',\ell')}
\end{aligned}$$

$$\begin{aligned}
E[TL] &= \sum_{g=1}^2 \sum_{\ell=1}^2 \left(4m_{\ell}^{(g)} - \frac{1}{2} \right) E \left[\left(T_{(12)}^{(g,\ell)} \right)^2 \right] \\
&= 2(p_1 + p_2 + 1)I_{p_1+p_2}.
\end{aligned}$$

With the above results, the proof is complete. $\square$

Corollary A.1. Suppose that A , B , and C are $p \times p$ constant matrices,

$$M = \frac{1}{m_1} \sum_{g=1}^2 \sqrt{m_1^{(g)}} T^{(g,1)}, \quad L^* = \frac{1}{2m_1} \sum_{g=1}^2 \frac{1}{\sqrt{m_1^{(g)}}} T^{(g,1)}.$$

Then, the following expectations can be obtained:

$$\begin{aligned}
E[M] &= O_p, \\
E[MAM] &= \frac{1}{m_1} \{A' + \text{tr}(A)I_p\}, \\
E[MAMB M] &= \frac{1}{m_1^2} \frac{1}{\sqrt{N}} \left\{ A'B' + A'\text{tr}(B') + BA + BA' \right. \\
&\quad \left. + B'\text{tr}(A') + B'A + I_{p_1+p_2} \text{tr}(A')\text{tr}(B') \right. \\
&\quad \left. + I_{p_1+p_2} \text{tr}(AB') \right\} + o(N^{-1/2}), \\
E[MAMB MCM] &= \frac{1}{m_1^2} \left\{ A'BC' + A'B\text{tr}(C) + BC'\text{tr}(A) + B\text{tr}(A)\text{tr}(C) \right. \\
&\quad \left. + B'\text{tr}(AC') + B'C'A' + CA'B' + CBA + C'BA' \right. \\
&\quad \left. + C'A'\text{tr}(B) + I_{p_1+p_2} \text{tr}(AB'C) + I_{p_1+p_2} \text{tr}(AC)\text{tr}(B) \right\} \\
&\quad + o(1), \\
E[MAL^*] &= \frac{1}{m_1^2} \{A' + \text{tr}(A)I_p\}.
\end{aligned}$$

Proof. Equations are can be obtained in the same way as result of Lemma A.5. $\square$

References

- [1] H. Akaike (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, **19**, 716–723.
- [2] T.W. Anderson and I. Olkin (1985). Maximum-likelihood estimation of the parameters of a multivariate normal distribution, *Linear Algebra and its Applications*, **70**, 147–171.
- [3] W. Y. Chang and D. St. P. Richards (2009). Finite-sample inference with monotone incomplete multivariate normal data, *Journal of Multivariate Analysis*, **100**, 1883–1899.
- [4] Y. Fujikoshi (1985). Selection of Variables in Two-group Discriminant analysis by error rate and akaike’s information criteria, *Journal of Multivariate Analysis*, **17**, 27–37.
- [5] A. K. Gupta and D. K. Nagar (1999). *Matrix Variate Distributions*, Chapman & Hall, New York.
- [6] M. Hyodo and T. Kubokawa (2014). A variable selection criterion for linear discriminant rule and its optimality in high dimensional and large sample data, *Journal of Multivariate Analysis*, **123**, 364–379.
- [7] T. Kanda and Y. Fujikoshi (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.
- [8] K. Krishnamoorthy and M. K. Pannala (1999). Confidence estimation of a normal mean vector with incomplete data, *The Canadian Journal of Statistics*, **27**, 395–407.
- [9] Y. Okada, N. Matsuchi, M. Hyodo, N. Shutoh (2024+). Redundancy test for some variables in linear discriminant analysis with two-step monotone incomplete data, (submitted to *SUT Journal of Mathematics*).
- [10] N. Shutoh (2012). An asymptotic approximation for EPMC in linear discriminant analysis based on monotone missing data, *Journal of Statistical Planning and Inference*, **142**, 110–125.
- [11] N. Shutoh (2012). An asymptotic expansion for the distribution of the linear discriminant function based on monotone missing data, *Journal of Statistical Computation and Simulation*, **82**, 241–259.
- [12] N. Shutoh, M. Hyodo and T. Seo (2011). An asymptotic approximation for EPMC in linear discriminant analysis based on two-step monotone missing samples, *Journal*

- of Multivariate Analysis*, **102**, 252–263.
- [13] A. Yagi and T. Seo (2017). Tests for normal mean vectors with monotone incomplete data, *American Journal of Mathematical and Management Sciences*, **36**, 1–20.
- [14] J. Yu, K. Krishnamoorthy and M. K. Pannala (2006). Two-sample inference for normal mean vectors based on monotone missing data, *Journal of Multivariate Analysis*, **97**, 2162–2176.