Journal of Multivariate Analysis 101 (2010) 603–620

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Finite-sample inference with monotone incomplete multivariate normal data, II

Wan-Ying Chang [a], Donald St. P. Richards [b,c,*]

[a] *Washington Department of Fish and Wildlife, Olympia, WA 98501, USA*
[b] *Department of Statistics, Penn State University, University Park, PA 16802, USA*
[c] *The Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, NC 27709, USA*

ARTICLE INFO

*Article history:*
Received 5 January 2008
Available online 30 September 2009

*AMS 2000 subject classifications:*
Primary 62H10
secondary 60D10
62E15

*Keywords:*
Likelihood ratio tests
Locally most powerful invariant tests
Matrix $F$-distribution
Maximum likelihood estimation
Missing completely at random
Multivariate analysis of variance
Testing independence
Sphericity test
Unbiased test statistics
Wishart distribution

ABSTRACT

We continue our recent work on inference with two-step, monotone incomplete data from a multivariate normal population with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Under the assumption that $\boldsymbol{\Sigma}$ is block-diagonal when partitioned according to the two-step pattern, we derive the distributions of the diagonal blocks of $\hat{\boldsymbol{\Sigma}}$ and of the estimated regression matrix, $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}$. We represent $\hat{\boldsymbol{\Sigma}}$ in terms of independent matrices; derive its exact distribution, thereby generalizing the Wishart distribution to the setting of monotone incomplete data; and obtain saddlepoint approximations for the distributions of $\hat{\boldsymbol{\Sigma}}$ and its partial Iwasawa coordinates. We prove the unbiasedness of a modified likelihood ratio criterion for testing $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, where $\boldsymbol{\Sigma}_0$ is a given matrix, and obtain the null and non-null distributions of the test statistic. In testing $H_0 : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are given, we prove that the likelihood ratio criterion is unbiased and obtain its null and non-null distributions. For the sphericity test, $H_0 : \boldsymbol{\Sigma} \propto \boldsymbol{I}_{p+q}$, we obtain the null distribution of the likelihood ratio criterion. In testing $H_0 : \boldsymbol{\Sigma}_{12} = \boldsymbol{0}$ we show that a modified locally most powerful invariant statistic has the same distribution as a Bartlett–Pillai–Nanda trace statistic in multivariate analysis of variance.



## 1. Introduction

In this paper, we continue our work in [1] on inference with two-step, monotone incomplete, multivariate normal data that are of the form

$$\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{Y}_1 \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_2 \\ \boldsymbol{Y}_2 \end{pmatrix} \cdots \begin{pmatrix} \boldsymbol{X}_n \\ \boldsymbol{Y}_n \end{pmatrix} \quad \boldsymbol{Y}_{n+1} \quad \boldsymbol{Y}_{n+2} \quad \cdots \quad \boldsymbol{Y}_N, \tag{1.1}$$

where each $\boldsymbol{X}_j$ is $p \times 1$, each $\boldsymbol{Y}_j$ is $q \times 1$, the complete data $(\boldsymbol{X}_j', \boldsymbol{Y}_j')'$, $j = 1, \ldots, n$, are drawn from $\mathrm{N}_{p+q}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, a multivariate normal population with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, the incomplete data $\boldsymbol{Y}_j$, $j = n+1, \ldots, N$, are observations on the last $q$ characteristics of the population, and all $N$ observations are mutually independent.

Closed-form expressions for $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, the maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, are well-known (Anderson [2], Anderson and Olkin [3], Giguère and Styan [4], Jinadasa and Tracy [5]), and those formulas have been utilized in inference for

* Corresponding author at: Department of Statistics, Penn State University, University Park, PA 16802, USA.
*E-mail address:* richards@stat.psu.edu (D.St.P. Richards).

0047-259X/$ – see front matter 
doi:10.1016/j.jmva.2009.09.011

$\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (Bhargava [6,7], Morrison [8], Little and Rubin [9], Kanda and Fujikoshi [10]); we note that a closed-form expression for $\hat{\boldsymbol{\Sigma}}$ requires the assumption that data are missing completely at random, an assumption stated and discussed in [1] and maintained here. In this paper, we continue our program of research on finite-sample inference for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by means of results on the exact distributions of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$; having derived in [1] the exact distribution of $\hat{\boldsymbol{\mu}}$ and made applications to inference for $\boldsymbol{\mu}$, we now turn our attention to inference for $\boldsymbol{\Sigma}$.

Assuming that $\boldsymbol{\Sigma}$ is block-diagonal when partitioned into $p \times p$ and $q \times q$ submatrices, we derive in Section 3 the distributions of the diagonal blocks of $\hat{\boldsymbol{\Sigma}}$ and the estimated regression matrix, $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}$. We also obtain a stochastic representation for $\hat{\boldsymbol{\Sigma}}$ and derive its exact distribution, thereby extending the Wishart distribution to the setting of monotone incomplete data, and we obtain saddlepoint approximations for $\hat{\boldsymbol{\Sigma}}$ and its partial Iwasawa coordinates.

In Section 4, we consider four tests of hypotheses on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. For $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, where $\boldsymbol{\Sigma}_0$ is specified, we derive the non-null moments of the likelihood ratio criterion and a stochastic representation for its null distribution, and we show that the criterion is not unbiased; we also construct a modified likelihood ratio criterion, and prove unbiasedness and a monotonicity property of its power function. In the case of $H_0 : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ is given, we prove that the likelihood ratio criterion is unbiased, and derive its non-null moments and its null distribution. For the sphericity test, $H_0 : \boldsymbol{\Sigma} \propto \boldsymbol{I}_{p+q}$, the identity matrix, we derive the null moments and distribution of the likelihood ratio criterion. In testing independence between the first $p$ and last $q$ characteristics of the population, Eaton and Kariya [11] derived a locally most powerful invariant criterion; the null distribution theory of that statistic appearing to be recondite, we modify it and prove that the modified statistic is distributed as a Bartlett–Pillai–Nanda trace statistic in multivariate analysis of variance.

## 2. Preliminary results

We retain throughout this paper the notation and conventions of [1], writing all vectors and matrices in boldface type. We denote by $\mathbf{0}$ any zero vector or matrix, the dimension of which will be clear from the context, and we denote by $\boldsymbol{I}_d$ the identity matrix of order $d$. We write $\boldsymbol{A} > \mathbf{0}$ to denote that a matrix $\boldsymbol{A}$ is positive definite (symmetric), and we write $\boldsymbol{A} \geq \boldsymbol{B}$ to mean that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite. We write $\boldsymbol{W} \sim W_d(a, \boldsymbol{\Lambda})$, a Wishart distribution, with $a > d - 1$ and $\boldsymbol{\Lambda} > \mathbf{0}$, whenever $\boldsymbol{W}$ is a $d \times d$ random matrix with density function

$$\frac{1}{2^{ad/2}|\boldsymbol{\Lambda}|^{a/2}\Gamma_d(a/2)}|\boldsymbol{W}|^{\frac{1}{2}a-\frac{1}{2}(p+1)} \exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{\Lambda}^{-1}\boldsymbol{W}\right), \tag{2.1}$$

$\boldsymbol{W} > \mathbf{0}$, where the multivariate gamma function [12, p. 62] is

$$\Gamma_d(a) = \pi^{d(d-1)/4}\prod_{j=1}^{d}\Gamma\left(a - \frac{1}{2}(j-1)\right). \tag{2.2}$$

We partition $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in conformity with (1.1), writing $\boldsymbol{\mu} = \begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}$ where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are of dimensions $p$ and $q$, respectively, and $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}'$, and $\boldsymbol{\Sigma}_{22}$ are of orders $p \times p$, $p \times q$, and $q \times q$, respectively. We assume throughout that $n > q + 2$ to ensure that all means and variances are finite and that all integrals arising are absolutely convergent. We denote by $\tau$ the proportion, $n/N$, of data which are complete and denote $1 - \tau$ by $\bar{\tau}$.

Define sample means

$$\bar{\boldsymbol{X}} = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{X}_j, \qquad \bar{\boldsymbol{Y}}_1 = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{Y}_j, \qquad \bar{\boldsymbol{Y}}_2 = \frac{1}{N-n}\sum_{j=n+1}^{N}\boldsymbol{Y}_j, \qquad \bar{\boldsymbol{Y}} = \frac{1}{N}\sum_{j=1}^{N}\boldsymbol{Y}_j,$$

and corresponding matrices of sums of squares and products,

$$\begin{aligned} \boldsymbol{A}_{11} &= \sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}})(\boldsymbol{X}_j - \bar{\boldsymbol{X}})', \qquad \boldsymbol{A}_{12} = \boldsymbol{A}_{21}' = \sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}})(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1)', \\ \boldsymbol{A}_{22,n} &= \sum_{j=1}^{n}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1)(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1)', \qquad \boldsymbol{A}_{22,N} = \sum_{j=1}^{N}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})'. \end{aligned} \tag{2.3}$$

By [2,3,8,5], the maximum likelihood estimator of $\boldsymbol{\mu}$ is $\hat{\boldsymbol{\mu}} = \begin{pmatrix}\hat{\boldsymbol{\mu}}_1\\ \hat{\boldsymbol{\mu}}_2\end{pmatrix}$, where

$$\hat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{X}} - \bar{\tau}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2), \qquad \hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{Y}}. \tag{2.4}$$

## 3. The distribution of $\hat{\boldsymbol{\Sigma}}$

Let $\boldsymbol{A}_{11\cdot 2,n} := \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$. By [2,3] (cf. [4,8]), the maximum likelihood estimator of $\boldsymbol{\Sigma}$ is $\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\boldsymbol{\Sigma}}_{11} & \hat{\boldsymbol{\Sigma}}_{12} \\ \hat{\boldsymbol{\Sigma}}_{21} & \hat{\boldsymbol{\Sigma}}_{22} \end{pmatrix}$, where

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_{11} &= \frac{1}{n}\boldsymbol{A}_{11\cdot 2,n} + \frac{1}{N}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{22,N}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}, \\ \hat{\boldsymbol{\Sigma}}_{12} &= \hat{\boldsymbol{\Sigma}}_{21}' = \frac{1}{N}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{22,N}, \\ \hat{\boldsymbol{\Sigma}}_{22} &= \frac{1}{N}\boldsymbol{A}_{22,N}. \end{aligned} \tag{3.1}$$

### 3.1. A representation for $\hat{\boldsymbol{\Sigma}}$

**Proposition 3.1.** *With the notation above, we have*

$$n\hat{\boldsymbol{\Sigma}} = \tau \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22,n} \end{pmatrix} + \bar{\tau} \begin{pmatrix} \boldsymbol{A}_{11\cdot 2,n} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} + \tau \begin{pmatrix} \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_q \end{pmatrix} \begin{pmatrix} \boldsymbol{B} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_q \end{pmatrix}, \tag{3.2}$$

*where* $\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22,n} \end{pmatrix} \sim W_{p+q}(n-1, \boldsymbol{\Sigma})$ *and* $\boldsymbol{B} \sim W_q(N-n, \boldsymbol{\Sigma}_{22})$ *are mutually independent. Moreover,* $N\hat{\boldsymbol{\Sigma}}_{22} \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$.

**Proof.** We write $\boldsymbol{A}_{22,N}$ in the form

$$\boldsymbol{A}_{22,N} = \sum_{j=1}^{n}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1 + \bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1 + \bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}})' + \sum_{j=n+1}^{N}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_2 + \bar{\boldsymbol{Y}}_2 - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_2 + \bar{\boldsymbol{Y}}_2 - \bar{\boldsymbol{Y}})',$$

and expand each term as a sum of products to obtain

$$\boldsymbol{A}_{22,N} = \boldsymbol{A}_{22,n} + \boldsymbol{B}, \tag{3.3}$$

where

$$\boldsymbol{B} = \boldsymbol{B}_1 + \boldsymbol{B}_2 \tag{3.4}$$

with

$$\boldsymbol{B}_1 = \sum_{j=n+1}^{N}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_2)(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_2)' \tag{3.5}$$

and

$$\boldsymbol{B}_2 = \frac{n(N-n)}{N}(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2)(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2)'.$$

Substituting (3.3) and (3.4) into (3.1), we obtain (3.2). (For $p = 1$, (3.3) is due to Morrison [8, Eq. (3.4)].)

By the independence of the sample mean and covariance matrix, and the independence of the individual observations in (1.1), the matrix

$$\sum_{j=1}^{n} \begin{pmatrix} \boldsymbol{X}_j - \bar{\boldsymbol{X}} \\ \boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1 \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_j - \bar{\boldsymbol{X}} \\ \boldsymbol{Y}_j - \bar{\boldsymbol{Y}}_1 \end{pmatrix}' \equiv \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22,n} \end{pmatrix}$$

is independent of $\{\bar{\boldsymbol{Y}}_1, \boldsymbol{Y}_{n+1}, \ldots, \boldsymbol{Y}_N\}$. Therefore $\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22,n} \end{pmatrix}$ is independent of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, and hence also is independent of $\boldsymbol{B}$.

Note also that $\boldsymbol{A}_{22,n}$, $\boldsymbol{B}_1$, and $\boldsymbol{B}_2$ are mutually independent Wishart matrices, with $\boldsymbol{A}_{22,n} \sim W_q(n-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{B}_1 \sim W_q(N-n-1, \boldsymbol{\Sigma}_{22})$, and $\boldsymbol{B}_2 \sim W_q(1, \boldsymbol{\Sigma}_{22})$. Therefore, by (3.4), $\boldsymbol{B} \sim W_q(N-n, \boldsymbol{\Sigma}_{22})$ and hence $N\hat{\boldsymbol{\Sigma}}_{22} = \boldsymbol{A}_{22,N} = \boldsymbol{A}_{22,n} + \boldsymbol{B} \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$. □

We now establish some results whose proofs were postponed from [1, Section 4].

**Proposition 3.2.** *Suppose that* $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$. *Then* $\boldsymbol{A}_{22,n}$, $\boldsymbol{A}_{11\cdot 2,n}$, $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$, $\boldsymbol{B}_1$, $\bar{\boldsymbol{X}}$, $\bar{\boldsymbol{Y}}_1$, *and* $\bar{\boldsymbol{Y}}_2$ *are mutually independent. Also,* $\boldsymbol{B}_2$ *and* $\bar{\boldsymbol{Y}}$ *are independent.*

**Proof.** By the independence of the mean and covariance matrix of a normal random sample, and by the mutual independence of the data, we see that $\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22,n} \end{pmatrix}$ and $\{\boldsymbol{B}_1, \bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}_1, \bar{\boldsymbol{Y}}_2\}$ are mutually independent. Since $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$ then $\bar{\boldsymbol{X}}$ is independent of $\{\boldsymbol{B}_1, \bar{\boldsymbol{Y}}_1, \bar{\boldsymbol{Y}}_2\}$ and, by [13, pp. 142–143], the matrices $\boldsymbol{A}_{22,n}$, $\boldsymbol{A}_{11\cdot 2,n}$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$ are mutually independent. Thus, $\boldsymbol{A}_{22,n}, \boldsymbol{A}_{11\cdot 2,n}, \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}, \bar{\boldsymbol{X}}$ and $\{\boldsymbol{B}_1, \bar{\boldsymbol{Y}}_1, \bar{\boldsymbol{Y}}_2\}$ are mutually independent.

Next, $\bar{\boldsymbol{Y}}_1$ and $\{\boldsymbol{B}_1, \bar{\boldsymbol{Y}}_2\}$ are mutually independent since they are constructed from disjoint sets of independent observations. And by again applying the independence of the mean and covariance matrix of a normal random sample, we see that $\boldsymbol{B}_1$ is independent of $\bar{\boldsymbol{Y}}_2$. Therefore $\boldsymbol{A}_{22,n}, \boldsymbol{A}_{11\cdot 2,n}$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}, \boldsymbol{B}_1, \bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}_1$, and $\bar{\boldsymbol{Y}}_2$ are mutually independent.

Finally, we show that $\boldsymbol{B}_2$ is independent of $\bar{\boldsymbol{Y}}$. Since $\boldsymbol{B}_2 \propto (\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2)(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2)'$ then we need only show that $\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2$ is independent of $\bar{\boldsymbol{Y}}$. The pair $(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2, \bar{\boldsymbol{Y}})$, being a linear function of $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N$, is jointly normally distributed; hence, to establish their independence, it suffices to verify that $E(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2)(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}_2)'$, their cross-covariance matrix, is zero. We write this matrix in the form

$$E(\bar{\boldsymbol{Y}}_1 - \bar{\boldsymbol{Y}}_2)(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}_2)' = E\left((\bar{\boldsymbol{Y}}_1 - \boldsymbol{\mu}_2) - (\bar{\boldsymbol{Y}}_2 - \boldsymbol{\mu}_2)\right)\left(\tau(\bar{\boldsymbol{Y}}_1 - \boldsymbol{\mu}_2) + \bar{\tau}(\bar{\boldsymbol{Y}}_2 - \boldsymbol{\mu}_2)\right)',$$

expand the right-hand side, and evaluate the expectation of all four terms in that expansion. For $j, k = 1, 2$, $E(\bar{\boldsymbol{Y}}_j - \boldsymbol{\mu}_2)(\bar{\boldsymbol{Y}}_k - \boldsymbol{\mu}_2)'$ equals $\boldsymbol{0}$ if $j \neq k$ and equals $\mathrm{Cov}(\bar{\boldsymbol{Y}}_j)$ if $j = k$; hence the cross-covariance matrix equals $\tau\mathrm{Cov}(\bar{\boldsymbol{Y}}_1) - \bar{\tau}\mathrm{Cov}(\bar{\boldsymbol{Y}}_2) = (\tau n^{-1} - \bar{\tau}(N-n)^{-1})\boldsymbol{\Sigma}_{22} = \boldsymbol{0}$, since $\tau n^{-1} = \bar{\tau}(N-n)^{-1} = N^{-1}$. The proof now is complete. □

For the remainder of this section, we assume that $p \leq q$. As in [1, Section 4], we denote by $O(q)$ the group of all $q \times q$ orthogonal matrices, and by $S_{p,q}$ the *Stiefel manifold* of all $p \times q$ matrices $\boldsymbol{H}_1$ such that $\boldsymbol{H}_1\boldsymbol{H}_1' = \boldsymbol{I}_p$. As noted in [1], the uniform distribution on $S_{p,q}$ is the unique probability distribution which is left-invariant under $O(p)$ and right-invariant under $O(q)$. If a random matrix $\boldsymbol{H} \in O(q)$ is distributed according to the Haar probability measure, and if we write $\boldsymbol{H}$ in the form $\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{pmatrix}$, where $\boldsymbol{H}_1 \in S_{p,q}$, then $\boldsymbol{H}_1$ is uniformly distributed on $S_{p,q}$. Conversely, a uniformly distributed $\boldsymbol{H}_1 \in S_{p,q}$ may be completed to form a random $q \times q$ orthogonal matrix $\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{pmatrix}$ having the Haar probability distribution on $O(q)$.

A $q \times q$ random matrix $\boldsymbol{F} \geq \boldsymbol{0}$ is said to have a *matrix F-distribution*, denoted as $\boldsymbol{F} \sim F_{a,b}^{(q)}$, with degrees of freedom $(a, b)$, $a \geq 0, b > q-1$, if $\boldsymbol{F} \stackrel{\mathcal{L}}{=} \boldsymbol{B}^{-1/2}\boldsymbol{A}\boldsymbol{B}^{-1/2}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are mutually independent Wishart matrices with $\boldsymbol{A} \sim \mathrm{W}_q(a, \boldsymbol{\Sigma}_{22})$ and $\boldsymbol{B} \sim \mathrm{W}_q(b, \boldsymbol{\Sigma}_{22})$. If $a \leq q-1$ then $\boldsymbol{A}$ is singular, so $\boldsymbol{F}$ also is singular, almost surely. If both $a, b > q-1$ then $\boldsymbol{F}$ is nonsingular, almost surely, and its density function is

$$\frac{\Gamma_q((a+b)/2)}{\Gamma_q(a/2)\Gamma_q(b/2)}|\boldsymbol{F}|^{\frac{1}{2}a-\frac{1}{2}(q+1)/2}|\boldsymbol{I}_q + \boldsymbol{F}|^{-(a+b)/2},$$

$\boldsymbol{F} > \boldsymbol{0}$. From this result, we see that the distribution of $\boldsymbol{F}$ is orthogonally invariant, i.e., $\boldsymbol{F} \stackrel{\mathcal{L}}{=} \boldsymbol{H}\boldsymbol{F}\boldsymbol{H}'$ for $\boldsymbol{H} \in O(q)$. It is also well-known [12, pp. 312–313] that if $\boldsymbol{A}$ and $\boldsymbol{B}$ are independent nonsingular Wishart matrices with $\boldsymbol{A} \sim \mathrm{W}_q(a, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{B} \sim \mathrm{W}_q(b, \boldsymbol{\Sigma}_{22})$ then both $\boldsymbol{A}^{1/2}\boldsymbol{B}^{-1}\boldsymbol{A}^{1/2}$ and $\boldsymbol{B}^{-1/2}\boldsymbol{A}\boldsymbol{B}^{-1/2}$ are distributed as $F_{a,b}^{(q)}$. Further, if $\boldsymbol{F} \sim F_{a,b}^{(q)}$ then $\boldsymbol{F}^{-1} \sim F_{b,a}^{(q)}$. If $\boldsymbol{F} \sim F_{a,b}^{(q)}$ then, assuming without loss of generality that $\boldsymbol{\Sigma}_{22} = \boldsymbol{I}_q$, we obtain $|\boldsymbol{F}| \stackrel{\mathcal{L}}{=} |\boldsymbol{A}|/|\boldsymbol{B}|$; recalling that $|\boldsymbol{A}| \stackrel{\mathcal{L}}{=} \prod_{j=1}^{q} \chi_{a-j+1}^2$, a product of independent chi-squared variables, with a similar result also holding for $|\boldsymbol{B}|$, we obtain $|\boldsymbol{F}| \stackrel{\mathcal{L}}{=} \prod_{j=1}^{p} F_{a-j+1,b-j+1}^{(1)}$.

**Lemma 3.3.** *Let* $\boldsymbol{F} \sim F_{a,b}^{(q)}$, $\boldsymbol{H}_1$ *be uniformly distributed on* $S_{p,q}$, *and* $\boldsymbol{F}$ *and* $\boldsymbol{H}_1$ *be independent. Then* $\boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' \sim F_{a,b-q+p}^{(p)}$ *and* $\boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' \stackrel{\mathcal{L}}{=} \boldsymbol{F}_{11}$, *the upper* $p \times p$ *principal submatrix of* $\boldsymbol{F}$.

**Proof.** By augmenting $\boldsymbol{H}_1$ to a Haar-distributed matrix $\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{pmatrix}$ on $O(q)$, we obtain

$$\begin{pmatrix} \boldsymbol{F}_{11} & \boldsymbol{F}_{12} \\ \boldsymbol{F}_{21} & \boldsymbol{F}_{22} \end{pmatrix} \equiv \boldsymbol{F} \stackrel{\mathcal{L}}{=} \boldsymbol{H}\boldsymbol{F}\boldsymbol{H}' = \begin{pmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{pmatrix} \boldsymbol{F} \begin{pmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{pmatrix}' = \begin{pmatrix} \boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' & \boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_2' \\ \boldsymbol{H}_2\boldsymbol{F}\boldsymbol{H}_1' & \boldsymbol{H}_2\boldsymbol{F}\boldsymbol{H}_2' \end{pmatrix},$$

proving that $\boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' \stackrel{\mathcal{L}}{=} \boldsymbol{F}_{11}$. Next, since $\boldsymbol{F} \sim F_{a,b}^{(q)}$ then $\boldsymbol{F} \stackrel{\mathcal{L}}{=} \boldsymbol{A}^{1/2}\boldsymbol{B}^{-1}\boldsymbol{A}^{1/2}$, where $\boldsymbol{A} \sim \mathrm{W}_q(a, \boldsymbol{I}_q)$, $\boldsymbol{B} \sim \mathrm{W}_q(b, \boldsymbol{I}_q)$, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are independent. Then, with $\boldsymbol{M} = (\boldsymbol{I}_p \vdots \boldsymbol{0})\boldsymbol{A}^{1/2}$,

$$\boldsymbol{F}_{11} = (\boldsymbol{I}_p \vdots \boldsymbol{0})\boldsymbol{F}(\boldsymbol{I}_p \vdots \boldsymbol{0})' \stackrel{\mathcal{L}}{=} (\boldsymbol{I}_p \vdots \boldsymbol{0})\boldsymbol{A}^{1/2}\boldsymbol{B}^{-1}\boldsymbol{A}^{1/2}(\boldsymbol{I}_p \vdots \boldsymbol{0})' \equiv \boldsymbol{M}\boldsymbol{B}^{-1}\boldsymbol{M}'.$$

By [12, p. 95], conditional on $\boldsymbol{M}$, $(\boldsymbol{M}\boldsymbol{B}^{-1}\boldsymbol{M}')^{-1} \sim \mathrm{W}_p(b-q+p, (\boldsymbol{M}\boldsymbol{M}')^{-1})$; hence the conditional density function of $\boldsymbol{R} = \boldsymbol{F}_{11}^{-1}$ given $\boldsymbol{S} = \boldsymbol{M}\boldsymbol{M}'$ is

$$f(\boldsymbol{R}|\boldsymbol{S}) = \text{const.} \times |\boldsymbol{R}|^{\frac{1}{2}(b-q+p)-\frac{1}{2}(p+1)}|\boldsymbol{S}|^{\frac{1}{2}(b-q+p)} \exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{S}\boldsymbol{R}\right),$$

$\boldsymbol{R}, \boldsymbol{S} > \boldsymbol{0}$. Since $\boldsymbol{S} = \boldsymbol{M}\boldsymbol{M}' = (\boldsymbol{I}_p \,\vdots\, \boldsymbol{0})\boldsymbol{A}(\boldsymbol{I}_p \,\vdots\, \boldsymbol{0}) \equiv \boldsymbol{A}_{11} \sim \mathrm{W}_p(a, \boldsymbol{I}_p)$ then the joint density function of $\boldsymbol{R}$ and $\boldsymbol{S}$ is

$$\begin{aligned} f(\boldsymbol{R}, \boldsymbol{S}) &= f(\boldsymbol{R}|\boldsymbol{S}) f(\boldsymbol{S}) \\ &\propto |\boldsymbol{R}|^{\frac{1}{2}(b-q+p)-\frac{1}{2}(p+1)} |\boldsymbol{S}|^{\frac{1}{2}(b-q+p)} \exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{S}\boldsymbol{R}\right) \cdot |\boldsymbol{S}|^{\frac{1}{2}(a-p-1)} \exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{S}\right) \\ &= |\boldsymbol{R}|^{\frac{1}{2}(b-q+p)-\frac{1}{2}(p+1)} |\boldsymbol{S}|^{\frac{1}{2}(a+b-q+p)-\frac{1}{2}(p+1)} \exp\left(-\frac{1}{2}\operatorname{tr}(\boldsymbol{I}_p + \boldsymbol{R})\boldsymbol{S}\right) \end{aligned}$$

for $\boldsymbol{R}, \boldsymbol{S} > \boldsymbol{0}$. Integrating over $\boldsymbol{S}$, we obtain the density function of $\boldsymbol{R}$ as

$$f(\boldsymbol{R}) = \text{const.} \times |\boldsymbol{R}|^{\frac{1}{2}(b-q+p)-\frac{1}{2}(p+1)} |\boldsymbol{I}_p + \boldsymbol{R}|^{-\frac{1}{2}(a+b-q+p)},$$

$\boldsymbol{R} > \boldsymbol{0}$. Therefore $\boldsymbol{R} \sim F^{(p)}_{b-q+p,a}$, so $\boldsymbol{F}_{11} = \boldsymbol{R}^{-1} \sim F^{(p)}_{a,b-q+p}$. □

**Proposition 3.4.** *Suppose that* $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$. *Then*

$$\boldsymbol{\Sigma}_{11}^{-1/2} \hat{\boldsymbol{\Sigma}}_{11} \boldsymbol{\Sigma}_{11}^{-1/2} \overset{\mathcal{L}}{=} \frac{1}{n}\boldsymbol{W}_1 + \frac{1}{N}\boldsymbol{W}_2^{1/2}(\boldsymbol{I}_p + \boldsymbol{F})\boldsymbol{W}_2^{1/2}, \tag{3.6}$$

*where* $\boldsymbol{W}_1 \sim \mathrm{W}_p(n-q-1, \boldsymbol{I}_p)$, $\boldsymbol{W}_2 \sim \mathrm{W}_p(q, \boldsymbol{I}_p)$, $\boldsymbol{F} \sim F^{(p)}_{N-n,n-q+p-1}$, *and* $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, *and* $\boldsymbol{F}$ *are independent.*

**Proof.** Let $\boldsymbol{W}_1 = \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{A}_{11\cdot 2,n}\boldsymbol{\Sigma}_{11}^{-1/2}$ and $\boldsymbol{K} = \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1/2}$. By (3.1) and (3.3),

$$\begin{aligned} N\boldsymbol{\Sigma}_{11}^{-1/2} \hat{\boldsymbol{\Sigma}}_{11} \boldsymbol{\Sigma}_{11}^{-1/2} &\overset{\mathcal{L}}{=} \frac{N}{n}\boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{A}_{11\cdot 2,n}\boldsymbol{\Sigma}_{11}^{-1/2} + \boldsymbol{\Sigma}_{11}^{-1/2}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}(\boldsymbol{A}_{22,n} + \boldsymbol{B})\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\boldsymbol{\Sigma}_{11}^{-1/2} \\ &= \frac{N}{n}\boldsymbol{W}_1 + \boldsymbol{K}(\boldsymbol{I}_q + \boldsymbol{A}_{22,n}^{-1/2}\boldsymbol{B}\boldsymbol{A}_{22,n}^{-1/2})\boldsymbol{K}'. \end{aligned}$$

Since $\boldsymbol{A}_{11\cdot 2,n} \sim \mathrm{W}_p(n-q-1, \boldsymbol{\Sigma}_{11})$ then $\boldsymbol{W}_1 \sim \mathrm{W}_p(n-q-1, \boldsymbol{I}_p)$. By [12, p. 93] $\boldsymbol{A}_{11\cdot 2,n}$, and hence $\boldsymbol{W}_1$, is independent of $\{\boldsymbol{A}_{12}, \boldsymbol{A}_{22,n}\}$ and $\boldsymbol{B}$. Since $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$ then $\boldsymbol{K}|\boldsymbol{A}_{22,n} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$ and, because this conditional distribution does not depend on $\boldsymbol{A}_{22,n}$, it is the unconditional distribution of $\boldsymbol{K}$. Therefore $\boldsymbol{W}_1$, $\boldsymbol{K}$, $\boldsymbol{A}_{22,n}$, and $\boldsymbol{B}$ are mutually independent.

Note also that the distribution of $\boldsymbol{K}$ is right-invariant under $O(q)$, i.e., $\boldsymbol{K} \overset{\mathcal{L}}{=} \boldsymbol{K}\boldsymbol{H}$ for all $\boldsymbol{H} \in O(q)$. By polar coordinates on matrix space ([14, p. 482], [15, p. 163]), $\boldsymbol{K} \overset{\mathcal{L}}{=} \boldsymbol{W}_2^{1/2}\boldsymbol{H}_1$ where $\boldsymbol{W}_2$ and $\boldsymbol{H}_1$ are independent, $\boldsymbol{W}_2 = \boldsymbol{K}\boldsymbol{K}' \sim \mathrm{W}_p(q, \boldsymbol{I}_p)$, and $\boldsymbol{H}_1$ is uniformly distributed on the Stiefel manifold $S_{p,q}$ [12, pp. 67–72].

Since $\boldsymbol{B} \sim \mathrm{W}_q(N-n, \boldsymbol{\Sigma}_{22})$ and $\boldsymbol{A}_{22,n} \sim \mathrm{W}_q(n-1, \boldsymbol{\Sigma}_{22})$ then $\boldsymbol{F} = \boldsymbol{A}_{22,n}^{-1/2}\boldsymbol{B}\boldsymbol{A}_{22,n}^{-1/2} \sim F^{(q)}_{N-n,n-1}$. Therefore

$$\begin{aligned} N\boldsymbol{\Sigma}_{11}^{-1/2} \hat{\boldsymbol{\Sigma}}_{11} \boldsymbol{\Sigma}_{11}^{-1/2} &\overset{\mathcal{L}}{=} \frac{N}{n}\boldsymbol{W}_1 + \boldsymbol{W}_2^{1/2}\boldsymbol{H}_1(\boldsymbol{I}_q + F)\boldsymbol{H}_1'\boldsymbol{W}_2^{1/2} \\ &= \frac{N}{n}\boldsymbol{W}_1 + \boldsymbol{W}_2^{1/2}(\boldsymbol{I}_p + \boldsymbol{H}_1 F\boldsymbol{H}_1')\boldsymbol{W}_2^{1/2}. \end{aligned}$$

By Lemma 3.3, $\boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' \sim F^{(p)}_{N-n,n-q+p-1}$, and the proof now is complete. □

**Remark 3.5.** Since the $F$-matrix in (3.6) is positive semidefinite, it follows that the right-hand side of (3.6) is stochastically greater than $\boldsymbol{W}_1 + \boldsymbol{W}_2$ in the sense that the difference

$$\frac{N}{n}\boldsymbol{W}_1 + \boldsymbol{W}_2^{1/2}(\boldsymbol{I}_p + \boldsymbol{F})\boldsymbol{W}_2^{1/2} - (\boldsymbol{W}_1 + \boldsymbol{W}_2) = \frac{N-n}{n}\boldsymbol{W}_1 + \boldsymbol{W}_2^{1/2}\boldsymbol{F}\boldsymbol{W}_2^{1/2}$$

is positive semidefinite, almost surely; we write this as

$$N\boldsymbol{\Sigma}_{11}^{-1/2} \hat{\boldsymbol{\Sigma}}_{11} \boldsymbol{\Sigma}_{11}^{-1/2} \overset{\mathcal{L}}{\geq} \frac{N}{n}\boldsymbol{W}_1 + \boldsymbol{W}_2 \overset{\mathcal{L}}{\geq} \boldsymbol{W}_1 + \boldsymbol{W}_2 \sim \mathrm{W}_p(n-1, \boldsymbol{I}_p).$$

Hence, we obtain the stochastic ordering $N^p|\hat{\boldsymbol{\Sigma}}_{11}|/|\boldsymbol{\Sigma}_{11}| \overset{\mathcal{L}}{\geq} |\boldsymbol{W}_1 + \boldsymbol{W}_2|$, so for all $\delta \geq 0$,

$$P(N^p|\hat{\boldsymbol{\Sigma}}_{11}|/|\boldsymbol{\Sigma}_{11}| \geq \delta) \geq P(|\boldsymbol{W}_1 + \boldsymbol{W}_2| \geq \delta).$$

As an application, we construct a one-sided confidence interval for $|\boldsymbol{\Sigma}_{11}|$ when $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$. Since $\boldsymbol{W}_1 + \boldsymbol{W}_2 \sim \mathrm{W}_p(n-1, \boldsymbol{I}_p)$ then $|\boldsymbol{W}_1 + \boldsymbol{W}_2|$ is distributed according to a product of independent chi-squared variables. If $\delta_\alpha$ is an upper $\alpha$% percentage point for $|\boldsymbol{W}_1 + \boldsymbol{W}_2|$, i.e., $P(|\boldsymbol{W}_1 + \boldsymbol{W}_2| \geq \delta_\alpha) = \alpha$, then

$$P(N^p|\hat{\boldsymbol{\Sigma}}_{11}|/|\boldsymbol{\Sigma}_{11}| \geq \delta_\alpha) \geq P(|\boldsymbol{W}_1 + \boldsymbol{W}_2| \geq \delta_\alpha) = \alpha.$$

Therefore the interval $(0, N^p|\hat{\boldsymbol{\Sigma}}_{11}|/\delta_\alpha)$ is a one-sided confidence interval for $|\boldsymbol{\Sigma}_{11}|$ with confidence level at least $100(1-\alpha)$%.

### 3.2. The distribution of the estimated regression matrix

We now consider the marginal distribution of $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}$ and some of its properties, making no assumptions about $\boldsymbol{\Sigma}_{12}$.

**Theorem 3.6.** *The distribution of* $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}$ *satisfies the stochastic representation*

$$\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} \stackrel{\mathcal{L}}{=} \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} + \boldsymbol{\Sigma}_{11\cdot 2}^{1/2}\boldsymbol{W}^{-1/2}\boldsymbol{K}\boldsymbol{\Sigma}_{22}^{-1/2}, \tag{3.7}$$

*where* $\boldsymbol{W}$ *and* $\boldsymbol{K}$ *are independent,* $\boldsymbol{W} \sim \mathrm{W}_p(n-q+p-1, \boldsymbol{I}_p)$, *and* $\boldsymbol{K} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$. *In particular,* $E(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}) = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$.

**Proof.** By (3.1), $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}$. Define $\boldsymbol{B}_{12} = \boldsymbol{A}_{12} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{A}_{22,n}$; then it is easily seen that $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} = \boldsymbol{B}_{12}\boldsymbol{A}_{22,n}^{-1} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$, or equivalently, $\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} = \boldsymbol{B}_{12}\boldsymbol{A}_{22,n}^{-1}$. By proceeding as in the proof of Theorem 3.1 in [1], we obtain $\boldsymbol{B}_{12}|\boldsymbol{A}_{22,n} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{11\cdot 2} \otimes \boldsymbol{A}_{22,n})$. Therefore, $\boldsymbol{B}_{12} \stackrel{\mathcal{L}}{=} \boldsymbol{\Sigma}_{11\cdot 2}^{1/2}\boldsymbol{K}\boldsymbol{A}_{22,n}^{1/2}$ where $\boldsymbol{K} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$, and it follows that

$$\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}) = \boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}\boldsymbol{B}_{12}\boldsymbol{A}_{22,n}^{-1} \stackrel{\mathcal{L}}{=} \boldsymbol{K}\boldsymbol{A}_{22,n}^{-1/2}. \tag{3.8}$$

Let $\boldsymbol{T} \in \mathbb{R}^{p\times q}$, the space of all $p \times q$ matrices; then the characteristic function of (3.8) is

$$\begin{aligned} E\exp(\mathrm{itr}\,\boldsymbol{T}'\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})) &= E\exp(\mathrm{itr}\,\boldsymbol{T}'\boldsymbol{K}\boldsymbol{A}_{22,n}^{-1/2}) \\ &= E\exp(\mathrm{itr}\,(\boldsymbol{T}\boldsymbol{A}_{22,n}^{-1/2})'\boldsymbol{K}) \\ &= E\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{A}_{22,n}^{-1/2}\boldsymbol{T}'\boldsymbol{T}\boldsymbol{A}_{22,n}^{-1/2}\right) \\ &= E\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{T}'\boldsymbol{T}\boldsymbol{A}_{22,n}^{-1}\right). \end{aligned}$$

Since $\boldsymbol{A}_{22,n} \sim \mathrm{W}_q(n-1, \boldsymbol{\Sigma}_{22})$ then this characteristic function equals

$$\frac{2^{-(n-1)q/2}\,|\boldsymbol{\Sigma}_{22}|^{-\frac{1}{2}(n-1)}}{\Gamma_q((n-1)/2)}\int_{\boldsymbol{A}_{22,n}>\boldsymbol{0}}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{T}'\boldsymbol{T}\boldsymbol{A}_{22,n}^{-1}\right)|\boldsymbol{A}_{22,n}|^{\frac{1}{2}(n-1)-\frac{1}{2}(q+1)}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{A}_{22,n}\right)\mathrm{d}\boldsymbol{A}_{22,n}.$$

This integral can be expressed in terms of $\boldsymbol{B}_{\delta}^{(q)}$, the Bessel function of matrix argument of the second kind defined by Herz [14]. Applying a formula from [14, p. 506], we have

$$B_{\delta}^{(q)}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2) = |\boldsymbol{\Lambda}_1|^{-\delta}\int_{\boldsymbol{W}>\boldsymbol{0}}\exp(-\mathrm{tr}\,(\boldsymbol{W}\boldsymbol{\Lambda}_1 + \boldsymbol{W}^{-1}\boldsymbol{\Lambda}_2))\,|\boldsymbol{W}|^{-\delta-\frac{1}{2}(q+1)}\mathrm{d}\boldsymbol{W}, \tag{3.9}$$

where $\boldsymbol{W}$, $\boldsymbol{\Lambda}_1$, and $\boldsymbol{\Lambda}_2$ are $q \times q$ positive definite matrices, so it follows that

$$E\exp(\mathrm{itr}\,\boldsymbol{T}'\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})) = \frac{1}{\Gamma_q((n-1)/2)}B_{-\frac{1}{2}(n-1)}^{(q)}\left(\frac{1}{4}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\boldsymbol{T}\right).$$

Since $\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\boldsymbol{T}$ and $\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'$ have the same set of non-zero eigenvalues and hence the same rank then, by [14, p. 509, Theorem 5.10],

$$B_{-\frac{1}{2}(n-1)}^{(q)}\left(\frac{1}{4}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\boldsymbol{T}\right) = \frac{\Gamma_q((n-1)/2)}{\Gamma_p((n-q+p-1)/2)}B_{-\frac{1}{2}(n-q+p-1)}^{(p)}\left(\frac{1}{4}\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\right),$$

and therefore

$$E\exp(\mathrm{itr}\,\boldsymbol{T}'\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})) = \frac{1}{\Gamma_p((n-q+p-1)/2)}B_{-\frac{1}{2}(n-q+p-1)}^{(p)}\left(\frac{1}{4}\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\right).$$

On applying (3.9) to express this Bessel function as an integral over the space of $p \times p$ positive definite matrices, we obtain

$$E\exp(\mathrm{itr}\,\boldsymbol{T}'\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})) = E\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\boldsymbol{W}^{-1}\right), \tag{3.10}$$

$\boldsymbol{W} \sim \mathrm{W}_p(n-q+p-1, \boldsymbol{I}_p)$. However the right-hand side of (3.10) equals

$$\begin{aligned} E\exp\left(-\frac{1}{2}\mathrm{tr}\,(\boldsymbol{W}^{-1/2}\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1/2})(\boldsymbol{W}^{-1/2}\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1/2})'\right) &= E\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Sigma}_{22}^{-1/2}\boldsymbol{T}'\boldsymbol{W}^{-1/2}\boldsymbol{K}\right) \\ &= E\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{T}'\boldsymbol{W}^{-1/2}\boldsymbol{K}\boldsymbol{\Sigma}_{22}^{-1/2}\right), \end{aligned}$$

$\boldsymbol{K} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$. Equivalently, $\boldsymbol{\Sigma}_{11\cdot2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}) \stackrel{\mathcal{L}}{=} \boldsymbol{W}^{-1/2}\boldsymbol{K}\boldsymbol{\Sigma}_{22}^{-1/2}$ and we then obtain (3.7). Finally, by taking expectations in (3.7) we obtain $E(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1}) = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$. □

**Remark 3.7.** We note that, by (3.7),

$$\boldsymbol{\Sigma}_{11\cdot2}^{-1/2}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\boldsymbol{\Sigma}_{22}(\hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})'\boldsymbol{\Sigma}_{11\cdot2}^{-1/2} \stackrel{\mathcal{L}}{=} \boldsymbol{W}^{-1/2}(\boldsymbol{K}\boldsymbol{K}')\boldsymbol{W}^{-1/2}. \tag{3.11}$$

Since $\boldsymbol{K}\boldsymbol{K}' \sim \mathrm{W}_p(q, \boldsymbol{I}_p)$ then the right-hand side of (3.11) has an $F^{(p)}_{q,n-q+p-1}$ distribution.

### *3.3. The distributions of $\hat{\Sigma}$ and $\hat{\Delta}$*

Let $\boldsymbol{\Sigma}$ be partitioned as before, and let $\boldsymbol{\Delta}_{11} = \boldsymbol{\Sigma}_{11\cdot2}$, $\boldsymbol{\Delta}_{12} = \boldsymbol{\Delta}_{21}' = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$, and $\boldsymbol{\Delta}_{22} = \boldsymbol{\Sigma}_{22}$ be the *partial Iwasawa coordinates* of $\boldsymbol{\Sigma}$ [16], and set $\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} \end{pmatrix}$. There is a unique correspondence between $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$, and also between $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\Delta}}$, the corresponding maximum likelihood estimators [15, *loc. cit.*]. Moreover, $\hat{\boldsymbol{\Delta}} := \begin{pmatrix} \hat{\boldsymbol{\Delta}}_{11} & \hat{\boldsymbol{\Delta}}_{12} \\ \hat{\boldsymbol{\Delta}}_{21} & \hat{\boldsymbol{\Delta}}_{22} \end{pmatrix}$ where, by (3.1),

$$\hat{\boldsymbol{\Delta}}_{11} = \hat{\boldsymbol{\Sigma}}_{11\cdot2} = \frac{1}{n}\boldsymbol{A}_{11\cdot2,n}, \qquad \hat{\boldsymbol{\Delta}}_{12} = \hat{\boldsymbol{\Sigma}}_{12}\hat{\boldsymbol{\Sigma}}_{22}^{-1} = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}, \qquad \hat{\boldsymbol{\Delta}}_{22} = \hat{\boldsymbol{\Sigma}}_{22} = \frac{1}{N}\boldsymbol{A}_{22,N}. \tag{3.12}$$

To obtain $f_{\hat{\boldsymbol{\Delta}}}$, the density function of $\hat{\boldsymbol{\Delta}}$, we need a preliminary result.

**Lemma 3.8.** *Let $\Xi_1$, $\Xi_2$, and $\Xi_3$ be absolutely continuous random matrices of the same dimension such that $(\Xi_1, \Xi_2)$ and $\Xi_3$ are independent. Then the conditional density function of $\Xi_1$ given $\Xi_2 + \Xi_3 = \boldsymbol{\xi}$, is*

$$f_{\Xi_1|\Xi_2+\Xi_3=\xi}(\boldsymbol{\xi}_1) = \frac{1}{f_{\Xi_2+\Xi_3}(\boldsymbol{\xi})}\int f_{\Xi_1|\Xi_2=\xi_2}(\boldsymbol{\xi}_1)f_{\Xi_2}(\boldsymbol{\xi}_2)f_{\Xi_3}(\boldsymbol{\xi}-\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2. \tag{3.13}$$

**Proof.** By a direct calculation,

$$\begin{aligned} f_{\Xi_2+\Xi_3}(\boldsymbol{\xi})f_{\Xi_1|\Xi_2+\Xi_3=\xi}(\boldsymbol{\xi}_1) &= f_{\Xi_1,\Xi_2+\Xi_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}) \\ &= \int f_{\Xi_1,\Xi_2,\Xi_3}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}-\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2 \\ &= \int f_{\Xi_1,\Xi_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)f_{\Xi_3}(\boldsymbol{\xi}-\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2 \\ &= \int f_{\Xi_1|\Xi_2=\xi_2}(\boldsymbol{\xi}_1)f_{\Xi_2}(\boldsymbol{\xi}_2)f_{\Xi_3}(\boldsymbol{\xi}-\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2. \end{aligned}$$

Dividing both sides of this equation by $f_{\Xi_2+\Xi_3}(\boldsymbol{\xi})$ completes the proof. □

In deriving the distribution of $\hat{\boldsymbol{\Delta}}$ we shall need the multivariate beta function,

$$B_q(a, b) = \frac{\Gamma_q(a)\Gamma_q(b)}{\Gamma_q(a+b)}, \tag{3.14}$$

$\mathrm{Re}(a), \mathrm{Re}(b) > (q-1)/2$; and the confluent hypergeometric function of matrix argument,

$${}_1F_1^{(q)}(a; b; \boldsymbol{M}) = \frac{1}{B_q(a, b-a)}\int_{\boldsymbol{0}<\boldsymbol{U}<\boldsymbol{I}_q} |\boldsymbol{U}|^{a-\frac{1}{2}(q+1)}|\boldsymbol{I}_q - \boldsymbol{U}|^{b-a-\frac{1}{2}(q+1)}\exp(\mathrm{tr}\,\boldsymbol{M}\boldsymbol{U})\mathrm{d}\boldsymbol{U}, \tag{3.15}$$

where $\boldsymbol{M}$ is $q \times q$ and symmetric; $\mathrm{Re}(b-a), \mathrm{Re}(a) > (q-1)/2$; and the region $\{\boldsymbol{0} < \boldsymbol{U} < \boldsymbol{I}_q\}$ consists of all $q \times q$ matrices $\boldsymbol{U}$ such that $\boldsymbol{U}$ and $\boldsymbol{I}_q - \boldsymbol{U}$ both are positive definite ([14], [12, p. 264]). For general $a$, $b$, the functions (3.15) satisfy the reduction formula

$${}_1F_1^{(q)}(a; a; \boldsymbol{M}) = \exp(\mathrm{tr}\,\boldsymbol{M}), \tag{3.16}$$

and Kummer's formula ([14, Eq. (2.8)], [12, p. 265]),

$${}_1F_1^{(q)}(a; b; \boldsymbol{M}) = \exp(\mathrm{tr}\,\boldsymbol{M})\,{}_1F_1^{(q)}(b-a; b; -\boldsymbol{M}). \tag{3.17}$$

If $\boldsymbol{M}$ is of rank $p \le q$, and $\boldsymbol{M}_0$ is any $p \times p$ symmetric matrix whose non-zero eigenvalues coincide with those of $\boldsymbol{M}$ then by Herz [14, Theorems 3.10, p. 497 and 4.15, p. 505],

$${}_1F_1^{(q)}(a; b; \boldsymbol{M}) = {}_1F_1^{(p)}(a; b; \boldsymbol{M}_0). \tag{3.18}$$

**Theorem 3.9.** *Let $n > p+q$ and $N-n > q-1$. Then $f_{\hat{\boldsymbol{\Delta}}}$, the density function of $\hat{\boldsymbol{\Delta}}$, evaluated at $\boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} \end{pmatrix}$, a $(p+q)\times(p+q)$ positive definite matrix, is*

$$f_{\hat{\boldsymbol{\Delta}}}(\boldsymbol{T}) = f_{\hat{\boldsymbol{\Delta}}_{11}}(\boldsymbol{T}_{11}) f_{\hat{\boldsymbol{\Delta}}_{22}}(\boldsymbol{T}_{22}) f_{\hat{\boldsymbol{\Delta}}_{12}|\hat{\boldsymbol{\Delta}}_{22}=\boldsymbol{T}_{22}}(\boldsymbol{T}_{12}), \tag{3.19}$$

*where the marginal density of $\hat{\boldsymbol{\Delta}}_{11}$ is*

$$f_{\hat{\boldsymbol{\Delta}}_{11}}(\boldsymbol{T}_{11}) = \frac{(\frac{1}{2}n)^{(n-q-1)p/2} |\boldsymbol{T}_{11}|^{\frac{1}{2}(n-q-1)-\frac{1}{2}(p+1)} \exp\left(-\frac{1}{2}n \operatorname{tr} \boldsymbol{T}_{11}\boldsymbol{\Delta}_{11}^{-1}\right)}{|\boldsymbol{\Delta}_{11}|^{(n-q-1)/2}\Gamma_p((n-q-1)/2)}, \tag{3.20}$$

*the marginal density of $\hat{\boldsymbol{\Delta}}_{22}$ is*

$$f_{\hat{\boldsymbol{\Delta}}_{22}}(\boldsymbol{T}_{22}) = \frac{(\frac{1}{2}N)^{(N-1)q/2} |\boldsymbol{T}_{22}|^{\frac{1}{2}(N-1)-\frac{1}{2}(q+1)} \exp\left(-\frac{1}{2}N \operatorname{tr} \boldsymbol{T}_{22}\boldsymbol{\Delta}_{22}^{-1}\right)}{|\boldsymbol{\Delta}_{22}|^{(N-1)/2}\Gamma_q((N-1)/2)}, \tag{3.21}$$

*and the conditional density function of $\hat{\boldsymbol{\Delta}}_{12}$ given $\hat{\boldsymbol{\Delta}}_{22}$ is*

$$\begin{aligned} f_{\hat{\boldsymbol{\Delta}}_{12}|\hat{\boldsymbol{\Delta}}_{22}=\boldsymbol{T}_{22}}(\boldsymbol{T}_{12}) &= (2\pi)^{-pq/2} 2^{-q(N-1)/2} N^{q(N+p-1)/2} \frac{\Gamma_q(\frac{1}{2}(n+p-1))}{\Gamma_q(\frac{1}{2}(n-1))\Gamma_q(\frac{1}{2}(N+p-1))} \\ &\quad \times |\boldsymbol{\Delta}_{11}|^{-q/2} |\boldsymbol{\Delta}_{22}|^{-(N-1)/2} \exp\left(-\frac{1}{2}\operatorname{tr} N\boldsymbol{\Delta}_{22}^{-1}\boldsymbol{T}_{22}\right) |\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)} \\ &\quad \times {}_1F_1^{(p)}\left(\frac{1}{2}(n+p-1); \frac{1}{2}(N+p-1); -\frac{1}{2}N\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\right). \end{aligned} \tag{3.22}$$

**Proof.** By Proposition 3.1,

$$\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22,n} \end{pmatrix} \sim W_{p+q}(\boldsymbol{\Sigma}, n-1); \tag{3.23}$$

consequently, by [12, p. 93], $\boldsymbol{A}_{11\cdot 2,n}$ and $\{\boldsymbol{A}_{12}, \boldsymbol{A}_{22,n}\}$ are mutually independent, and hence so are $\boldsymbol{A}_{11\cdot 2,n}$ and $\{\boldsymbol{A}_{12}, \boldsymbol{A}_{22,N}\}$. Therefore $\hat{\boldsymbol{\Delta}}_{11}$ and $\{\hat{\boldsymbol{\Delta}}_{12}, \hat{\boldsymbol{\Delta}}_{22}\}$ are mutually independent, so the joint density of $\hat{\boldsymbol{\Delta}}$ is of the form (3.19). By (3.23), $n\hat{\boldsymbol{\Delta}}_{11} = \boldsymbol{A}_{11\cdot 2,n} \sim W_p(n-q-1, \boldsymbol{\Delta}_{11})$ and then (3.20) is obtained by a transformation of the Wishart density (2.1). Also, since $N\hat{\boldsymbol{\Delta}}_{22} = \boldsymbol{A}_{22,N} \sim W_q(N-1, \boldsymbol{\Delta}_{22})$ then (3.21) is obtained similarly.

By [12, p. 93], $\hat{\boldsymbol{\Delta}}_{12}|\boldsymbol{A}_{22,n} = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}|\boldsymbol{A}_{22,n} \sim N(\boldsymbol{\Delta}_{12}, \boldsymbol{\Delta}_{11} \otimes \boldsymbol{A}_{22,n}^{-1})$. Therefore, for $\boldsymbol{T}_{12} \in \mathbb{R}^{p\times q}$ and a $q \times q$ matrix $\boldsymbol{U} > \boldsymbol{0}$,

$$\begin{aligned} f_{\hat{\boldsymbol{\Delta}}_{12}|N^{-1}\boldsymbol{A}_{22,n}=\boldsymbol{U}}(\boldsymbol{T}_{12}) &\equiv f_{\hat{\boldsymbol{\Delta}}_{21}|\boldsymbol{A}_{22,n}=N\boldsymbol{U}}(\boldsymbol{T}_{12}) \\ &= (2\pi)^{-pq/2} |\boldsymbol{\Delta}_{11}|^{-q/2} N^{pq/2} |\boldsymbol{U}|^{p/2} \exp\left(-\frac{1}{2}N \operatorname{tr} \boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{U}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\right). \end{aligned} \tag{3.24}$$

Since $\boldsymbol{A}_{22,n} \sim W_q(n-1, \boldsymbol{\Delta}_{22})$ then $N^{-1}\boldsymbol{A}_{22,n}$ has density function

$$f_{N^{-1}\boldsymbol{A}_{22,n}}(\boldsymbol{U}) = \frac{N^{(n-1)q/2} |\boldsymbol{U}|^{\frac{1}{2}(n-1)-\frac{1}{2}(q+1)} \exp\left(-\frac{1}{2}N \operatorname{tr} \boldsymbol{U}\boldsymbol{\Delta}_{22}^{-1}\right)}{2^{(n-1)q/2}|\boldsymbol{\Delta}_{22}|^{(n-1)/2}\Gamma_q((n-1)/2)}, \tag{3.25}$$

$\boldsymbol{U} > \boldsymbol{0}$. Similarly, in (3.3), $\boldsymbol{B} \sim W_q(N-n, \boldsymbol{\Delta}_{22})$ so $N^{-1}\boldsymbol{B}$ has marginal density function

$$f_{N^{-1}\boldsymbol{B}}(\boldsymbol{U}) = \frac{N^{q(N-n)/2} |\boldsymbol{U}|^{\frac{1}{2}(N-n)-\frac{1}{2}(q+1)} \exp\left(-\frac{1}{2}N \operatorname{tr} \boldsymbol{U}\boldsymbol{\Delta}_{22}^{-1}\right)}{2^{q(N-n)/2}|\boldsymbol{\Delta}_{22}|^{(N-n)/2}\Gamma_q((N-n)/2)}, \tag{3.26}$$

$\boldsymbol{U} > \boldsymbol{0}$. To evaluate $f_{\hat{\boldsymbol{\Delta}}_{12}|\hat{\boldsymbol{\Delta}}_{22}}$, we apply Lemma 3.8 with $\boldsymbol{\Xi}_1 = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1} \equiv \hat{\boldsymbol{\Delta}}_{12}$, $\boldsymbol{\Xi}_2 = N^{-1}\boldsymbol{A}_{22,n}$, and $\boldsymbol{\Xi}_3 = N^{-1}\boldsymbol{B}$. Noting that $\boldsymbol{\Xi}_2 + \boldsymbol{\Xi}_3 = N^{-1}(\boldsymbol{A}_{22,n} + \boldsymbol{B}) \equiv \hat{\boldsymbol{\Delta}}_{22}$, it follows from (3.13) that we need to evaluate the integral

$$\int_{\boldsymbol{0}<\boldsymbol{U}<\boldsymbol{T}_{22}} f_{\boldsymbol{\Xi}_1|\boldsymbol{\Xi}_2=\boldsymbol{U}}(\boldsymbol{T}_{12}) f_{\boldsymbol{\Xi}_2}(\boldsymbol{U}) f_{\boldsymbol{\Xi}_3}(\boldsymbol{T}_{22}-\boldsymbol{U}) \mathrm{d}\boldsymbol{U}.$$

Introducing the notation $\boldsymbol{M}_1 = \frac{1}{2}N(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})$ and $\boldsymbol{M}_2 = \frac{1}{2}N\boldsymbol{\Delta}_{22}^{-1}$, and collecting terms in $\boldsymbol{U}$ from (3.24)–(3.26), we find that we are to evaluate

$$\int_{\boldsymbol{0}<\boldsymbol{U}<\boldsymbol{T}_{22}} |\boldsymbol{U}|^{\frac{1}{2}(n+p-1)-\frac{1}{2}(q+1)} |\boldsymbol{T}_{22}-\boldsymbol{U}|^{\frac{1}{2}(N-n)-\frac{1}{2}(q+1)} \exp(-\operatorname{tr} \boldsymbol{M}_1\boldsymbol{U}) \mathrm{d}\boldsymbol{U}.$$

Changing variables from $\boldsymbol{U}$ to $\boldsymbol{T}_{22}^{1/2}\boldsymbol{U}\boldsymbol{T}_{22}^{1/2}$ transforms this integral to

$$|\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)}\int_{\boldsymbol{0}<\boldsymbol{U}<\boldsymbol{I}_q}|\boldsymbol{U}|^{\frac{1}{2}(n+p-1)-\frac{1}{2}(q+1)}|\boldsymbol{I}_q-\boldsymbol{U}|^{\frac{1}{2}(N-n)-\frac{1}{2}(q+1)}\exp(-\operatorname{tr}\boldsymbol{T}_{22}^{1/2}\boldsymbol{M}_1\boldsymbol{T}_{22}^{1/2}\boldsymbol{U})\mathrm{d}\boldsymbol{U}$$
$$=B_q\left(\frac{1}{2}(n+p-1),\frac{1}{2}(N-n)\right)|\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)}\,{}_1F_1^{(q)}\left(\frac{1}{2}(n+p-1);\frac{1}{2}(N+p-1);-\boldsymbol{M}_1\boldsymbol{T}_{22}\right),\quad(3.27)$$

where the last equality follows from (3.15).

Combining and simplifying (3.24)–(3.27), we obtain

$$f_{\hat{\boldsymbol{\Delta}}_{12}|\hat{\boldsymbol{\Delta}}_{22}=\boldsymbol{T}_{22}}(\boldsymbol{T}_{12})=(2\pi)^{-pq/2}2^{-q(N-1)/2}N^{q(N+p-1)/2}\frac{B_q(\frac{1}{2}(n+p-1),\frac{1}{2}(N-n))}{\Gamma_q(\frac{1}{2}(n-1))\Gamma_q(\frac{1}{2}(N-n))}$$
$$\times|\boldsymbol{\Delta}_{11}|^{-q/2}|\boldsymbol{\Delta}_{22}|^{-(N-1)/2}\exp\left(-\frac{1}{2}\operatorname{tr}N\boldsymbol{\Delta}_{22}^{-1}\boldsymbol{T}_{22}\right)|\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)}$$
$$\times\,{}_1F_1^{(q)}\left(\frac{1}{2}(n+p-1);\frac{1}{2}(N+p-1);-\frac{1}{2}N(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}\right),\quad(3.28)$$

where $\boldsymbol{T}_{12}\in\mathbb{R}^{p\times q}$, $\boldsymbol{T}_{22}>\boldsymbol{0}$. By (3.14),

$$\frac{B_q(\frac{1}{2}(n+p-1),\frac{1}{2}(N-n))}{\Gamma_q(\frac{1}{2}(n-1))\Gamma_q(\frac{1}{2}(N-n))}=\frac{\Gamma_q(\frac{1}{2}(n+p-1))}{\Gamma_q(\frac{1}{2}(n-1))\Gamma_q(\frac{1}{2}(N+p-1))}.$$

Note that the matrix $\boldsymbol{M}_1$ is of rank $p$; therefore, its non-zero eigenvalues are the eigenvalues of $\frac{1}{2}N\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'$. It now follows from (3.18) that

$${}_1F_1^{(q)}\left(\frac{1}{2}(n+p-1);\frac{1}{2}(N+p-1);-\frac{1}{2}N(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}\right)$$
$$={}_1F_1^{(p)}\left(\frac{1}{2}(n+p-1);\frac{1}{2}(N+p-1);-\frac{1}{2}N\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\right).$$

Applying these last two results to (3.28), we obtain (3.22). □

**Corollary 3.10.** *Under the assumptions of Theorem 3.9, the density function of $\hat{\boldsymbol{\Sigma}}$ is*

$$f_{\hat{\boldsymbol{\Sigma}}}(\boldsymbol{T})=|\boldsymbol{T}_{22}|^{-p}f_{\hat{\boldsymbol{\Delta}}_{11}}(\boldsymbol{T}_{11}-\boldsymbol{T}_{12}\boldsymbol{T}_{22}^{-1}\boldsymbol{T}_{21})f_{\hat{\boldsymbol{\Delta}}_{22}}(\boldsymbol{T}_{22})f_{\hat{\boldsymbol{\Delta}}_{12}|\hat{\boldsymbol{\Delta}}_{22}=\boldsymbol{T}_{22}}(\boldsymbol{T}_{12}\boldsymbol{T}_{22}^{-1}),$$

*where* $\boldsymbol{T}=\begin{pmatrix}\boldsymbol{T}_{11}&\boldsymbol{T}_{12}\\\boldsymbol{T}_{21}&\boldsymbol{T}_{22}\end{pmatrix}>\boldsymbol{0}$.

**Proof.** We apply the transformation from $\hat{\boldsymbol{\Delta}}$ to $\hat{\boldsymbol{\Sigma}}$ given by (3.12). The Jacobian of this transformation is

$$J(\hat{\boldsymbol{\Delta}}_{11}\to\hat{\boldsymbol{\Sigma}}_{11})\cdot J(\hat{\boldsymbol{\Delta}}_{12}\to\hat{\boldsymbol{\Sigma}}_{12})\cdot J(\hat{\boldsymbol{\Delta}}_{22}\to\hat{\boldsymbol{\Sigma}}_{22})=1\cdot|\hat{\boldsymbol{\Sigma}}_{22}^{-1}|^{p}\cdot 1=|\hat{\boldsymbol{\Sigma}}_{22}|^{-p}.$$

Therefore, the density function of $\hat{\boldsymbol{\Sigma}}$ is

$$f_{\hat{\boldsymbol{\Sigma}}}(\boldsymbol{T})=f_{\hat{\boldsymbol{\Delta}}_{11},\hat{\boldsymbol{\Delta}}_{12},\hat{\boldsymbol{\Delta}}_{22}}(\boldsymbol{T}_{11}-\boldsymbol{T}_{12}\boldsymbol{T}_{22}^{-1}\boldsymbol{T}_{21},\boldsymbol{T}_{12}\boldsymbol{T}_{22}^{-1},\boldsymbol{T}_{22})|\boldsymbol{T}_{22}|^{-p},$$

which equals the stated formula. □

**Remark 3.11.** If the density function of $\hat{\boldsymbol{\Delta}}$ is to be integrated over subsets of the space of positive definite matrices, we recommend that the saddlepoint approximations of Butler and Wood [17] be applied to approximate the function ${}_1F_1^{(p)}$, as follows. Let $\boldsymbol{T}$ be a positive definite $p\times p$ matrix with eigenvalues $t_1,\ldots,t_p$. For $a<b$, define

$$\hat{s}_i=\begin{cases}\left[t_i-b+\left((t_i-b)^2+4at_i\right)^{1/2}\right]\Big/2t_i, & \text{if } t_i\neq 0\\ a/b, & \text{if } t_i=0\end{cases}$$

$i=1,\ldots,p$,

$$\widetilde{J}_{1,1}=\prod_{i=1}^{p}\prod_{j=1}^{p}\left[a(1-\hat{s}_i)(1-\hat{s}_j)+(b-a)\hat{s}_i\hat{s}_j\right],$$

and

$$\hat{J}_{1,1} = \prod_{i=1}^{p}\prod_{j=1}^{p}\left[\frac{\hat{s}_i\hat{s}_j}{a} + \frac{(1-\hat{s}_i)(1-\hat{s}_j)}{b-a}\right].$$

For $a, b-a > (p-1)/2$ the *raw Laplace approximation* to ${}_1F_1^{(p)}(a; b; \boldsymbol{T})$ is

$${}_1\widetilde{F}_1^{(p)}(a; b; \boldsymbol{T}) = \frac{2^{p/2}\pi^{p(p+1)/4}}{B_p(a, b-a)}\widetilde{J}_{1,1}^{-1/2}\prod_{i=1}^{p}\left[\hat{s}_i^a(1-\hat{s}_i)^{b-a}\exp(t_i\hat{s}_i)\right], \tag{3.29}$$

and the *calibrated Laplace approximation* is

$${}_1\hat{F}_1^{(p)}(a; b; \boldsymbol{T}) = b^{pb-p(p+1)/4}\hat{J}_{1,1}^{-1/2}\prod_{i=1}^{p}\left[\left(\frac{\hat{s}_i}{a}\right)^a\left(\frac{1-\hat{s}_i}{b-a}\right)^{b-a}e^{t_i\hat{s}_i}\right], \tag{3.30}$$

both of which satisfy (3.16) and (3.17).

## 4. Tests of hypotheses about $\mu$ and $\Sigma$

### 4.1. Testing that $\boldsymbol{\Sigma}$ equals a given matrix

Consider the problem of testing $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ against $H_a : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$, where $\boldsymbol{\Sigma}_0$ is a given positive definite matrix, on the basis of the monotone sample (1.1). Hao and Krishnamoorthy [18] used an invariance argument to show that, without loss of generality, we may assume $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_{p+q}$, and they proved that the likelihood ratio statistic for testing $H_0$ against $H_a$ is

$$\begin{aligned}\lambda_1 = {} & (e/N)^{Nq/2}|\boldsymbol{A}_{22,N}|^{N/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{22,N}\right)\\ & \times (e/n)^{np/2}|\boldsymbol{A}_{11\cdot 2,n}|^{n/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{11\cdot 2,n}\right)\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\right).\end{aligned} \tag{4.1}$$

In the case of a complete sample, it is well-known that the likelihood ratio statistic for this problem is not unbiased, so the same can be expected to hold for $\lambda_1$. Hao and Krishnamoorthy [18] then modified $\lambda_1$ in the usual way, replacing sample sizes by degrees of freedom to obtain

$$\begin{aligned}\lambda_2 = {} & (e/(N-1))^{(N-1)q/2}|\boldsymbol{A}_{22,N}|^{(N-1)/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{22,N}\right)\\ & \times (e/(n-q-1))^{(n-q-1)p/2}|\boldsymbol{A}_{11\cdot 2,n}|^{(n-q-1)/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{11\cdot 2,n}\right)\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\right),\end{aligned} \tag{4.2}$$

and they derived an approximation to the asymptotic null distribution of this statistic. We shall prove that a sufficient condition for $\lambda_2$ to be unbiased is that, under $H_a$, $|\boldsymbol{\Sigma}_{11}| \leq 1$. Since $\lambda_2$ might not always be unbiased, we propose a new statistic,

$$\begin{aligned}\lambda_3 = {} & (e/(N-1))^{(N-1)q/2}|\boldsymbol{A}_{22,N}|^{(N-1)/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{22,N}\right)\\ & \times (e/(n-q-1))^{(n-q-1)p/2}|\boldsymbol{A}_{11\cdot 2,n}|^{(n-q-1)/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{11\cdot 2,n}\right)\\ & \times (e/q)^{qp/2}|\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}|^{q/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\right),\end{aligned} \tag{4.3}$$

and establish that it is always unbiased. The crucial difference between $\lambda_2$ and $\lambda_3$ is that the term $|\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}|^{q/2}$ in (4.3) causes certain integrals to be invariant under some matrix transformations, and those invariance properties cause $\lambda_3$ to be unbiased.

We now calculate the non-null moments of $\lambda_3$, and thereby identify its exact null distribution, derive approximations to that distribution, and establish unbiasedness. In the next result, we denote by $e_{p,q,n,N}$ the constant in (4.3).

**Theorem 4.1.** *For $h = 0, 1, 2, \ldots$ the h-th non-null moment of $\lambda_3$ is*

$$E(\lambda_3^h) = e_{p,q,n,N}^h \, 2^{((N-1)q+(n-1)p)h/2} \frac{\Gamma_q((N-1)(1+h)/2)}{\Gamma_q((N-1)/2)} \frac{\Gamma_p((n-q-1)(1+h)/2)}{\Gamma_p((n-q-1)/2)} \frac{\Gamma_p(q(1+h)/2)}{\Gamma_p(q/2)}$$
$$\times |\boldsymbol{\Sigma}_{22}|^{(N-1)h/2}|\boldsymbol{I}_q + h\boldsymbol{\Sigma}_{22}|^{-(N-1)(1+h)/2}|\boldsymbol{\Sigma}_{11\cdot 2}|^{(n-q-1)h/2}|\boldsymbol{I}_p + h\boldsymbol{\Sigma}_{11\cdot 2}|^{-(n-q-1)(1+h)/2}$$
$$\times |\boldsymbol{\Sigma}_{11}|^{qh/2}|\boldsymbol{I}_p + h\boldsymbol{\Sigma}_{11}|^{-q(1+h)/2}. \quad (4.4)$$

**Proof.** Under $H_a$, we apply invariance arguments to allow us to assume, without loss of generality, that $\boldsymbol{\Sigma}$ is diagonal [18, p. 66]. Then, by Proposition 3.2, $\boldsymbol{A}_{22,N}$, $\boldsymbol{A}_{11\cdot 2,n}$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$ are mutually independent, and $\boldsymbol{A}_{22,N} \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{A}_{11\cdot 2,n} \sim W_p(n-q-1, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21} \sim W_p(q, \boldsymbol{\Sigma}_{11})$. Therefore

$$\lambda_3 \overset{\mathcal{L}}{=} e_{p,q,n,N}|\boldsymbol{W}_1|^{(N-1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{W}_1\right)|\boldsymbol{W}_2|^{(n-q-1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{W}_2\right)|\boldsymbol{W}_3|^{q/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{W}_3\right), \quad (4.5)$$

where $\boldsymbol{W}_1 \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{W}_2 \sim W_p(n-q-1, \boldsymbol{\Sigma}_{11})$, $\boldsymbol{W}_3 \sim W_p(q, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, and $\boldsymbol{W}_3$ are mutually independent.

For $\boldsymbol{W} \sim W_d(a, \boldsymbol{\Sigma})$, it follows from (2.1) that

$$E\left(|\boldsymbol{W}|^{\alpha/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{W}\right)\right)^h = 2^{\alpha dh/2}\frac{\Gamma_d((\alpha h + a)/2)}{\Gamma_d(a/2)}|\boldsymbol{\Sigma}|^{ah/2}|\boldsymbol{I}_d + h\boldsymbol{\Sigma}|^{-(\alpha h+a)/2}, \quad (4.6)$$

$\mathrm{Re}(\alpha h + a) > p - 1$. Applying this formula to each Wishart matrix in (4.5) and simplifying the resulting expression, we obtain (4.4). □

By expressing each determinant in (4.4) as a product of its eigenvalues, we thereby deduce a stochastic representation for $\lambda_3$ as a product of independent random variables. We state this result explicitly in the null case, bearing in mind that we have then assumed $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_{p+q}$.

**Corollary 4.2.** *Under the null hypothesis $H_0 : \boldsymbol{\Sigma} = \boldsymbol{I}_{p+q}$, we have*

$$\lambda_3 \overset{\mathcal{L}}{=} e_{p,q,n,N}\, e^{-Q_0/2}\prod_{j=1}^{q} Q_{j,1}^{(N-1)/2}e^{-Q_{j,1}/2}\cdot\prod_{j=1}^{p} Q_{j,2}^{(n-q-1)/2}e^{-Q_{j,2}/2}Q_{j,3}^{q/2}e^{-Q_{j,3}/2}, \quad (4.7)$$

*where $Q$ and all $Q_{j,k}$ are mutually independent, $Q \sim \chi^2_{\frac{1}{2}q(q-1)+p(p-1)}$; $Q_{j,1} \sim \chi^2_{N-j}$, $j = 1, \ldots, q$; $Q_{j,2} \sim \chi^2_{n-q-j}$, and $Q_{j,3} \sim \chi^2_{q-j+1}$, $j = 1, \ldots, p$.*

**Proof.** Substituting $\boldsymbol{\Sigma} = \boldsymbol{I}_{p+q}$ in (4.4), we obtain the null moments of $\lambda_3$, *viz.*,

$$E(\lambda_3^h) = e_{p,q,n,N}^h\, 2^{((N-1)q+(n-1)p)h/2}\,(1+h)^{-((N-1)q+(n-1)p)(1+h)/2}$$
$$\times \frac{\Gamma_q((N-1)(1+h)/2)}{\Gamma_q((N-1)/2)}\frac{\Gamma_p((n-q-1)(1+h)/2)}{\Gamma_p((n-q-1)/2)}\frac{\Gamma_p(q(1+h)/2)}{\Gamma_p(q/2)}.$$

Substituting $\boldsymbol{\Sigma} = \boldsymbol{I}_d$ at (4.6), the right-hand side of that formula reduces to

$$2^{adh/2}\frac{\Gamma_d(a(1+h)/2)}{\Gamma_d(a/2)}(1+h)^{-ad(1+h)/2}$$
$$= (1+h)^{-d(d-1)/4}\prod_{j=1}^{d}\left[2^{ah/2}\frac{\Gamma(\frac{1}{2}(a-j+1)+\frac{1}{2}ah)}{\Gamma(\frac{1}{2}(a-j+1))}(1+h)^{-(a-j+1+ah)/2}\right].$$

On recognizing that each of the $d+1$ terms in this latter product is the $h$-th moment of a function of a chi-squared random variable, we deduce that if $\boldsymbol{W} \sim W_d(a, \boldsymbol{I}_d)$ then

$$|\boldsymbol{W}|^{a/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{W}\right) \overset{\mathcal{L}}{=} e^{-Q_0/2}\prod_{j=1}^{d} Q_j^{a/2}\,e^{-Q_j/2},$$

where $Q_0, \ldots, Q_d$ are independent chi-squared variables, $Q_0 \sim \chi^2_{d(d-1)/2}$, and $Q_j \sim \chi^2_{a-j+1}$ for $j = 1, \ldots, d$. Applying this result to each matrix in (4.5), we obtain

$$\lambda_3 \overset{\mathcal{L}}{=} e_{p,q,n,N}\, e^{-(Q_{0,1}+Q_{0,2}+Q_{0,3})/2}\prod_{j=1}^{q} Q_{j,1}^{(N-1)/2}e^{-Q_{j,1}/2}\cdot\prod_{j=1}^{p} Q_{j,2}^{(n-q-1)/2}e^{-Q_{j,2}/2}\cdot\prod_{j=1}^{p} Q_{j,3}^{q/2}e^{-Q_{j,3}/2},$$

where the $Q_{j,k}$ are independent, $Q_{0,1} \sim \chi^2_{q(q-1)/2}$, $Q_{j,1} \sim \chi^2_{N-j}, j = 1, \ldots, q$; $Q_{0,2} \sim \chi^2_{p(p-1)/2}$, $Q_{j,2} \sim \chi^2_{n-q-j}, j = 1, \ldots, p$; and $Q_{0,3} \sim \chi^2_{p(p-1)/2}$, $Q_{j,3} \sim \chi^2_{q-j+1}, j = 1, \ldots, p$. Letting $Q = Q_{0,1} + Q_{0,2} + Q_{0,3}$, so that $Q \sim \chi^2_{\frac{1}{2}q(q-1)+p(p-1)}$, we obtain (4.7). □

A complete treatment of the exact distribution of $\lambda_3$ would take us too far afield, so we restrict our attention to its asymptotic distribution and approximations thereof. With regard to the null distribution of $\lambda_3$, we apply the results of [12, p. 359] (see also [18, p. 68]) to each of the three terms in the representation of $\lambda_3$ as a product of independent random entities in (4.3) or (4.5). Under $H_0$, the asymptotic distribution of $\lambda_3$ for large $n$ and $N$ is given by

$$-2 \ln \lambda_3 \approx \sum_{j=1}^{3} \rho_j^{-1} \chi^2_{d_j}, \tag{4.8}$$

where $\chi^2_{d_j}, j = 1, 2, 3$, are independent, $d_1 = q(q+1)/2$, $d_2 = d_3 = p(p+1)/2$, and

$$\rho_1 = 1 - \frac{2q^2 + 3q - 1}{6(N-1)(q+1)}, \qquad \rho_2 = 1 - \frac{2p^2 + 3p - 1}{6(n-q-1)(p+1)}, \qquad \rho_3 = 1 - \frac{2p^2 + 3p - 1}{6q(p+1)}.$$

Let $\rho_{(1)}$ and $\rho_{(3)}$ denote the smallest and largest of $\rho_1, \rho_2, \rho_3$, respectively. On applying to the right-hand side of (4.8) the results of Kotz et al. [19, Section 5], we obtain the asymptotic distribution function of $-2 \ln \lambda_3$ in the form

$$P(-2 \ln \lambda_3 \le t) \simeq P(\chi^2_{d_1+d_2+d_3} \le t/\beta_1),$$

$t > 0$, where $\beta_1 = (\rho_{(1)}^{-1} + \rho_{(3)}^{-1})/2$. This approximation is the first term in the Laguerre series expansions of [19], and additional terms in our approximation may be obtained accordingly from their series. Alternatively, by applying the results of [19, Section 6], we also obtain

$$P(-2 \ln \lambda_3 \le t) \simeq c_0(\beta_2) P(\chi^2_{d_1+d_2+d_3} \le t/\beta_2),$$

where $\beta_2 = (d_1 + d_2 + d_3)/(d_1\rho_1 + d_2\rho_2 + d_3\rho_3)$ and $c_0(\beta_2) = \prod_{j=1}^{3} (\beta_2 \rho_j)^{d_j/2}$.

Saddlepoint approximations to the distribution of (4.8) are noteworthy for they generally are superior to standard asymptotic approximations in the case of small sample sizes. Let

$$K(\zeta) = -\frac{1}{2} \sum_{j=1}^{3} d_j \ln(1 - 2\rho_j^{-1} \zeta)$$

denote the cumulant-generating function of the right-hand side of (4.8). Applying the results of Kuonen [20, Eq. (3)], we obtain

$$P(-2 \ln \lambda_3 \le t) \simeq \Phi\left(w + w^{-1} \ln(v w^{-1})\right),$$

$t > 0$, where $\Phi$ denotes the standard normal distribution function, $\hat{\zeta}$ is the unique solution of the equation $K'(\zeta) = t$, $w = \operatorname{sign}(\hat{\zeta})[2\{\hat{\zeta} t - K(\hat{\zeta})\}]^{1/2}$, and $v = \hat{\zeta}\, [K''(\hat{\zeta})]^{1/2}$.

We remark also that although the above results constitute a saddlepoint approximation only to the asymptotic distribution of $\lambda_3$, the methods of Booth et al. [21] may be applied to obtain a saddlepoint approximation to the exact distribution of $\lambda_3$.

We consider next the unbiasedness of $\lambda_2$ and $\lambda_3$. The proof of the following result follows the argument of Sugiura and Nagao [22] (see [12, p. 367]).

**Theorem 4.3.** *The statistic $\lambda_3$ is unbiased. Further, if $|\boldsymbol{\Sigma}_{11}| \le 1$ then $\lambda_2$ is unbiased.*

**Proof.** As before, without loss of generality, we assume under $H_a$ that $\boldsymbol{\Sigma}$ is diagonal. By (4.5), a critical region of size $\alpha$ for $\lambda_3$ is the set $\mathfrak{C}_3 = \{(\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3) : \lambda_3/e_{p,q,n,N} \le k_\alpha\}$, where $\boldsymbol{W}_1 \sim \mathrm{W}_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{W}_2 \sim \mathrm{W}_p(n-q-1, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{W}_3 \sim \mathrm{W}_p(q, \boldsymbol{\Sigma}_{11})$ are mutually independent, and the constant $k_\alpha$ is such that $P(\lambda_3 \in \mathfrak{C}_3 | H_0) = \alpha$. Denote by $c_q(N-1, \boldsymbol{\Sigma}_{22})$, $c_p(n-q-1, \boldsymbol{\Sigma}_{11})$, and $c_p(q, \boldsymbol{\Sigma}_{11})$ the normalizing constants in the Wishart density functions of $\boldsymbol{W}_1, \boldsymbol{W}_2$, and $\boldsymbol{W}_3$, respectively. Again applying (4.5), we obtain

$$\begin{aligned} P(\lambda_3 \in \mathfrak{C}_3 | H_a) = \int_{(\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3) \in \mathfrak{C}_3} & c_q(N-1, \boldsymbol{\Sigma}_{22}) |\boldsymbol{W}_1|^{\frac{1}{2}(N-1) - \frac{1}{2}(q+1)} \exp\left(-\frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{W}_1\right) \\ & \times c_p(n-q-1, \boldsymbol{\Sigma}_{11}) |\boldsymbol{W}_2|^{\frac{1}{2}(n-q-1) - \frac{1}{2}(p+1)} \exp\left(-\frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{W}_2\right) \\ & \times c_p(q, \boldsymbol{\Sigma}_{11}) |\boldsymbol{W}_3|^{\frac{1}{2}q - \frac{1}{2}(p+1)} \exp\left(-\frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{W}_3\right) \prod_{j=1}^{3} \mathrm{d}\boldsymbol{W}_j. \end{aligned}$$

Making the transformation

$$(\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3) = (\boldsymbol{\Sigma}_{22}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{22}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2}\widetilde{\boldsymbol{W}}_2\boldsymbol{\Sigma}_{11}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2}\widetilde{\boldsymbol{W}}_3\boldsymbol{\Sigma}_{11}^{1/2}) \tag{4.9}$$

in this integral, we obtain

$$\begin{aligned} P(\lambda_3 \in \mathfrak{C}_3|H_a) = \int_{(\widetilde{\boldsymbol{W}}_1,\widetilde{\boldsymbol{W}}_2,\widetilde{\boldsymbol{W}}_3)\in\mathfrak{C}_3^*} & c_q(N-1, \boldsymbol{I}_q)|\widetilde{\boldsymbol{W}}_1|^{\frac{1}{2}(N-1)-\frac{1}{2}(q+1)} \exp\left(-\frac{1}{2}\operatorname{tr}\widetilde{\boldsymbol{W}}_1\right) \\ & \times c_p(n-q-1, \boldsymbol{I}_p)|\widetilde{\boldsymbol{W}}_2|^{\frac{1}{2}(n-q-1)-\frac{1}{2}(p+1)} \exp\left(-\frac{1}{2}\operatorname{tr}\widetilde{\boldsymbol{W}}_2\right) \\ & \times c_p(q, \boldsymbol{I}_p)|\widetilde{\boldsymbol{W}}_3|^{\frac{1}{2}q-\frac{1}{2}(p+1)} \exp\left(-\frac{1}{2}\operatorname{tr}\widetilde{\boldsymbol{W}}_3\right)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j, \end{aligned}$$

where

$$\mathfrak{C}_3^* = \{(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) : (\boldsymbol{\Sigma}_{22}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{22}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{11}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{11}^{1/2}) \in \mathfrak{C}_3\}. \tag{4.10}$$

Under $H_0$, $\mathfrak{C}_3^* = \mathfrak{C}_3$; denoting the null joint density function of $(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)$ by $f_0$, we have

$$\begin{aligned} P(\lambda_3 \in \mathfrak{C}_3|H_a) - P(\lambda_3 \in \mathfrak{C}_3|H_0) &= \left\{\int_{\mathfrak{C}_3^*} - \int_{\mathfrak{C}_3}\right\} f_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j \\ &= \left\{\int_{\mathfrak{C}_3^*\setminus\mathfrak{C}_3} - \int_{\mathfrak{C}_3\setminus\mathfrak{C}_3^*}\right\} f_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j. \end{aligned}$$

For $(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) \in \mathfrak{C}_3 \setminus \mathfrak{C}_3^* \subset \mathfrak{C}_3$,

$$|\widetilde{\boldsymbol{W}}_1|^{\frac{1}{2}(N-1)} \exp\left(-\frac{1}{2}\operatorname{tr}\widetilde{\boldsymbol{W}}_1\right)|\widetilde{\boldsymbol{W}}_2|^{\frac{1}{2}(n-q-1)} \exp\left(-\frac{1}{2}\operatorname{tr}\widetilde{\boldsymbol{W}}_2\right)|\widetilde{\boldsymbol{W}}_3|^{q/2} \exp\left(-\frac{1}{2}\operatorname{tr}\widetilde{\boldsymbol{W}}_3\right) \le k_\alpha,$$

and hence $f_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) \le k_\alpha\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)$, where

$$\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) = c_q(N-1, \boldsymbol{I}_q)\, c_p(n-q-1, \boldsymbol{I}_p)\, c_p(q, \boldsymbol{I}_p)|\widetilde{\boldsymbol{W}}_1|^{-\frac{1}{2}(q+1)}|\widetilde{\boldsymbol{W}}_2|^{-\frac{1}{2}(p+1)}|\widetilde{\boldsymbol{W}}_3|^{-\frac{1}{2}(p+1)},$$

$\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3 > \boldsymbol{0}$. For $(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) \in \mathfrak{C}_3^* \setminus \mathfrak{C}_3 \subset \mathfrak{C}_3^*$,

$$f_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) > k_\alpha\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3);$$

therefore

$$\begin{aligned} P(\lambda_3 \in \mathfrak{C}_3|H_a) - P(\lambda_3 \in \mathfrak{C}_3|H_0) &> k_\alpha\left\{\int_{\mathfrak{C}_3^*\setminus\mathfrak{C}_3} - \int_{\mathfrak{C}_3\setminus\mathfrak{C}_3^*}\right\}\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j \\ &= k_\alpha\left\{\int_{\mathfrak{C}_3^*} - \int_{\mathfrak{C}_3}\right\}\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j. \end{aligned}$$

Now substitute $(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) = (\boldsymbol{\Sigma}_{22}^{-1/2}W_1\boldsymbol{\Sigma}_{22}^{-1/2}, \boldsymbol{\Sigma}_{11}^{-1/2}W_2\boldsymbol{\Sigma}_{11}^{-1/2}, \boldsymbol{\Sigma}_{11}^{-1/2}W_3\boldsymbol{\Sigma}_{11}^{-1/2})$. Since the measure $|\widetilde{\boldsymbol{W}}_1|^{-\frac{1}{2}(q+1)}$ $|\widetilde{\boldsymbol{W}}_2|^{-\frac{1}{2}(p+1)}|\widetilde{\boldsymbol{W}}_3|^{-\frac{1}{2}(p+1)}\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j$ is invariant under this transformation, we obtain

$$\int_{\mathfrak{C}_3^*}\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j = \int_{\mathfrak{C}_3}\widetilde{f}_0(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3)\prod_{j=1}^{3}\mathrm{d}\widetilde{\boldsymbol{W}}_j.$$

Therefore $P(\lambda_3 \in \mathfrak{C}_3|H_a) - P(\lambda_3 \in \mathfrak{C}_3|H_0) > 0$, which proves that $\lambda_3$ is unbiased.

In the case of $\lambda_2$, let $\mathfrak{C}_2 = \{(\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3) : \lambda_2/e_{2,p,q,n,N} \le k_\alpha\}$ denote the critical region of size $\alpha$ and $k_\alpha$ be the corresponding percentage point, where $e_{2,p,q,n,N}$ denotes the constant term in (4.2). We again apply the transformation (4.9) and, similar to (4.10), define $\mathfrak{C}_2^* = \{(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) : (\boldsymbol{\Sigma}_{22}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{22}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{11}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2}\widetilde{\boldsymbol{W}}_1\boldsymbol{\Sigma}_{11}^{1/2}) \in \mathfrak{C}_2\}$. By an argument analogous to that given for $\lambda_3$, we obtain

$$P(\lambda_2 \in \mathfrak{C}_2|H_a) - P(\lambda_2 \in \mathfrak{C}_2|H_0) > k_\alpha\left(|\boldsymbol{\Sigma}_{11}|^{-qp/2} - 1\right)\left\{\int_{\mathfrak{C}_2^*} - \int_{\mathfrak{C}_2}\right\}|W_3|^{-\frac{1}{2}(p+1)}\widetilde{f}_0(\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_3)\prod_{j=1}^{3}\mathrm{d}\boldsymbol{W}_j.$$

For $|\boldsymbol{\Sigma}_{11}|^{-qp/2} - 1 \ge 0$, or equivalently $|\boldsymbol{\Sigma}_{11}| \le 1$, it follows that $\lambda_2$ is unbiased. □

Next, we show that the statistic $\lambda_1$ in (4.1) is not unbiased for all $n$ and $N$. Here, the proof follows the classical approach of Das Gupta [23] (see also [12, p. 357]).

**Proposition 4.4.** *For testing $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ against $H_a : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$, the likelihood ratio test statistic $\lambda_1$ in (4.1) is not unbiased.*

**Proof.** As before, we shall assume without loss of generality that $\boldsymbol{\Sigma}$ is diagonal, say, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{1,1}, \ldots, \sigma_{p+q,p+q})$. By Proposition 3.2, the matrices $\boldsymbol{A}_{22,N}$, $\boldsymbol{A}_{11\cdot 2,n}$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$ are mutually independent with $\boldsymbol{A}_{22,N} \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{A}_{11\cdot 2,n} \sim W_p(n-q-1, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21} \sim W_p(q, \boldsymbol{\Sigma}_{11})$. By (4.1),

$$
\begin{aligned}
(e/N)^{-Nq/2}(e/n)^{-np/2}\lambda_1 &= |\boldsymbol{A}_{22,N}|^{N/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{A}_{22,N}\right)|\boldsymbol{A}_{11\cdot 2,n}|^{n/2}\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{A}_{11\cdot 2,n}\right)\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\right) \\
&= \left[\frac{|\boldsymbol{A}_{22,N}|}{\prod_{j=p+1}^{p+q}(\boldsymbol{A}_{22,N})_{jj}}\right]^{N/2}\prod_{j=p+1}^{p+q}(\boldsymbol{A}_{22,N})_{jj}^{N/2}\exp\left(-\frac{1}{2}(\boldsymbol{A}_{22,N})_{jj}\right) \\
&\quad\times\left[\frac{|\boldsymbol{A}_{11\cdot 2,n}|}{\prod_{j=1}^{p}(\boldsymbol{A}_{11\cdot 2,n})_{jj}}\right]^{N/2}\prod_{j=1}^{p}(\boldsymbol{A}_{11\cdot 2,n})_{jj}^{N/2}\exp\left(-\frac{1}{2}(\boldsymbol{A}_{11\cdot 2,n})_{jj}\right)\exp\left(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\right).
\end{aligned}
$$

The rest of the proof now proceeds as in the classical case. The random variables $(\boldsymbol{A}_{22,N})_{jj}$, $j = p+1, \ldots, p+q$, and $|\boldsymbol{A}_{22,N}|/\prod_{j=p+1}^{p+q}(\boldsymbol{A}_{22,N})_{jj}$ are mutually independent. Moreover, the distribution of $|\boldsymbol{A}_{22,N}|/\prod_{j=p+1}^{p+q}(\boldsymbol{A}_{22,N})_{jj}$ does not depend on $\boldsymbol{\Sigma}_{22}$ and $(\boldsymbol{A}_{22,N})_{jj}/\sigma_{j,j} \sim \chi^2_{N-1}$. By [12, p. 356, Lemma 8.4.3], there exists $\sigma^*_{p+q} \in (1, N/(N-1))$ such that, for any $c > 0$,

$$
\begin{aligned}
&P\left((\boldsymbol{A}_{22,N})_{p+q,p+q}^{N/2}\exp\left(-\frac{1}{2}(\boldsymbol{A}_{22,N})_{p+q,p+q}\right) \geq k|\sigma_{p+q,p+q} = 1\right) \\
&\quad< P\left((\boldsymbol{A}_{22,N})_{p+q,p+q}^{N/2}\exp\left(-\frac{1}{2}(\boldsymbol{A}_{22,N})_{p+q,p+q}\right) \geq c|\sigma_{p+q,p+q} = \sigma^*_{p+q}\right).
\end{aligned}
$$

The conclusion is obtained when we evaluate $P(\lambda_1 \geq c)$ by conditioning on the variables $\{(\boldsymbol{A}_{22,N})_{jj}, j = p+1, \ldots, p+q-1\}$, $|\boldsymbol{A}_{22,N}|/\prod_{j=p+1}^{p+q}(\boldsymbol{A}_{22,N})_{jj}$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$. □

As in the classical case, we can obtain a result which is stronger than the unbiasedness property of $\lambda_3$ [12, p. 358]; however, we also note that it does not provide the unbiasedness property of $\lambda_2$ which was deduced in Theorem 4.3. The proof of the following result is similar to the classical case.

**Theorem 4.5.** *For $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{1,1}, \ldots, \sigma_{p+q,p+q})$, the power function of the modified likelihood ratio statistic $\lambda_3$ increases monotonically with $|\sigma_{j,j} - 1|$, $1 \leq j \leq p+q$.*

**Proof.** By [12, p. 357, Corollary 8.4.4],

$$
P\left((\boldsymbol{A}_{22,N})_{p+q,p+q}^{(N-1)/2}\exp\left(-\frac{1}{2}(\boldsymbol{A}_{22,N})_{p+q,p+q}\right) \leq k|\sigma_{p+q,p+q}\right)
$$

increases monotonically as $|\sigma_{p+q,p+q} - 1|$ increases, and an analogous result holds for $\boldsymbol{A}_{11\cdot 2,n}$ and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$. The conclusion is now obtained by a conditioning argument similar to the one applied in Proposition 4.4. □

*4.2. Testing that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ equal a given vector and matrix*

On the basis of the monotone sample (1.1), consider the problem of testing $H_0 : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ against $H_a : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) \neq (\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are completely specified. Hao and Krishnamoorthy [18, Eq. (4.1)] showed that the likelihood ratio test statistic is

$$
\lambda_4 = \lambda_1 \exp\left(-\frac{1}{2}(n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}} + N\bar{\boldsymbol{Y}}'\bar{\boldsymbol{Y}})\right), \tag{4.11}
$$

where $\lambda_1$ is the test statistic in (4.1). By invariance arguments we may assume, without loss of generality, that $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = (\mathbf{0}, \boldsymbol{I}_{p+q})$ and that $\boldsymbol{\Sigma}$ is diagonal under $H_a$. Substituting (4.1) into (4.11), we obtain

$$\lambda_4 = (e/N)^{Nq/2}|\boldsymbol{A}_{22,N}|^{N/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{22,N}\right)(e/n)^{np/2}|\boldsymbol{A}_{11\cdot 2,n}|^{n/2}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{11\cdot 2,n}\right)$$
$$\times \exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\right)\exp\left(-\frac{1}{2}(n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}} + N\bar{\boldsymbol{Y}}'\bar{\boldsymbol{Y}})\right).$$

By (3.3) and Proposition 3.2 we have that $\boldsymbol{A}_{22,N}$, $\boldsymbol{A}_{11\cdot 2,n}$, $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$, $\bar{\boldsymbol{X}}$, and $\bar{\boldsymbol{Y}}$ are mutually independent under $H_0$ and $\boldsymbol{A}_{22,N} \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{A}_{11\cdot 2,n} \sim W_p(n-q-1, \boldsymbol{\Sigma}_{11})$, $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21} \sim W_p(q, \boldsymbol{\Sigma}_{11})$, $\bar{\boldsymbol{X}} \sim N_p(\boldsymbol{\mu}_1, n^{-1}\boldsymbol{\Sigma}_{11})$, and $\bar{\boldsymbol{Y}} \sim N_q(\boldsymbol{\mu}_2, N^{-1}\boldsymbol{\Sigma}_{22})$. In particular, the individual terms on the right-hand side of (4.11) are mutually independent.

To identify the exact null distribution of $\lambda_4$ and investigate its unbiasedness properties, we proceed as in the case of $\lambda_3$. We omit the proof of the following result since the details are similar to those in the previous subsection.

**Theorem 4.6.** *The likelihood ratio statistic $\lambda_4$ for testing $H_0 : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\mathbf{0}, \boldsymbol{I}_{p+q})$ against $H_a : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) \neq (\mathbf{0}, \boldsymbol{I}_{p+q})$ is unbiased. For $h = 0, 1, 2, \ldots$ the h-th non-null moment of $\lambda_4$ is*

$$E(\lambda_4^h) = \left(\frac{2e}{N}\right)^{Nqh/2}\left(\frac{2e}{n}\right)^{nph/2}\frac{\Gamma_q((Nh+N-1)/2)}{\Gamma_q((N-1)/2)}\frac{\Gamma_p((nh+n-q-1)/2)}{\Gamma_p((n-q-1)/2)}$$
$$\times |\boldsymbol{\Sigma}_{22}|^{Nh/2}|\boldsymbol{I}_q + h\boldsymbol{\Sigma}_{22}|^{-(Nh+N-1)/2}|\boldsymbol{\Sigma}_{11}|^{nh/2}|\boldsymbol{I}_p + h\boldsymbol{\Sigma}_{11}|^{-(nh+n-1)/2}$$
$$\times \exp\left(-(n\boldsymbol{\mu}_1'\boldsymbol{\mu}_1 + N\boldsymbol{\mu}_2'\boldsymbol{\mu}_2)h\right)|\boldsymbol{I}_p + 2h\boldsymbol{\Sigma}_{11}|^{-1/2}|\boldsymbol{I}_q + 2h\boldsymbol{\Sigma}_{22}|^{-1/2}$$
$$\times \exp\left(2h^2[n\boldsymbol{\mu}_1'(\boldsymbol{I}_p + 2h\boldsymbol{\Sigma}_{11})^{-1}\boldsymbol{\mu}_1 + N\boldsymbol{\mu}_2'(\boldsymbol{I}_q + 2h\boldsymbol{\Sigma}_{22})^{-1}\boldsymbol{\mu}_2]\right) \tag{4.12}$$

*and, under $H_0$,*

$$\lambda_4 \stackrel{\mathcal{L}}{=} (2e/N)^{Nq/2}(2e/n)^{np/2}e^{-(Q_1+2Q_2)/2}\left(\prod_{j=1}^{p} Q_{j,1}^{n/2}\, e^{-Q_{j,1}/2}\right)\left(\prod_{j=1}^{q} Q_{j,2}^{N/2}\, e^{-Q_{j,2}/2}\right), \tag{4.13}$$

*where $Q_1 \sim \chi^2_{\frac{1}{2}p(p-1)+\frac{1}{2}q(q-1)+pq}$; $Q_2 \sim \chi^2_{p+q}$; $Q_{j,1} \sim \chi^2_{n-q-j}$, $1 \le j \le p$; $Q_{j,2} \sim \chi^2_{N-j}$; and all such $\chi^2$ variables are mutually independent.*

We remark that, in the non-null case, the distribution of $\lambda_4$ may also be obtained from (4.12); the final result is similar to (4.13) and involves noncentral chi-square random variables.

### *4.3. The sphericity test*

Consider the problem of testing sphericity, in which the null hypothesis is $H_0 : \boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_{p+q}$ and the alternative hypothesis is $H_a : \boldsymbol{\Sigma} \neq \sigma^2\boldsymbol{I}_{p+q}$, where $\sigma^2 > 0$ is unspecified. Bhargava [6, Section 6] derived the likelihood ratio test statistic for a problem more general than the sphericity test and obtained the null distribution of a modified form of that statistic in terms of independent chi-squared random variables. We shall treat the sphericity problem in a form closer to that of the classical approach ([13, p. 431], [12, p. 433]), deriving its moments and a stochastic representation for its null distribution.

First, we derive the likelihood ratio criterion. Under $H_0$, it is simple to show that the maximum likelihood estimators of $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$ and $\sigma^2$ are, respectively, $\hat{\boldsymbol{\mu}}_{10} = \bar{\boldsymbol{X}}$, $\hat{\boldsymbol{\mu}}_{20} = \bar{\boldsymbol{Y}}$, and

$$\hat{\sigma}_0^2 = \frac{1}{np+Nq}\left[\sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}})'(\boldsymbol{X}_j - \bar{\boldsymbol{X}}) + \sum_{j=1}^{N}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})'(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})\right]$$
$$= \frac{1}{np+Nq}\left[\operatorname{tr}\boldsymbol{A}_{11} + \operatorname{tr}\boldsymbol{A}_{22,N}\right].$$

Under $H_a$, the maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given in (2.4) and (3.1), respectively. By a straightforward calculation, we deduce that the likelihood ratio criterion for testing $H_0$ against $H_a$ is

$$\lambda_5 = \frac{|n^{-1}\boldsymbol{A}_{11\cdot 2,n}|^{n/2}|N^{-1}\boldsymbol{A}_{22,N}|^{N/2}}{\left((np+Nq)^{-1}(\operatorname{tr}\boldsymbol{A}_{11} + \operatorname{tr}\boldsymbol{A}_{22,N})\right)^{(np+Nq)/2}}. \tag{4.14}$$

For the classical case [13, p. 433], it is well-known that the likelihood ratio statistic is the quotient of an arithmetic and a geometric mean, and that result leads to an immediate proof that the statistic is no larger than 1. Generalizing that result, we now apply an arithmetic–geometric mean inequality to prove directly that $\lambda_5 \le 1$. Let $\mathcal{A}_1$ and $\mathcal{G}_1$ denote the arithmetic and geometric means, respectively, of the eigenvalues of $n^{-1}\boldsymbol{A}_{11\cdot 2,n}$, and let $\mathcal{A}_2$ and $\mathcal{G}_2$ denote the same for $N^{-1}\boldsymbol{A}_{22,N}$. Because $\boldsymbol{A}_{11} = \boldsymbol{A}_{11\cdot 2,n} + \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$ then

$$\begin{aligned}\lambda_5^2 &= \frac{|n^{-1}\boldsymbol{A}_{11\cdot 2,n}|^n |N^{-1}\boldsymbol{A}_{22,N}|^N}{\big((np+Nq)^{-1}(\operatorname{tr}\boldsymbol{A}_{11\cdot 2,n} + \operatorname{tr}\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21} + \operatorname{tr}\boldsymbol{A}_{22,N})\big)^{np+Nq}} \\ &\le \frac{|n^{-1}\boldsymbol{A}_{11\cdot 2,n}|^n |N^{-1}\boldsymbol{A}_{22,N}|^N}{\big((np+Nq)^{-1}(\operatorname{tr}\boldsymbol{A}_{11\cdot 2,n} + \operatorname{tr}\boldsymbol{A}_{22,N})\big)^{np+Nq}} \\ &\equiv \frac{\mathcal{G}_1^{np}\mathcal{G}_2^{Nq}}{\big((np+Nq)^{-1}(np\mathcal{A}_1 + Nq\mathcal{A}_2)\big)^{np+Nq}}. \end{aligned} \tag{4.15}$$

By the weighted arithmetic–geometric mean inequality (Marshall and Olkin [24, p. 455]),

$$(np+Nq)^{-1}(np\mathcal{A}_1 + Nq\mathcal{A}_2) \ge (\mathcal{A}_1^{np}\mathcal{A}_2^{Nq})^{1/(np+Nq)}.$$

Therefore, since $\mathcal{G}_j \le \mathcal{A}_j$, $j = 1, 2$, we obtain

$$\lambda_5^2 \le \frac{\mathcal{G}_1^{np}\mathcal{G}_2^{Nq}}{\Big((\mathcal{A}_1^{np}\mathcal{A}_2^{Nq})^{1/(np+Nq)}\Big)^{np+Nq}} \equiv \left(\frac{\mathcal{G}_1}{\mathcal{A}_1}\right)^{np}\left(\frac{\mathcal{G}_2}{\mathcal{A}_2}\right)^{Nq} \le 1.$$

We remark also that (4.15) shows how $\lambda_5$ may be expressed entirely in terms of the eigenvalues of $\boldsymbol{A}_{11\cdot 2,n}$, $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$, and $\boldsymbol{A}_{22,N}$.

**Theorem 4.7.** *For $h = 0, 1, 2, \ldots$ the h-th null moment of $\lambda_5$ is*

$$\begin{aligned} E(\lambda_5^h) = {} & \frac{(np+Nq)^{(np+Nq)h/2}}{n^{nph/2}N^{Nqh/2}} \frac{\Gamma_p\left(\frac{1}{2}(nh+n-q-1)\right)\Gamma_q\left(\frac{1}{2}(Nh+N-1)\right)}{\Gamma_p\left(\frac{1}{2}(n-q-1)\right)\Gamma_q\left(\frac{1}{2}(N-1)\right)} \\ & \times \frac{\Gamma\left(\frac{1}{2}((n-1)p+(N-1)q)\right)}{\Gamma\left(\frac{1}{2}((n-1)p+(N-1)q)+\frac{1}{2}(np+Nq)h\right)}. \end{aligned} \tag{4.16}$$

*Under $H_0$,*

$$\lambda_5 \stackrel{\mathcal{L}}{=} \frac{(np+Nq)^{(np+Nq)/2}}{n^{np/2}N^{Nq/2}} \left(\prod_{j=1}^{p} U_j\right)^{n/2} \left(\prod_{j=p+1}^{p+q} U_j\right)^{N/2} \left(\prod_{j=2}^{p} U_{1j}\right)^{n/2} \left(\prod_{j=p+1}^{p+q} U_j\right)^{N/2}, \tag{4.17}$$

*where*

$$(U_1, \ldots, U_{p+q}) \sim SD_{p+q}\left(\underbrace{\frac{1}{2}(n-q-1), \ldots, \frac{1}{2}(n-q-1)}_{p}, \underbrace{\frac{1}{2}(N-1), \ldots, \frac{1}{2}(N-1)}_{q}\right),$$

*a singular Dirichlet distribution; $U_{1j} \sim \beta(\frac{1}{2}(n-q-i+1), \frac{1}{2}(i-1))$, $2 \le j \le p$; $U_{2j} \sim \beta(\frac{1}{2}(N-i+1), \frac{1}{2}(i-1))$, $2 \le j \le q$; and $(U_1, \ldots, U_{p+q})$, $U_{12}, \ldots, U_{1p}$, $U_{22}, \ldots, U_{2p}$ are mutually independent.*

**Proof.** Under $H_0$ an invariance argument allows us to assume that $\sigma^2 = 1$, and hence $\boldsymbol{\Sigma} = \boldsymbol{I}_{p+q}$. Then $\boldsymbol{A}_{11\cdot 2,n}$, $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$, and $\boldsymbol{A}_{22,N}$ are mutually independent. By (4.14),

$$E(\lambda_5^h) = n^{-nph/2}N^{-Nqh/2}(np+Nq)^{(np+Nq)h/2}E|\boldsymbol{W}_1|^{nh/2}|\boldsymbol{W}_3|^{Nh/2}(\operatorname{tr}\boldsymbol{W}_1 + \operatorname{tr}\boldsymbol{W}_2 + \operatorname{tr}\boldsymbol{W}_3)^{-(np+Nq)h/2}, \tag{4.18}$$

where $\boldsymbol{W}_1 \sim W_p(n-q-1, \boldsymbol{I}_p)$, $\boldsymbol{W}_2 \sim W_p(q, \boldsymbol{I}_p)$, and $\boldsymbol{W}_3 \sim W_q(N-1, \boldsymbol{I}_q)$ are independent. When the density function of $\boldsymbol{W}_1$ is multiplied by the term $|\boldsymbol{W}_1|^{nh/2}$, the outcome is a constant multiple of the density function of $\widetilde{\boldsymbol{W}}_1 \sim W_p(nh+n-q-1, \boldsymbol{I}_p)$. Similarly, when the density function of $\boldsymbol{W}_3$ is multiplied by the term $|\boldsymbol{W}_3|^{Nh/2}$, the outcome is a constant multiple of the density function of $\widetilde{\boldsymbol{W}}_3 \sim W_q(Nh+N-1, \boldsymbol{I}_q)$. Therefore

$$\begin{aligned} & E|\boldsymbol{W}_1|^{nh/2}|\boldsymbol{W}_3|^{Nh/2}(\operatorname{tr}\boldsymbol{W}_1 + \operatorname{tr}\boldsymbol{W}_2 + \operatorname{tr}\boldsymbol{W}_3)^{-(np+Nq)h/2} \\ & \quad = \frac{c_p(n-q-1, \boldsymbol{I}_p)c_q(N-1, \boldsymbol{I}_q)}{c_p(nh+n-q-1, \boldsymbol{I}_p)c_q(Nh+N-1, \boldsymbol{I}_q)} E(\operatorname{tr}\widetilde{\boldsymbol{W}}_1 + \operatorname{tr}\boldsymbol{W}_2 + \operatorname{tr}\widetilde{\boldsymbol{W}}_3)^{-(np+Nq)h/2}, \end{aligned} \tag{4.19}$$

where $c_p(n-p-1, \boldsymbol{I}_p)$ denotes the usual Wishart normalizing constant. By [12, p. 107, Theorem 3.2.20], we have $\operatorname{tr} \widetilde{\boldsymbol{W}}_1 \sim \chi^2_{(nh+n-q-1)p}$, $\operatorname{tr} \boldsymbol{W}_2 \sim \chi^2_{qp}$, and $\operatorname{tr} \widetilde{\boldsymbol{W}}_3 \sim \chi^2_{(Nh+N-1)q}$, and hence $\operatorname{tr} \widetilde{\boldsymbol{W}}_1 + \operatorname{tr} \boldsymbol{W}_2 + \operatorname{tr} \widetilde{\boldsymbol{W}}_3 \sim \chi^2_{(nh+n-1)p+(Nh+N-1)q}$. Applying the formula

$$E(\chi^2_r)^{-\delta/2} = \frac{\Gamma((r-\delta)/2)}{2^{\delta/2}\Gamma(r/2)},$$

$\delta < r$, to $\operatorname{tr} \widetilde{\boldsymbol{W}}_1 + \operatorname{tr} \boldsymbol{W}_2 + \operatorname{tr} \widetilde{\boldsymbol{W}}_3$ in (4.19) and substituting the result in (4.18), we obtain

$$E(\lambda_5^h) = \frac{(np+Nq)^{(np+Nq)h/2}}{n^{nph/2}N^{Nqh/2}} \frac{c_p(n-q-1, \boldsymbol{I}_p)c_q(N-1, \boldsymbol{I}_q)}{c_p(nh+n-q-1, \boldsymbol{I}_p)c_q(Nh+N-1, \boldsymbol{I}_q)} \times \frac{\Gamma\left(\frac{1}{2}((n-1)p+(N-1)q)\right)}{2^{(np+Nq)h/2}\Gamma\left(\frac{1}{2}((n-1)p+(N-1)q)+\frac{1}{2}(np+Nq)h\right)}.$$

Substituting from (2.2) for the multivariate gamma function, we obtain (4.16).

To prove (4.17), we rewrite (4.16) as a product of four ratios,

$$E(\lambda_5^h) = \frac{\Gamma_p\left(\frac{1}{2}(nh+n-q-1)\right)\Gamma^p\left(\frac{1}{2}(n-q-1)\right)}{\Gamma_p\left(\frac{1}{2}(n-q-1)\right)\Gamma^p\left(\frac{1}{2}(nh+n-q-1)\right)} \frac{\Gamma_q\left(\frac{1}{2}(Nh+N-1)\right)\Gamma^q\left(\frac{1}{2}(N-1)\right)}{\Gamma_q\left(\frac{1}{2}(N-1)\right)\Gamma^q\left(\frac{1}{2}(Nh+N-1)\right)} \times \frac{\Gamma\left(\frac{1}{2}((n-1)p+(N-1)q)\right)}{\Gamma\left(\frac{1}{2}((n-1)p+(N-1)q)+\frac{1}{2}(np+Nq)h\right)} \frac{\Gamma^p\left(\frac{1}{2}(nh+n-q-1)\right)\Gamma^q\left(\frac{1}{2}(Nh+N-1)\right)}{\Gamma^p\left(\frac{1}{2}(n-q-1)\right)\Gamma^q\left(\frac{1}{2}(N-1)\right)}. \tag{4.20}$$

The first ratio in this product is the $h$-th moment of a classical sphericity statistic; see [13, p. 435, Eq. (16)], from which we deduce that the ratio is the $h$-th moment of a product of powers of independent beta random variables, $(\prod_{j=2}^{p} U_{1j})^{n/2}$, where $U_{1j} \sim \beta(\frac{1}{2}(n-q-i+1), \frac{1}{2}(i-1))$, $2 \le j \le p$. Similarly, the second ratio in (4.20) is the $h$-th moment of $(\prod_{j=2}^{q} U_{2j})^{N/2}$, with independent $U_{2j} \sim \beta(\frac{1}{2}(N-i+1), \frac{1}{2}(i-1))$, $2 \le j \le q$. By applying the formula for the density function of the singular Dirichlet distribution (see [24, p. 307], Eq. (11)), we find that the product of the last two ratios in (4.20) is the $h$-th moment of $(\prod_{j=1}^{p} U_j)^{n/2}(\prod_{j=p+1}^{p+q} U_j)^{N/2}$, where $(U_1, \ldots, U_{p+q})$ is as stated earlier. Combining these results, we obtain (4.17). □

We have been unable to determine whether or not $\lambda_5$ is unbiased; in particular, the methods of Gleser [25] or Sugiura and Nagao [22] seem inapplicable to this problem. On the other hand, the non-null distribution of $\lambda_5$ can be obtained using the methods given here, suitably generalizing the approach provided by Muirhead [12, p. 339 ff.].

### *4.4. Testing independence between subsets of the variables*

Consider the problem of testing $H_0 : \boldsymbol{\Sigma}_{12} = \boldsymbol{0}$ against $H_a : \boldsymbol{\Sigma}_{12} \neq \boldsymbol{0}$ with the sample (1.1). Eaton and Kariya [11] showed that the likelihood ratio test statistic ignores the incomplete data $\boldsymbol{Y}_j$, $j = n+1, \ldots, N$, and they proved that, among the class of affinely invariant test procedures, the test that rejects $H_0$ for small values of

$$\lambda_6 = \operatorname{tr} \boldsymbol{A}_{22,n}(\boldsymbol{A}_{22,n} + \boldsymbol{B}_1)^{-1} - np^{-1} \operatorname{tr} \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}(\boldsymbol{A}_{22,n} + \boldsymbol{B}_1)^{-1}\boldsymbol{A}_{21}$$

is locally most powerful invariant, where $\boldsymbol{A}_{11}$, $\boldsymbol{A}_{12}$, $\boldsymbol{A}_{22,n}$, and $\boldsymbol{B}_1$ are given in (2.3) and (3.5), respectively; cf. [26–28]. To date, the distribution theory of $\lambda_6$ remains explored and seems recondite. On the other hand, by omitting the term $np^{-1}$, we obtain the modified statistic,

$$\begin{aligned}\lambda_7 &= \operatorname{tr} \boldsymbol{A}_{22,n}(\boldsymbol{A}_{22,n} + \boldsymbol{B}_1)^{-1} - \operatorname{tr} \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}(\boldsymbol{A}_{22,n} + \boldsymbol{B}_1)^{-1}\boldsymbol{A}_{21} \\ &= \operatorname{tr} \left(\boldsymbol{A}_{22,n} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\right)(\boldsymbol{A}_{22,n} + \boldsymbol{B}_1)^{-1}.\end{aligned}$$

The statistic $\lambda_7$ will not generally enjoy the same optimality properties as $\lambda_6$. However, $\lambda_6 \le \lambda_7$ for $n \ge p$, in which case if $H_0$ is rejected for small values of $\lambda_7$ then $H_0$ also is rejected by $\lambda_6$. Moreover, $\lambda_7$ has a null distribution which is simpler than that of $\lambda_6$. Indeed, with $\boldsymbol{W}_1 = \boldsymbol{A}_{22,n} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$ and $\boldsymbol{W}_2 = \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} + \boldsymbol{B}_1$, we have $\lambda_7 = \operatorname{tr} \boldsymbol{W}_1(\boldsymbol{W}_1 + \boldsymbol{W}_2)^{-1}$. By [13, pp. 142–143] we obtain that, under $H_0$, $\boldsymbol{W}_1 \sim W_q(n-p-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{W}_2 \sim W_q(N-n+p-1, \boldsymbol{\Sigma}_{22})$, and $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are independent. Therefore, under $H_0$, $\lambda_7$ is exactly of the form of the Bartlett–Nanda–Pillai criterion in MANOVA ([13, Section 8.6.3], [12, Section 10.6.3]), and its null distribution may be derived accordingly.

## 5. Concluding remarks

In the case of $k$-step monotone incomplete data, many open problems remain. Romer [29] has derived some results on exact stochastic representations for $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ for $k = 3$, but little is known for $k \ge 4$ and this has prevented extensions of results in [1,30]. The likelihood ratio test procedures in Section 4 have been extended [18, Section 3.4] to the $k$-step case, however it seems formidable to extend similarly the unbiasedness results in Section 4. Romer [29] has derived, by highly non-trivial methods, an exact stochastic representation for the analog of Hotelling's $T^2$-statistic in the two-step case, and the $k$-step case remains open.

As regards the case of non-monotone incomplete data, many problems remain unexplored. In the case of [11], the likelihood equations for $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are unsolved; indeed, Romer and Richards (unpublished notes) have proved that those equations have multiple solutions.

## Acknowledgments

The authors are grateful to the referees and an associate editor for their valuable comments on an earlier version of this article. The second author was supported in part by National Science Foundation grants DMS-0705210 and DMS-0112069.

## References

[1] W.-Y. Chang, D.St.P. Richards, Finite-sample inference with monotone incomplete multivariate normal data, I, J. Multivariate Anal. 100 (2009) 1883–1899.
[2] T.W. Anderson, Maximum likelihood estimators for a multivariate normal distribution when some observations are missing, J. Amer. Statist. Assoc. 52 (1957) 200–203.
[3] T.W. Anderson, I. Olkin, Maximum likelihood estimation of the parameters of a multivariate normal distribution, Linear Algebra Appl. 70 (1985) 147–171.
[4] M.A. Giguère, G.P.H. Styan, Multivariate normal estimation with missing data on several variates, in: Trans. Seventh Prague Conference on Information Theory, Statistical Decision Functions, Academia, Prague, 1978, pp. 129–139.
[5] K.G. Jinadasa, D.S. Tracy, Maximum likelihood estimation for multivariate normal distribution with monotone sample, Comm. Statist. Theory Methods 21 (1992) 41–50.
[6] R. Bhargava, Multivariate tests of hypotheses with incomplete data, Tech. Report No. 3, Applied Math. and Statistics Laboratories, Stanford University, Stanford, CA, 1962.
[7] R.P. Bhargava, Some one-sample hypothesis testing problems when there is a monotone sample from a multivariate normal population, Ann. Inst. Statist. Math. 27 (1975) 327–339.
[8] D.F. Morrison, Expectations and variances of maximum likelihood estimates of the multivariate normal distribution parameters with missing data, J. Amer. Statist. Assoc. 66 (1971) 602–604.
[9] R.J.A. Little, D.B. Rubin, Statistical Analysis With Missing Data, second ed., Wiley-Interscience, Hoboken, NJ, 2002.
[10] T. Kanda, Y. Fujikoshi, Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data, Amer. J. Math. Management Sci. 18 (1998) 161–190.
[11] M.L. Eaton, T. Kariya, Multivariate tests with incomplete data, Ann. Statist. 11 (1983) 654–665.
[12] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[13] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, third ed., Wiley, New York, 2003.
[14] C.S. Herz, Bessel functions of matrix argument, Ann. Math. 61 (1955) 474–523.
[15] M.L. Eaton, Multivariate Statistics. A Vector Space Approach, Wiley, New York, 1983.
[16] H. Fujisawa, A note on the maximum likelihood estimators for multivariate normal distribution with monotone data, Comm. Statist. Theory Methods 24 (1995) 1377–1382.
[17] R.W. Butler, A.T.A. Wood, Laplace approximations for hypergeometric functions with matrix argument, Ann. Statist. 30 (2002) 1155–1177.
[18] J. Hao, K. Krishnamoorthy, Inferences on a normal covariance matrix and generalized variance with monotone missing data, J. Multivariate Anal. 78 (2001) 62–82.
[19] S. Kotz, N.L. Johnson, D.W. Boyd, Series representations of distributions of quadratic forms in normal variables. I. Central case, Ann. Math. Statist. 38 (1967) 823–837.
[20] D. Kuonen, Saddlepoint approximations for distributions of quadratic forms in normal variables, Biometrika 86 (1999) 929–935.
[21] J.G. Booth, R.W. Butler, S. Huzurbazar, A.T.A. Wood, Saddlepoint approximations for $P$-values of some tests of covariance matrices, J. Statist. Comput. Simul. 53 (1995) 165–180.
[22] N. Sugiura, H. Nagao, Unbiasedness of some test criteria for the equality of one or two covariance matrices, Ann. Math. Statist. 39 (1968) 1686–1692.
[23] S. Das Gupta, Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions, Ann. Math. Statist. 40 (1969) 697–701.
[24] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.
[25] L.J. Gleser, A note on the sphericity test, Ann. Math. Statist. 37 (1966) 464–467.
[26] R.-J. Chou, W.D. Lo, On the local minimaxity of a test of independence in incomplete samples, Ann. Inst. Statist. Math. 38 (1986) 495–502.
[27] W.D. Lo, R.-J. Chou, Test of independence in incomplete samples, Bull. Inst. Math. Acad. Sinica 14 (1986) 117–127.
[28] M.S. Srivastava, S.R. Chakravorti, Asymptotic distributions of two test statistics for testing independence with missing data, Comm. Statist. A—Theory Methods 15 (1986) 571–588.
[29] M.M. Romer, The Statistical Analysis of Monotone Incomplete Multivariate Normal Data, Doctoral Dissertation, Penn State University, 2009.
[30] D.St.P. Richards, T. Yamada, The Stein phenomenon for monotone incomplete multivariate normal data, J. Multivariate Anal., 2009 (submitted for publication).