Finite-Sample Inference with Monotone Incomplete Multivariate Normal Data, II

Wan-Ying Chang* and Donald St. P. Richards†

July 24, 2008

Abstract

We continue our recent work on finite-sample, i.e., non-asymptotic, inference with two-step, monotone incomplete data from $N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 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 $\widehat{\boldsymbol{\Sigma}}$ and of the estimated regression matrix, $\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}$. We obtain a representation for $\widehat{\boldsymbol{\Sigma}}$ in terms of independent matrices, and derive its exact density function, thereby generalizing the Wishart distribution to the setting of monotone incomplete data, and obtain saddlepoint approximations for the distributions of $\widehat{\boldsymbol{\Sigma}}$ and its partial Iwasawa coordinates. We establish 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 that of a Bartlett-Pillai-Nanda trace-statistic in multivariate analysis of variance.

1 Introduction

In this paper, we continue our recent work [8] on inference with incomplete multivariate normal data. Our data are independent observations consisting of a random sample of n complete observations on all $d = p + q$ characteristics together with an additional $N - n$ incomplete observations on the last q characteristics only. We write the data in the form

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

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

*Washington Department of Fish and Wildlife, Olympia, WA 98501, USA.

†Department of Statistics, Penn State University, University Park, PA 16802; and The Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, NC 27709, USA.

†Supported in part by National Science Foundation grants DMS-0705210 and DMS-0112069.

2000 Mathematics Subject Classification: Primary 62H10; Secondary 60D10, 62E15.

Key words and phrases. Likelihood ratio tests, locally most powerful invariant tests, matrix F-distribution, maximum likelihood estimation, missing data, testing independence, sphericity test, unbiased test statistics, Wishart distribution.

characteristics of the same population. The data in (1.1) are called *two-step monotone*; cf. [8] for numerous references to the literature on these data.

Closed-form expressions for $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$, the maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, have long been available (cf. Anderson [1], Anderson and Olkin [2], Jinadasa and Tracy [18], and those formulas have been utilized in inference for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ (cf. Bhargava [4, 5], Morrison [27], Giguère and Styan [14], Little and Rubin [24], Kanda and Fujikoshi [20]). Here, we continue our program of research on inference for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by means of results on the exact distributions of $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$. Having derived in [8] the exact distribution of $\widehat{\boldsymbol{\mu}}$ and making applications to inference for $\boldsymbol{\mu}$, we now turn our attention to inference for $\boldsymbol{\Sigma}$.

Under the assumption that $\boldsymbol{\Sigma}$ is block-diagonal when partitioned according to the two-step pattern, we derive in Section 3 the distributions of the diagonal blocks of $\widehat{\boldsymbol{\Sigma}}$ and of the estimated regression matrix, $\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}$. We also obtain a stochastic representation for $\widehat{\boldsymbol{\Sigma}}$ and thereby derive its exact distribution, hence extending the Wishart distribution to the setting of monotone incomplete data. Further, we obtain saddlepoint approximations for $\widehat{\boldsymbol{\Sigma}}$ and its partial Iwasawa coordinates.

In Section 4, we consider 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, 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 of, and a stochastic representation for, the likelihood ratio criterion. In testing independence between the first p and the last q characteristics of the population, Eaton and Kariya [12] 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

Throughout this paper, we retain the notation and conventions of [8], writing all vectors and matrices in boldface type. In particular, we denote by $\mathbf{0}$ any zero vector or matrix, the dimension of which will be clear from the context, and we denote the identity matrix of order d by $\boldsymbol{I}_d$. We write $\boldsymbol{A} > \mathbf{0}$ to denote that a matrix $\boldsymbol{A}$ is positive definite (symmetric), and we write $\boldsymbol{A} \ge \boldsymbol{B}$ to mean that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite.

Suppose that $\boldsymbol{W} \sim \mathrm{W}_d(a, \boldsymbol{\Lambda})$, a Wishart distribution, where $a > d-1$ and $\boldsymbol{\Lambda} > \mathbf{0}$, i.e., $\boldsymbol{W}$ is a $d \times d$ random matrix with density function

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

$\boldsymbol{W} > \mathbf{0}$, where

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

is the multivariate gamma function ([28], p. 62).

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 encountered later are absolutely convergent. We use the notation $\tau = n/N$ for the proportion of data which are complete and denote $1 - \tau$ by $\bar{\tau}$, so that $\bar{\tau} = (N-n)/N$ is the proportion of incomplete observations.

Define sample means

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

and the corresponding matrices of sums of squares and products by

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

By Anderson [1] (cf. Morrison [27], Anderson and Olkin [2], Jinadasa and Tracy [18]), the maximum likelihood estimator of $\boldsymbol{\mu}$ is $\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1 \\ \widehat{\boldsymbol{\mu}}_2 \end{pmatrix}$, where

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

3 The distribution of $\widehat{\boldsymbol{\Sigma}}$

By Anderson [1] or Anderson and Olkin [2] (cf. Morrison [27], Giguère and Styan [14]), the maximum likelihood estimator of $\boldsymbol{\Sigma}$ is $\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_{11} & \widehat{\boldsymbol{\Sigma}}_{12} \\ \widehat{\boldsymbol{\Sigma}}_{21} & \widehat{\boldsymbol{\Sigma}}_{22} \end{pmatrix}$ where, in the notation of (2.3),

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

3.1 A representation for $\widehat{\boldsymbol{\Sigma}}$

Proposition 3.1. *Let* $\boldsymbol{A}_{11\cdot 2,n} := \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$. *Then*

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

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

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 [27], eq. (3.4).

As noted earlier (cf. (3.23), *infra*), $\bar{\boldsymbol{Y}}_1$ is independent of

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

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 $\boldsymbol{B}_2$, 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 $N\widehat{\boldsymbol{\Sigma}}_{22} = \boldsymbol{A}_{22,N} = \boldsymbol{A}_{22,n} + \boldsymbol{B} \sim W_q(N-1, \boldsymbol{\Sigma}_{22})$. □

We now establish some results that were stated in [8], 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 applying the usual independence of the mean and covariance matrix of a random sample from a multivariate normal population, and by the independence of the individual observations in the data set, 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 also, by Proposition 2.2(iii), 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 last, 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 transformation 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

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

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 $\operatorname{Cov}(\bar{\boldsymbol{Y}}_j)$ if $j = k$; hence the cross-covariance matrix equals $\tau \operatorname{Cov}(\bar{\boldsymbol{Y}}_1) - \bar{\tau} \operatorname{Cov}(\bar{\boldsymbol{Y}}_2) = (\tau n^{-1} - \bar{\tau}(N-n)^{-1})\boldsymbol{\Sigma}_{22} = \boldsymbol{0}$, 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 $p \le q$. As in [8], 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 [8], 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 Haar 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 $\boldsymbol{F} \ge \boldsymbol{0}$ is said to have a *matrix* F*-distribution*, denoted $\boldsymbol{F} \sim \mathrm{F}_{a,b}^{(q)}$, with degrees of freedom (a, b), $a \ge 0$, $b > q - 1$, 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 $a \le q - 1$ 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

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

$\boldsymbol{F} > \boldsymbol{0}$. 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 (see [28], 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 $\mathrm{F}_{a,b}^{(q)}$. Further, if $\boldsymbol{F} \sim \mathrm{F}_{a,b}^{(q)}$ then $\boldsymbol{F}^{-1} \sim \mathrm{F}_{b,a}^{(q)}$. If $\boldsymbol{F} \sim \mathrm{F}_{a,b}^{(q)}$ then, assuming $\boldsymbol{\Sigma}_{22} = \boldsymbol{I}_q$ (with no loss of generality), we obtain $|\boldsymbol{F}| \stackrel{\mathcal{L}}{=} |\boldsymbol{A}|/|\boldsymbol{B}|$; recalling that $|\boldsymbol{A}| \stackrel{\mathcal{L}}{=} \prod_{j=1}^{q} \chi^2_{a-j+1}$, a product of independent chi-squared variables, with a similar result also holding for $|\boldsymbol{B}|$, we obtain $|\boldsymbol{F}| \stackrel{\mathcal{L}}{=} \prod_{j=1}^{p} \mathrm{F}^{(1)}_{a-j+1,b-j+1}$.

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

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

proving that $\boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' \overset{\mathcal{L}}{=} \boldsymbol{F}_{11}$. Next, since $\boldsymbol{F} \sim \mathrm{F}^{(q)}_{a,b}$ then $\boldsymbol{F} \overset{\mathcal{L}}{=} \boldsymbol{A}^{1/2}\boldsymbol{B}^{-1}\boldsymbol{A}^{1/2}$, 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,

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

where $\boldsymbol{M} = (\boldsymbol{I}_p \vdots \boldsymbol{0})\boldsymbol{A}^{1/2}$. By [8], Proposition 2.2(iv), conditional on $\boldsymbol{M}$, $(\boldsymbol{M}\boldsymbol{B}^{-1}\boldsymbol{M}')^{-1} \sim \mathrm{W}_p(b - q + p, (\boldsymbol{M}\boldsymbol{M}')^{-1})$, so the conditional density function of $\boldsymbol{R} = \boldsymbol{F}_{11}^{-1}$ given $\boldsymbol{S} = \boldsymbol{M}\boldsymbol{M}'$ is

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

$\boldsymbol{R}, \boldsymbol{S} > \boldsymbol{0}$. 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

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

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

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

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

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

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

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 [8], Proposition 2.2(i, ii), $\boldsymbol{A}_{11\cdot 2,n}$, 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., $\boldsymbol{K} \stackrel{\mathcal{L}}{=} \boldsymbol{KH}$ for all $\boldsymbol{H} \in O(q)$. By polar coordinates on matrix space ([17], p. 482; [11], p. 163), $\boldsymbol{K} \stackrel{\mathcal{L}}{=} \boldsymbol{W}_2^{1/2}\boldsymbol{H}_1$ where $\boldsymbol{W}_2$ and $\boldsymbol{H}_1$ are independent, $\boldsymbol{W}_2 = \boldsymbol{KK}' \sim \mathrm{W}_p(q, \boldsymbol{I}_p)$, and $\boldsymbol{H}_1$ is uniformly distributed on the Stiefel manifold $S_{p,q}$ (see [28], 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 $\boldsymbol{F} = \boldsymbol{A}_{22,n}^{-1/2}\boldsymbol{B}\boldsymbol{A}_{22,n}^{-1/2} \sim \mathrm{F}^{(q)}_{N-n,n-1}$. Therefore

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

By Lemma 3.3, $\boldsymbol{H}_1\boldsymbol{F}\boldsymbol{H}_1' \sim \mathrm{F}^{(p)}_{N-n,n-q+p-1}$, 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

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

In particular, $N^p|\widehat{\boldsymbol{\Sigma}}_{11}|/|\boldsymbol{\Sigma}_{11}| \stackrel{\mathcal{L}}{\geq} |\boldsymbol{W}_1 + \boldsymbol{W}_2|$, so that for all $\delta \geq 0$,

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

As an application of this stochastic ordering, we construct a $100(1-\alpha)\%$ 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 δ_α is an upper $\alpha\%$ significance point for $|\boldsymbol{W}_1 + \boldsymbol{W}_2|$, i.e., $P(|\boldsymbol{W}_1 + \boldsymbol{W}_2| \geq \delta_\alpha) = \alpha$, then

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

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

3.2 The distribution of the estimated regression matrix

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

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

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

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,* $\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}$ *is an unbiased estimator of* $\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$*.*

Proof. By (3.1), $\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}$. Set $\boldsymbol{A}_{12} = \boldsymbol{B}_{12} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{A}_{22,n}$; then, as in the proof of Theorem 3.1 in [8], $\boldsymbol{B}_{12}|\boldsymbol{A}_{22,n} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{11\cdot 2} \otimes \boldsymbol{A}_{22,n})$. Therefore

$$\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} \overset{\mathcal{L}}{=} (\boldsymbol{B}_{12} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{A}_{22,n})\boldsymbol{A}_{22,n}^{-1} = \boldsymbol{B}_{12}\boldsymbol{A}_{22,n}^{-1} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1},$$

or $\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \overset{\mathcal{L}}{=} \boldsymbol{B}_{12}\boldsymbol{A}_{22,n}^{-1}$. Also, $\boldsymbol{B}_{12} \overset{\mathcal{L}}{=} \boldsymbol{\Sigma}_{11\cdot 2}^{1/2}\boldsymbol{K}\boldsymbol{A}_{22,n}^{1/2}$, where $\boldsymbol{K} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_p \otimes \boldsymbol{I}_q)$, so

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

For $\boldsymbol{T} \in \mathbb{R}^{p\times q}$, the characteristic function of $\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})$ is

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

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

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

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 [17]. Applying a formula of Herz [17], p. 506, we have

$$B_\delta^{(q)}(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2) = |\boldsymbol{\Lambda}_1|^{-\delta}\int_{\boldsymbol{W}>\boldsymbol{0}} \exp\big(-\mathrm{tr}\,(\boldsymbol{W}\boldsymbol{\Lambda}_1 + \boldsymbol{W}^{-1}\boldsymbol{\Lambda}_2)\big)\,|\boldsymbol{W}|^{-\delta-\frac{1}{2}(q+1)}\,d\boldsymbol{W}, \tag{3.8}$$

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

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

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 [17], p. 509, Theorem 5.10,

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

therefore

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

On applying (3.8) to express this Bessel function as an integral over the space of $p \times p$ positive definite matrices, we obtain

$$E\exp\big(i\mathrm{tr}\,\boldsymbol{T}'\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\big) = E\exp(-\tfrac{1}{2}\mathrm{tr}\,\boldsymbol{T}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{T}'\boldsymbol{W}^{-1}), \tag{3.9}$$

$\boldsymbol{W} \sim \mathrm{W}_p(n-q+p-1, \boldsymbol{I}_p)$. However the right-hand side of (3.9) equals

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

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

Remark 3.7. By (3.7),

$$\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2}(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})\boldsymbol{\Sigma}_{22}(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})'\boldsymbol{\Sigma}_{11\cdot 2}^{-1/2} \stackrel{\mathcal{L}}{=} \boldsymbol{W}^{-1/2}(\boldsymbol{K}\boldsymbol{K}')\boldsymbol{W}^{-1/2}. \quad (3.10)$$

Since $\boldsymbol{K}\boldsymbol{K}' \sim \mathrm{W}_p(q, \boldsymbol{I}_p)$ then the right-hand side of (3.10) has an $\mathrm{F}^{(p)}_{q,n-q+p-1}$ distribution.

For the case in which $\boldsymbol{\Sigma} = \boldsymbol{I}_{p+q}$, (3.10) implies that $(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})' \sim \mathrm{F}^{(p)}_{q,n-q+p-1}$. This result allows us to conduct exploratory analyses alternative to likelihood ratio testing to assess the plausibility of the hypothesis $H_0 : \boldsymbol{\Sigma} = \boldsymbol{I}_{p+q}$. To that end, we construct the matrix $(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})'$ from the monotone data and compare the empirical distribution of its eigenvalues with the distribution of the eigenvalues of an $\mathrm{F}^{(p)}_{q,n-q+p-1}$-distributed matrix, or we may compare the determinant $|(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})'|$ with $|\boldsymbol{F}|$, the determinant of the corresponding $\mathrm{F}^{(p)}_{q,n-q+p-1}$-distributed matrix. Here, we recall that $|\boldsymbol{F}| \stackrel{\mathcal{L}}{=} \prod_{j=1}^{p} \mathrm{F}^{(1)}_{q-j+1,n-q+p-j}$, a product of independent, F-distributed random variables. By simulating the distribution of this product, we may ascertain whether the observed value of $|(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})'|$ is greater than specified percentage points of the distribution of $|\boldsymbol{F}|$. We may also perform exploratory analyses using, say, the trace or extreme eigenvalues of $(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})(\widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1})'$, comparing the sample estimates with the percentage points of corresponding functions of the $\mathrm{F}^{(p)}_{q,n-q+p-1}$-distributed matrix.

3.3 The distribution of $\widehat{\boldsymbol{\Delta}}$

It is well-known (see [11], Proposition 8.7) that a generic positive definite matrix $\boldsymbol{\Sigma}$ can be expressed as

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{\Delta}_{12} \\ \boldsymbol{0} & \boldsymbol{I}_q \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Delta}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{0} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{I}_q \end{pmatrix}, \quad (3.11)$$

where

$$\boldsymbol{\Delta}_{11} = \boldsymbol{\Sigma}_{11\cdot 2}, \quad \boldsymbol{\Delta}_{12} = \boldsymbol{\Delta}_{21}' = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}, \quad \boldsymbol{\Delta}_{22} = \boldsymbol{\Sigma}_{22}.$$

This defines the positive definite symmetric matrix $\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} \end{pmatrix}$, and the set of submatrices $\{\boldsymbol{\Delta}_{11}, \boldsymbol{\Delta}_{12}, \boldsymbol{\Delta}_{22}\}$ are also called the *partial Iwasawa coordinates* of $\boldsymbol{\Sigma}$ (Fujisawa [13]).

Since $\begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{0} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{I}_q \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{0} \\ -\boldsymbol{\Delta}_{21} & \boldsymbol{I}_q \end{pmatrix}$ then, by inverting (3.11), we obtain

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \boldsymbol{\Delta}_{11}^{-1} & -\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Delta}_{12} \\ -\boldsymbol{\Delta}_{21}\boldsymbol{\Delta}_{11}^{-1} & \boldsymbol{\Delta}_{22}^{-1} + \boldsymbol{\Delta}_{21}\boldsymbol{\Delta}_{11}^{-1}\boldsymbol{\Delta}_{12} \end{pmatrix}.$$

Therefore the correspondence between $\boldsymbol{\Delta}$ and $\boldsymbol{\Sigma}$ is one-to-one, with the transformation from $\boldsymbol{\Delta}$ to $\boldsymbol{\Sigma}$ given by

$$\boldsymbol{\Sigma}_{11} = \boldsymbol{\Delta}_{11} + \boldsymbol{\Delta}_{12}\boldsymbol{\Delta}_{22}\boldsymbol{\Delta}_{21}, \quad \boldsymbol{\Sigma}_{12} = \boldsymbol{\Delta}_{12}\boldsymbol{\Delta}_{22}, \quad \boldsymbol{\Sigma}_{22} = \boldsymbol{\Delta}_{22}.$$

The maximum likelihood estimator of $\boldsymbol{\Delta}$ is $\widehat{\boldsymbol{\Delta}} := \begin{pmatrix} \widehat{\boldsymbol{\Delta}}_{11} & \widehat{\boldsymbol{\Delta}}_{12} \\ \widehat{\boldsymbol{\Delta}}_{21} & \widehat{\boldsymbol{\Delta}}_{22} \end{pmatrix}$, where each $\widehat{\boldsymbol{\Delta}}_{ij}$ is the corresponding maximum likelihood estimator of $\boldsymbol{\Delta}_{ij}$. By (3.1),

$$\begin{aligned} \widehat{\boldsymbol{\Delta}}_{11} &= \widehat{\boldsymbol{\Sigma}}_{11\cdot 2} = \frac{1}{n}\boldsymbol{A}_{11\cdot 2,n}, \\ \widehat{\boldsymbol{\Delta}}_{12} &= \widehat{\boldsymbol{\Delta}}_{21}' = \widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1} = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}, \\ \widehat{\boldsymbol{\Delta}}_{22} &= \widehat{\boldsymbol{\Sigma}}_{22} = \frac{1}{N}\boldsymbol{A}_{22,N}. \end{aligned} \tag{3.12}$$

There also holds a one-to-one correspondence between $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\Delta}}$, and the inverse transformation from $\widehat{\boldsymbol{\Delta}}$ to $\widehat{\boldsymbol{\Sigma}}$ is the same as that from $\boldsymbol{\Delta}$ to $\boldsymbol{\Sigma}$.

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

Lemma 3.8. *Let $\boldsymbol{\Xi}_1$, $\boldsymbol{\Xi}_2$, and $\boldsymbol{\Xi}_3$ be absolutely continuous random matrices of the same dimension such that $(\boldsymbol{\Xi}_1, \boldsymbol{\Xi}_2)$ and $\boldsymbol{\Xi}_3$ are independent. Then the conditional density function of $\boldsymbol{\Xi}_1$ given $\boldsymbol{\Xi}_2 + \boldsymbol{\Xi}_3 = \boldsymbol{\xi}$, is*

$$f_{\boldsymbol{\Xi}_1|\boldsymbol{\Xi}_2+\boldsymbol{\Xi}_3=\boldsymbol{\xi}}(\boldsymbol{\xi}_1) = \frac{1}{f_{\boldsymbol{\Xi}_2+\boldsymbol{\Xi}_3}(\boldsymbol{\xi})} \int f_{\boldsymbol{\Xi}_1|\boldsymbol{\Xi}_2=\boldsymbol{\xi}_2}(\boldsymbol{\xi}_1)\, f_{\boldsymbol{\Xi}_2}(\boldsymbol{\xi}_2)\, f_{\boldsymbol{\Xi}_3}(\boldsymbol{\xi}-\boldsymbol{\xi}_2)\, d\boldsymbol{\xi}_2. \tag{3.13}$$

Proof. By a direct calculation,

$$\begin{aligned} f_{\boldsymbol{\Xi}_2+\boldsymbol{\Xi}_3}(\boldsymbol{\xi})\, f_{\boldsymbol{\Xi}_1|\boldsymbol{\Xi}_2+\boldsymbol{\Xi}_3=\boldsymbol{\xi}}(\boldsymbol{\xi}_1) &= f_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2+\boldsymbol{\Xi}_3}(\boldsymbol{\xi}_1,\boldsymbol{\xi}) \\ &= \int f_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2,\boldsymbol{\Xi}_3}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}-\boldsymbol{\xi}_2)\, d\boldsymbol{\xi}_2 \\ &= \int f_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)\, f_{\boldsymbol{\Xi}_3}(\boldsymbol{\xi}-\boldsymbol{\xi}_2)\, d\boldsymbol{\xi}_2 \\ &= \int f_{\boldsymbol{\Xi}_1|\boldsymbol{\Xi}_2=\boldsymbol{\xi}_2}(\boldsymbol{\xi}_1)\, f_{\boldsymbol{\Xi}_2}(\boldsymbol{\xi}_2)\, f_{\boldsymbol{\Xi}_3}(\boldsymbol{\xi}-\boldsymbol{\xi}_2) d\boldsymbol{\xi}_2. \end{aligned}$$

Dividing both sides of this equation by $f_{\boldsymbol{\Xi}_2+\boldsymbol{\Xi}_3}(\boldsymbol{\xi})$ completes the proof. □

In deriving the distribution of $\widehat{\boldsymbol{\Delta}}$ 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,

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

where $\boldsymbol{M}$ is $q \times q$ and symmetric; $\mathrm{Re}(b-a), \mathrm{Re}(a) > (q-1)/2$; and the region $\{\mathbf{0} < \boldsymbol{U} < \boldsymbol{I}_q\}$ consists of all $q \times q$ matrices $\boldsymbol{U}$ such that $\boldsymbol{U}$ and $\boldsymbol{I}_q - \boldsymbol{U}$ both are positive definite ([17]; [28], p. 264). For general a, b, these hypergeometric functions satisfy the reduction formula

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

and Kummer's formula ([17], eq. (2.8); [28], p. 265),

$$ {}_1F_1{}^{(q)}\Big(\begin{matrix} a \\ b \end{matrix}; \boldsymbol{M}\Big) = \exp(\operatorname{tr} \boldsymbol{M})\, {}_1F_1{}^{(q)}\Big(\begin{matrix} b-a \\ b \end{matrix}; -\boldsymbol{M}\Big). \tag{3.17}$$

If $\boldsymbol{M}$ is of rank $p \le q$ then ([17], Theorem 3.10, p. 497 and Theorem 4.15, p. 505)

$$ {}_1F_1{}^{(q)}\Big(\begin{matrix} a \\ b \end{matrix}; \boldsymbol{M}\Big) = {}_1F_1{}^{(p)}\Big(\begin{matrix} a \\ b \end{matrix}; \boldsymbol{M}_0\Big), \tag{3.18}$$

where $\boldsymbol{M}_0$ is any $p \times p$ symmetric matrix whose non-zero eigenvalues coincide with those of $\boldsymbol{M}$.

Theorem 3.9. *Let $n > p + q$ and $N - n > q - 1$. Then $f_{\widehat{\boldsymbol{\Delta}}}$, the density function of $\widehat{\boldsymbol{\Delta}}$, 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*

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

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

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

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

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

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

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

Proof. By Proposition 3.1,

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

consequently, by [8], Proposition 2.2(i), $\boldsymbol{A}_{11\cdot 2,n}$ and $\{\boldsymbol{A}_{12}, \boldsymbol{A}_{22,n}\}$ are mutually independent, hence so are $\boldsymbol{A}_{11\cdot 2,n}$ and $\{\boldsymbol{A}_{12}, \boldsymbol{A}_{22,N}\}$. Therefore $\widehat{\boldsymbol{\Delta}}_{11}$ and $\{\widehat{\boldsymbol{\Delta}}_{12}, \widehat{\boldsymbol{\Delta}}_{22}\}$ are mutually independent, so the joint density of $\widehat{\boldsymbol{\Delta}}$ can be expressed in the form (3.19). By (3.23), $n\widehat{\boldsymbol{\Delta}}_{11} = \boldsymbol{A}_{11\cdot 2,n} \sim \mathrm{W}_p(n-q-1, \boldsymbol{\Delta}_{11})$ and then (3.20) is obtained by a transformation of the Wishart density (2.1). Also, since $N\widehat{\boldsymbol{\Delta}}_{22} = \boldsymbol{A}_{22,N} \sim \mathrm{W}_q(N-1, \boldsymbol{\Delta}_{22})$ then (3.21) is obtained similarly.

By [8], Proposition 2.2(ii), $\widehat{\boldsymbol{\Delta}}_{12}|\boldsymbol{A}_{22,n} = \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}|\boldsymbol{A}_{22,n} \sim \mathrm{N}(\boldsymbol{\Delta}_{12}, \boldsymbol{\Delta}_{11} \otimes \boldsymbol{A}_{22,n}^{-1})$. Therefore, for $\boldsymbol{T}_{12} \in \mathbb{R}^{p\times q}$ and a $q \times q$ matrix $\boldsymbol{U} > \boldsymbol{0}$,

$$\begin{aligned} f_{\widehat{\boldsymbol{\Delta}}_{12}|N^{-1}\boldsymbol{A}_{22,n}=\boldsymbol{U}}(\boldsymbol{T}_{12}) &\equiv f_{\widehat{\boldsymbol{\Delta}}_{21}|\boldsymbol{A}_{22,n}=N\boldsymbol{U}}(\boldsymbol{T}_{12}) \\ &= (2\pi)^{-pq/2}\,|\boldsymbol{\Delta}_{11}|^{-q/2}\,N^{pq/2}\,|\boldsymbol{U}|^{p/2} \\ &\qquad \times \exp\big(-\tfrac{1}{2}N\,\mathrm{tr}\,\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{U}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\big). \end{aligned} \tag{3.24}$$

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

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

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

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

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

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

Introducing the temporary 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), (3.25), and (3.26), we find that we are to evaluate

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

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

$$\begin{aligned} |\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)} &\int_{\boldsymbol{0}<\boldsymbol{U}<\boldsymbol{I}_q} |\boldsymbol{U}|^{\frac{1}{2}(n+p-1)-\frac{1}{2}(q+1)}\,|\boldsymbol{I}_q-\boldsymbol{U}|^{\frac{1}{2}(N-n)-\frac{1}{2}(q+1)} \\ &\qquad \times \exp(-\mathrm{tr}\,\boldsymbol{T}_{22}^{1/2}\boldsymbol{M}_1\boldsymbol{T}_{22}^{1/2}\boldsymbol{U})\,d\boldsymbol{U} \\ &= B_q(\tfrac{1}{2}(n+p-1), \tfrac{1}{2}(N-n))\,|\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)} \\ &\qquad \times {}_1F_1^{(q)}\Big(\begin{matrix}(n+p-1)/2\\(N+p-1)/2\end{matrix}; -\boldsymbol{M}_1\boldsymbol{T}_{22}\Big), \end{aligned} \tag{3.27}$$

where the last equality follows from (3.15).

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

$$\begin{aligned} & f_{\widehat{\boldsymbol{\Delta}}_{12}|\widehat{\boldsymbol{\Delta}}_{22}=\boldsymbol{T}_{22}}(\boldsymbol{T}_{12}) \\ &\quad = (2\pi)^{-pq/2} 2^{-q(N-1)/2} N^{q(N+p-1)/2} \frac{B_q(\frac{1}{2}(n+p-1), \frac{1}{2}(N-n))}{\Gamma_q(\frac{1}{2}(n-1))\Gamma_q(\frac{1}{2}(N-n))} \\ &\qquad \times |\boldsymbol{\Delta}_{11}|^{-q/2} |\boldsymbol{\Delta}_{22}|^{-(N-1)/2} \exp(-\tfrac{1}{2}\mathrm{tr}\, N\boldsymbol{\Delta}_{22}^{-1}\boldsymbol{T}_{22})\ |\boldsymbol{T}_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)} \\ &\qquad \times {}_1F_1{}^{(q)}\Big(\begin{matrix} (n+p-1)/2 \\ (N+p-1)/2 \end{matrix}; -\tfrac{1}{2}N(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}\Big), \end{aligned} \tag{3.28}$$

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

$$\begin{aligned} & {}_1F_1{}^{(q)}\Big(\begin{matrix} (n+p-1)/2 \\ (N+p-1)/2 \end{matrix}; -\tfrac{1}{2}N(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}\Big) \\ &\qquad\qquad = {}_1F_1{}^{(p)}\Big(\begin{matrix} (n+p-1)/2 \\ (N+p-1)/2 \end{matrix}; -\tfrac{1}{2}N\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})\boldsymbol{T}_{22}(\boldsymbol{T}_{12}-\boldsymbol{\Delta}_{12})'\Big). \end{aligned}$$

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

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

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

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

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

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

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

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

which equals the stated formula. □

Remark 3.11. In practical situations in which the density function of $\widehat{\boldsymbol{\Delta}}$ is to be integrated over subsets of the space of positive definite matrices, we recommend that the saddlepoint approximations of Butler and Wood [7] be utilized. These approximations are as follows. Let $\boldsymbol{T}$ be a positive definite symmetric $p \times p$ matrix with eigenvalues $t_1, \ldots, t_p$. For $a < b$, define

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

$i = 1, \ldots, p$; it may be verified that $\widehat{s}_i$ is the unique solution in $(0, 1)$ of the quadratic equation $t_i s^2 - (t_i - b)s - a = 0$. Define

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

For $a, b - a > (p - 1)/2$ the *raw Laplace approximation* to ${}_1F_1{}^{(p)}\Big(\begin{smallmatrix} a \\ b \end{smallmatrix}; \boldsymbol{T}\Big)$, the confluent hypergeometric function of matrix argument, is

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

and the *calibrated Laplace approximation* is

$$\begin{aligned} {}_1\widehat{F}_1{}^{(p)}\left(\begin{matrix} a \\ b \end{matrix}; \boldsymbol{T}\right) &= {}_1\widetilde{F}_1{}^{(p)}\left(\begin{matrix} a \\ b \end{matrix}; \boldsymbol{T}\right) \Big/ {}_1\widetilde{F}_1{}^{(p)}\left(\begin{matrix} a \\ b \end{matrix}; \boldsymbol{0}\right) \\ &= b^{pb - p(p+1)/4}\, \widehat{J}_{1,1}^{-1/2} \prod_{i=1}^{p} \Big[\left(\frac{\widehat{s}_i}{a}\right)^{a} \left(\frac{1 - \hat{s}_i}{b - a}\right)^{b-a} e^{t_i \widehat{s}_i} \Big], \end{aligned} \tag{3.30}$$

where

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

Both the raw and calibrated Laplace approximations satisfy the reduction property (3.16) and the Kummer formula (3.17). Noting that each approximation involves only elementary functions of the s_j and t_j we recommend that, in the calculation of probabilities involving the eigenvalues of $\widehat{\boldsymbol{\Delta}}$, the hypergeometric function ${}_1F_1{}^{(p)}$ be approximated by (3.29) or (3.30).

4 Tests of hypotheses about $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$

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 specified positive definite matrix, on the basis of a monotone sample. Hao and Krishnamoorthy [16] showed that the likelihood ratio test statistic is

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

Because this statistic is not unbiased in the case of complete samples, Hao and Krishnamoorthy [16] modified λ_1 in the usual way, replacing sample sizes by degrees of freedom to obtain

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

and they derived an approximation to the asymptotic null distribution of this statistic. We shall prove that a sufficient condition for λ_2 to be unbiased is that $|\boldsymbol{\Sigma}_{11}| \le 1$. Since λ_2 might not always be unbiased, we propose a new statistic,

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

and establish that it is always unbiased. The crucial difference between λ_2 and λ_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 λ_3 to be unbiased.

We now calculate the null moments of λ_3, identify its exact null distribution, derive approximations to its null 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 moment of λ_3 is*

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

Proof. Under H_a, we apply invariance arguments to assume, without loss of generality, that $\boldsymbol{\Sigma}$ is diagonal ([16], 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 $\boldsymbol{A}_{22,N} \sim \mathrm{W}_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{A}_{11\cdot 2,n} \sim \mathrm{W}_p(n-q-1, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21} \sim \mathrm{W}_p(q, \boldsymbol{\Sigma}_{11})$. Therefore

$$\begin{aligned}\lambda_3 \stackrel{\mathcal{L}}{=} e_{p,q,n,N}\, & |\boldsymbol{W}_1|^{(N-1)/2} \exp\big(-\tfrac{1}{2}\mathrm{tr}\, \boldsymbol{W}_1\big) \\ & \times |\boldsymbol{W}_2|^{(n-q-1)/2} \exp\big(-\tfrac{1}{2}\mathrm{tr}\, \boldsymbol{W}_2\big)\, |\boldsymbol{W}_3|^{q/2} \exp\big(-\tfrac{1}{2}\mathrm{tr}\, \boldsymbol{W}_3\big),\end{aligned} \tag{4.5}$$

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

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

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

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

On writing each determinant in (4.4) as a product of its eigenvalues, we obtain a stochastic representation for the distribution of λ_3 as a product of independent random variables. We state this result explicitly in the null case. Under H_0, by applying invariance arguments, we may assume without loss of generality that $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_{p+q}$.

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

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

where Q_0 and all $Q_{j,k}$ are mutually independent, $Q_0 \sim \chi^2_{\frac{1}{2}q(q-1)+p(p-1)}$; $Q_{j,1} \sim \chi^2_{N-j}$, $j = 1,\ldots,q$; $Q_{j,2} \sim \chi^2_{n-q-j}$, 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 λ_3, *viz.*,

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

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

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

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 if that $\boldsymbol{W} \sim W_d(a, \boldsymbol{I}_d)$ then

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

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

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

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 $Q_0 = Q_{0,1} + Q_{0,2} + Q_{0,3} \sim \chi^2_{\frac{1}{2}q(q-1)+p(p-1)}$, we obtain (4.7). □

A complete treatment of the exact distribution of λ_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 λ_3, we apply the results of [28], p. 359 (see also [16], p. 68) to each of the three terms in the representation of λ_3 as a product of independent random entities in (4.3) or (4.5). Under H_0, the asymptotic distribution of λ_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

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

Let $\rho_{(1)}$ and $\rho_{(3)}$ denote the smallest and largest of ρ_1, ρ_2, ρ_3, respectively. On applying to the right-hand side of (4.8) the results of Kotz, *et al.*, [21], 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 [21], and additional terms in our approximation may be obtained accordingly from their series. Alternatively, by applying the results of [21], 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 also noteworthy for, in the case of small sample sizes, those approximations generally are superior to standard asymptotic approximations. Let

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

denote the cumulant-generating function of the right-hand side of (4.8). Applying the results of Kuonen [23], eq. (3) we obtain

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

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

We remark also that although the above results constitute a saddlepoint approximation only to the asymptotic distribution of λ_3, the methods of Booth, *et al.* [6] may be applied to obtain a saddlepoint approximation to the exact distribution of λ_3.

We consider next the unbiasedness of λ_2 and λ_3. The proof of the following result follows the argument of Sugiura and Nagao [30] (see [28], p. 367).

Theorem 4.3. *The statistic λ_3 is unbiased. Further, if $|\mathbf{\Sigma}_{11}| \le 1$ then λ_2 is unbiased.*

Proof. By (4.5), a critical region of size α for the test using λ_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 $\boldsymbol{W}_1 \sim \mathrm{W}_q(N-1, \mathbf{\Sigma}_{22})$, $\boldsymbol{W}_2 \sim \mathrm{W}_p(n-q-1, \mathbf{\Sigma}_{11\cdot 2})$, and $\boldsymbol{W}_3 \sim \mathrm{W}_p(q, \mathbf{\Sigma}_{11})$ are mutually independent, and the constant k_α is such that $P(\lambda_3 \in \mathfrak{C}_3 | H_0) = \alpha$. Denote by $c_q(N-1, \mathbf{\Sigma}_{22})$, $c_p(n-q-1, \mathbf{\Sigma}_{11\cdot 2})$, and $c_p(q, \mathbf{\Sigma}_{11})$ 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

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

Making the transformation

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

in this integral, we obtain

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

where

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

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

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

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

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

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

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

$\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3 > \boldsymbol{0}$. 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

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

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

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

Therefore $P(\lambda_3 \in \mathfrak{C}_3 | H_a) - P(\lambda_3 \in \mathfrak{C}_3 | H_0) > 0$, which proves that λ_3 is unbiased.

In the case of λ_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 α and k_α 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 $\mathfrak{C}_2^* = \{(\widetilde{\boldsymbol{W}}_1, \widetilde{\boldsymbol{W}}_2, \widetilde{\boldsymbol{W}}_3) : (\boldsymbol{\Sigma}_{22}^{1/2} \widetilde{\boldsymbol{W}}_1 \boldsymbol{\Sigma}_{22}^{1/2}, \boldsymbol{\Sigma}_{11\cdot 2}^{1/2} \widetilde{\boldsymbol{W}}_1 \boldsymbol{\Sigma}_{11\cdot 2}^{1/2}, \boldsymbol{\Sigma}_{11}^{1/2} \widetilde{\boldsymbol{W}}_1 \boldsymbol{\Sigma}_{11}^{1/2}) \in \mathfrak{C}_2\}$. By an argument analogous to that given for λ_3, we obtain

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

For $|\boldsymbol{\Sigma}_{11}|^{-qp/2} - 1 \ge 0$, equivalently $|\boldsymbol{\Sigma}_{11}| \le 1$, we see that λ_2 is unbiased. □

Next, we show that the statistic λ_1 in (4.1) is not unbiased for all n and N. Here, the proof follows the classical approach of Das Gupta [10] (see [28], p. 357).

Proposition 4.4. *For testing $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ against $H_a : \boldsymbol{\Sigma} \ne \boldsymbol{\Sigma}_0$, the likelihood ratio test statistic λ_1 in (4.1) is not unbiased.*

Proof. As before, we shall assume without loss of generality that $\boldsymbol{\Sigma}$ is diagonal, say, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{1,1}, \dots, \sigma_{p+q,p+q})$. 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 with $\boldsymbol{A}_{22,N} \sim \mathrm{W}_q(N-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{A}_{11\cdot 2,n} \sim \mathrm{W}_p(n-q-1, \boldsymbol{\Sigma}_{11})$, and $\boldsymbol{A}_{12} \boldsymbol{A}_{22,n}^{-1} \boldsymbol{A}_{21} \sim \mathrm{W}_p(q, \boldsymbol{\Sigma}_{11})$. By (4.1),

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

The rest of the proof now proceeds as in the classical case. The random variables $(\boldsymbol{A}_{22,N})_{jj}$, $j = p+1, \dots, p+q$, 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 $(\boldsymbol{A}_{22,N})_{jj} / \sigma_{jj} \sim \chi^2_{N-1}$. By [28], p. 356, Lemma 8.4.3 there exists $\sigma_{p+q}^* \in (1, N/(N-1))$ such that, for any $c > 0$,

$$\begin{aligned} &P\Big((\boldsymbol{A}_{22,N})_{p+q,p+q}^{N/2} \exp\big(-\tfrac{1}{2} (\boldsymbol{A}_{22,N})_{p+q,p+q}\big) \ge k | \sigma_{p+q} = 1 \Big) \\ &\qquad < P\Big((\boldsymbol{A}_{22,N})_{p+q,p+q}^{N/2} \exp\big(-\tfrac{1}{2} (\boldsymbol{A}_{22,N})_{p+q,p+q}\big) \ge c | \sigma_{p+q} = \sigma_{p+q}^* \Big). \end{aligned}$$

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

As in the classical case, we can obtain a result which is stronger than the unbiasedness property of λ_3 (see [28], p. 358); however, we also note that it does not provide the unbiasedness property of λ_2 which was deduced in Theorem 4.3. The proof of the following result is similar to the classical case.

Theorem 4.5. *For $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{p+q,p+q})$, the power function of the modified likelihood ratio statistic λ_3 increases monotonically with $|\sigma_{jj} - 1|$, $1 \leq j \leq p+q$.*

Proof. By [28], p. 357, Corollary 8.4.4, $P\big((\boldsymbol{A}_{22,N})_{p+q,p+q}^{(N-1)/2} \exp\big(-\tfrac{1}{2}(\boldsymbol{A}_{22,N})_{p+q,p+q}\big) \leq k|\sigma_{p+q}\big)$ 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 that used in the previous result. $\square$

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 [16], eq. (4.1) showed that the likelihood ratio test statistic is

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

where λ_1 is given in (4.1). By invariance arguments we may assume, without loss of generality, that $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = (\boldsymbol{0}, \boldsymbol{I}_{p+q})$ and that $\boldsymbol{\Sigma}$ is diagonal under H_a. Substituting (4.1) into (4.11), we obtain

$$\begin{aligned} \lambda_4 &= (e/N)^{Nq/2}\, |\boldsymbol{A}_{22,N}|^{N/2} \exp\big(-\tfrac{1}{2}\mathrm{tr}\, \boldsymbol{A}_{22,N}\big) \\ &\quad \times (e/n)^{np/2}\, |\boldsymbol{A}_{11\cdot 2,n}|^{n/2} \exp\big(-\tfrac{1}{2}\mathrm{tr}\, \boldsymbol{A}_{11\cdot 2,n}\big) \\ &\quad \times \exp\big(-\tfrac{1}{2}\mathrm{tr}\, \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}\big) \exp\big(-\tfrac{1}{2}(n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}} + N\bar{\boldsymbol{Y}}'\bar{\boldsymbol{Y}}\big). \end{aligned}$$

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

To identify the exact null distribution of λ_4 and investigate its unbiasedness properties, we proceed as in the case of λ_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 λ_4 for testing $H_0 : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\boldsymbol{0}, \boldsymbol{I}_{p+q})$ against $H_a : (\boldsymbol{\mu}, \boldsymbol{\Sigma}) \neq (\boldsymbol{0}, \boldsymbol{I}_{p+q})$ is unbiased. For $h = 0, 1, 2, \ldots$ the h-th moment of λ_4 is*

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

and, under H_0,

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

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 χ^2 variables are mutually independent.

We remark that, in the non-null case, the distribution of λ_4 may also be obtained from (4.12); the final result is similar to (4.13) but is more cumbersome to state, involving noncentral chi-square random variables.

4.3 The sphericity test

Consider the problem of testing sphericity, in which the null hypothesis is $H_0 : \mathbf{\Sigma} = \sigma^2 \boldsymbol{I}_{p+q}$ and the alternative hypothesis is $H_a : \mathbf{\Sigma} \neq \sigma^2 \boldsymbol{I}_{p+q}$, where $\sigma^2 > 0$ is unspecified. Bhargava [4], Section 6 derived the likelihood ratio test statistic for a problem more general than the sphericity problem 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 the classical approach (see [3], p. 431; [28], p. 333), and we derive 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 σ^2 are, respectively, $\widehat{\boldsymbol{\mu}}_{10} = \bar{\boldsymbol{X}}$, $\widehat{\boldsymbol{\mu}}_{20} = \bar{\boldsymbol{Y}}$, and

$$\begin{aligned}\widehat{\sigma}_0^2 &= \frac{1}{np+Nq}\Big[\sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}})'(\boldsymbol{X}_j - \bar{\boldsymbol{X}}) + \sum_{j=1}^{N}(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})'(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})\\ &= \frac{1}{np+Nq}\big[\operatorname{tr}\boldsymbol{A}_{11} + \operatorname{tr}\boldsymbol{A}_{22,N}\big].\end{aligned}$$

Under H_a, the maximum likelihood estimators of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ 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

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

In the classical case ([3], 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 $N^{-1}\boldsymbol{A}_{22,N}$. Since $\boldsymbol{A}_{11} = \boldsymbol{A}_{11\cdot 2,n} + \boldsymbol{A}_{12}\boldsymbol{A}_{22,n}^{-1}\boldsymbol{A}_{21}$ then

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

By the weighted arithmetic-geometric mean inequality (Marshall and Olkin [26], p. 455),

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

Therefore, since $\mathcal{G}_j \leq \mathcal{A}_j$, $j = 1, 2$, then

$$\lambda_5^2 \leq \frac{\mathcal{G}_1^{np}\,\mathcal{G}_2^{Nq}}{\big((\mathcal{A}_1^{np}\mathcal{A}_2^{Nq})^{1/(np+Nq)}\big)^{np+Nq}} \equiv \Big(\frac{\mathcal{G}_1}{\mathcal{A}_1}\Big)^{np}\Big(\frac{\mathcal{G}_2}{\mathcal{A}_2}\Big)^{Nq} \leq 1.$$

We remark also that (4.15) shows how λ_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 λ_5 is*

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

Under H_0,

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

where $(U_1, \ldots, U_{p+q}) \sim SD_{p+q}\big(\underbrace{\tfrac{1}{2}(n-q-1), \ldots, \tfrac{1}{2}(n-q-1)}_{p}, \underbrace{\tfrac{1}{2}(N-1), \ldots, \tfrac{1}{2}(N-1)}_{q}\big)$, a singular Dirichlet distribution; $U_{1j} \sim \beta(\frac{1}{2}(n-q-i+1), \frac{1}{2}(i-1))$, $2 \leq j \leq p$; $U_{2j} \sim \beta(\frac{1}{2}(N-i+1), \frac{1}{2}(i-1))$, $2 \leq j \leq q$; 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 $\sigma^2 = 1$, 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),

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

where $\boldsymbol{W}_1 \sim \mathrm{W}_p(n-q-1, \boldsymbol{I}_p)$, $\boldsymbol{W}_2 \sim \mathrm{W}_p(q, \boldsymbol{I}_p)$, and $\boldsymbol{W}_3 \sim \mathrm{W}_q(N-1, \boldsymbol{I}_q)$ 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 $\widetilde{\boldsymbol{W}}_1 \sim \mathrm{W}_p(nh+n-q-1, \boldsymbol{I}_p)$. 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 $\widetilde{\boldsymbol{W}}_3 \sim \mathrm{W}_q(Nh+N-1, \boldsymbol{I}_q)$. 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 [28, 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 $\operatorname{tr} \widetilde{\boldsymbol{W}}_3 \sim \chi^2_{(Nh+N-1)q}$, 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

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

$\delta < r$, 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

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

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,

$$\begin{aligned} E(\lambda_5^h) &= \frac{\Gamma_p\big(\frac{1}{2}(nh+n-q-1)\big)\,\Gamma^p\big(\frac{1}{2}(n-q-1)\big)}{\Gamma_p\big(\frac{1}{2}(n-q-1)\big)\,\Gamma^p\big(\frac{1}{2}(nh+n-q-1)\big)} \\ &\quad \times \frac{\Gamma_q\big(\frac{1}{2}(Nh+N-1)\big)\,\Gamma^q\big(\frac{1}{2}(N-1)\big)}{\Gamma_q\big(\frac{1}{2}(N-1)\big)\,\Gamma^q\big(\frac{1}{2}(Nh+N-1)\big)} \\ &\quad \times \frac{\Gamma\big(\frac{1}{2}((n-1)p+(N-1)q)\big)}{\Gamma\big(\frac{1}{2}((n-1)p+(N-1)q)+\frac{1}{2}(np+Nq)h\big)} \\ &\quad \times \frac{\Gamma^p\big(\frac{1}{2}(nh+n-q-1)\big)\,\Gamma^q\big(\frac{1}{2}(Nh+N-1)\big)}{\Gamma^p\big(\frac{1}{2}(n-q-1)\big)\,\Gamma^q\big(\frac{1}{2}(N-1)\big)}. \end{aligned} \tag{4.20}$$

The first ratio in this product is the h-th moment of a classical sphericity statistic; see [3], p. 435, eq. (16), from which we deduce that the first ratio is the h-th moment of a product of powers of independent beta-distributed 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 [26], p. 307, eq. (11)), it is straightforward to verify 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 λ_5 is unbiased; in particular, the methods of Gleser [15] or Sugiura and Nagao [30] seem inapplicable to this problem. On the other hand, the non-null distribution of λ_5 can be obtained using the methods given here, suitably generalizing the approach provided by Muirhead [28], p. 339 ff.

4.4 Testing independence between subsets of the variables

Consider the problem of testing the null hypothesis $H_0 : \boldsymbol{\Sigma}_{12} = \boldsymbol{0}$ against the alternative hypothesis $H_a : \boldsymbol{\Sigma}_{12} \neq \boldsymbol{0}$ with the monotone sample (1.1). Eaton and Kariya [12] showed that the likelihood ratio test statistic is based only on the complete observations $(\boldsymbol{X}_1, \boldsymbol{Y}_1), \ldots, (\boldsymbol{X}_n, \boldsymbol{Y}_n)$ and ignores the incomplete observations $\boldsymbol{Y}_{n+1}, \ldots, \boldsymbol{Y}_N$.

With the data (1.1), Eaton and Kariya [12] proved that, among the class of affinely invariant test procedures, the test that rejects H_0 for small values of

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

is a locally most powerful invariant test, 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. [9, 25, 29]. To date, the distribution theory of λ_6 has not been explored, and our calculations thereof did not lead to a useful form for the distribution. On the other hand, by omitting the term np^{-1}, we obtain the modified statistic,

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

The statistic λ_7 will not generally enjoy the same optimality properties as λ_6. However, for $n \ge p$ we have $\lambda_6 \le \lambda_7$, in which case if H_0 is rejected for small values of λ_7 then H_0 also is rejected by λ_6. Moreover, λ_7 has a null distribution which is simpler than that of λ_6. Indeed, with $\boldsymbol{W}_1 = \boldsymbol{A}_{22,n} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$ and $\boldsymbol{W}_2 = \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$, we have $\lambda_7 = \mathrm{tr}\, \boldsymbol{W}_1(\boldsymbol{W}_1 + \boldsymbol{W}_2 + \boldsymbol{B}_1)^{-1}$. Applying [8], Proposition 2.2(iii), we obtain that, under H_0, $\boldsymbol{W}_1 \sim \mathrm{W}_q(n-p-1, \boldsymbol{\Sigma}_{22})$, $\boldsymbol{W}_2 \sim \mathrm{W}_q(p, \boldsymbol{\Sigma}_{22})$; further, $\boldsymbol{B}_1 \sim \mathrm{W}_q(N-n-1, \boldsymbol{\Sigma}_{22})$; and $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, and $\boldsymbol{B}_1$ are independent; equivalently, $\lambda_7 \overset{\mathcal{L}}{=} \mathrm{tr}\, \boldsymbol{W}_1(\boldsymbol{W}_1 + \boldsymbol{W}_3)^{-1}$, where $\boldsymbol{W}_1$ is as before and $\boldsymbol{W}_3 = \boldsymbol{W}_2 + \boldsymbol{B}_1 \sim \mathrm{W}_q(N-n+p-1, \boldsymbol{\Sigma}_{22})$. It is now clear that, under H_0, the statistic λ_7 is exactly of the form of the Bartlett-Nanda-Pillai criterion in MANOVA ([3], Section 8.6.3; [28], Section 10.6.3] and its null distribution may be derived accordingly.

Acknowledgments. The authors are grateful to the referees and an associate editor for providing valuable comments on an earlier version of this article.

References

[1] T. W. Anderson, Maximum likelihood estimators for a multivariate normal distribution when some observations are missing. *J. Amer. Statist. Assoc.*, **52** (1957), 200–203.

[2] T. W. Anderson and I. Olkin, Maximum likelihood estimation of the parameters of a multivariate normal distribution. *Linear Algebra Appl.*, **70** (1985), 147–171.

[3] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, third edition, Wiley, New York, 2003.

[4] R. Bhargava, Multivariate tests of hypotheses with incomplete data, Technical Report No. 3, Applied Math. and Statistics Laboratories, Stanford University, Stanford, CA, 1962.

[5] 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), 329–339.

[6] J. G. Booth, R. W. Butler, S. Huzurbazar, and A. T. A. Wood, Saddlepoint approximations for P-values of some tests of covariance matrices, *J. Statist. Comput. Simulation*, **53** (1995), 165–180.

[7] R. W. Butler and A. T. A. Wood, Laplace approximations for hypergeometric functions with matrix argument, *Ann. Statist.*, **30** (2002), 1155–1177.

[8] W.-Y. Chang and D. St. P. Richards, Finite-sample inference with monotone incomplete multivariate normal data, I, 2008, preprint, Penn State University.

[9] R.-J. Chou and W. D. Lo, On the local minimaxity of a test of independence in incomplete samples, *Ann. Inst. Statist. Math.*, **38** (1986), 495–502.

[10] S. Das Gupta, Properties of power functions of some tests concerning dispersion matrices of multivariate normal distributions, *Ann. Math. Statist.*, **40** (1969), 697–701.

[11] M. L. Eaton, *Multivariate Statistics. A Vector Space Approach*, Wiley, New York, 1983.

[12] M. L. Eaton and T. Kariya, Multivariate tests with incomplete data. *Ann. Statist.*, **11** (1983), 654–665.

[13] H. Fujisawa, A note on the maximum likelihood estimators for multivariate normal distribution with monotone data, *Comm. Statist. Theory Methods*, **24** (1995), 1377–1382.

[14] M. A. Giguère and G. P. H. Styan, Multivariate normal estimation with missing data on several variates, In: *Trans. Seventh Prague Conference on Information Theory, Statistical Decision Functions*, pp. 129–139, Academia, Prague, 1978.

[15] L. J. Gleser, A note on the sphericity test, *Ann. Math. Statist.*, **37** (1966), 464–467.

[16] J. Hao and K. Krishnamoorthy, Inferences on a normal covariance matrix and generalized variance with monotone missing data, *J. Multivariate Anal.*, **78** (2001), 62–82.

[17] C. S. Herz, Bessel functions of matrix argument, *Ann. Math.*, **61** (1955), 474–523.

[18] K. G. Jinadasa and D. S. Tracy, Maximum likelihood estimation for multivariate normal distribution with monotone sample, *Comm. Statist. Theory Methods*, **21** (1992), 41–50.

[19] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, fifth edition, Prentice-Hall, Englewood Cliffs, NJ, 2002.

[20] T. Kanda and Y. Fujikoshi, Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data, *Amer. J. Math. Management Sci.*, **18** (1988), 161–190.

[21] S. Kotz, N. L. Johnson, and D. W. Boyd, Series representations of distributions of quadratic forms in normal variables. I. Central case, *Ann. Math. Statist.*, **38** (1967), 823–837.

[22] K. Krishnamoorthy, Estimation of normal covariance and precision matrices with incomplete data, *Comm. Statist. Theory Methods*, **20** (1991), 757–770.

[23] D. Kuonen, Saddlepoint approximations for distributions of quadratic forms in normal variables, *Biometrika*, **86** (1999), 929–935.

[24] R. J. A. Little and D. B. Rubin, *Statistical Analysis With Missing Data*, second edition, Wiley-Interscience, Hoboken, NJ, 2002.

[25] W. D. Lo and R.-J. Chou, Test of independence in incomplete samples. *Bull. Inst. Math. Acad. Sinica*, **14** (1986), 117–127.

[26] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

[27] 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.

[28] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

[29] M. S. Srivastava and S. R. Chakravorti, Asymptotic distributions of two test statistics for testing independence with missing data, *Comm. Statist. A—Theory Methods*, **15** (1986), 571–588.

[30] N. Sugiura and H. Nagao, Unbiasedness of some test criteria for the equality of one or two covariance matrices, *Ann. Math. Statist.*, **39** (1968), 1686–1692.