Journal of Multivariate Analysis 100 (2009) 473–496

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Third-order power comparisons for a class of tests for multivariate linear hypothesis under general distributions

Yoshihide Kakizawa

Faculty of Economics, Hokkaido University, Nishi 7, Kita 9, Kita-ku, Sapporo 060–0809, Japan

ARTICLE INFO

Article history:
Received 20 September 2007
Available online 10 June 2008

AMS 1991 subject classifications:
primary 62H10
62H15
secondary 62E20
62J05
62J10

Keywords:
Asymptotic expansion
Multivariate linear regression model
GMANOVA model
Bartlett's type adjustment
Cornish–Fisher's type size adjustment
Local power
Nonnormality
Differential operator

ABSTRACT

The purpose of this paper is, in multivariate linear regression model (Part I) and GMANOVA model (Part II), to investigate the effect of nonnormality upon the nonnull distributions of some multivariate test statistics under normality. It is shown that whatever the underlying distributions, the difference of local powers up to order N^{-1} after either Bartlett's type adjustment or Cornish–Fisher's type size adjustment under nonnormality coincides with that in Anderson [An Introduction to Multivariate Statistical Analysis, 2nd ed. and 3rd ed., Wiley, New York, 1984, 2003] under normality. The derivation of asymptotic expansions is based on the differential operator associated with the multivariate linear regression model under general distributions. The performance of higher-order results in finite samples, including monotone Bartlett's type adjustment and monotone Cornish–Fisher's type size adjustment, is examined using simulation studies.

1. Introduction

The study of Bartlett adjustability and higher-order power for several asymptotic χ^2 tests based on not only the likelihood under model specification but also Owen's empirical likelihood approach (we refer to his book [51]) and its generalization to the empirical discrepancy approach has received considerable attention — see e.g. Barndorff-Nielsen and Cox [2], Bickel and Ghosh [7], Chandra and Mukerjee [11], Cordeiro and Ferrari [16], DiCiccio et al. [18], Taniguchi [61], Mukerjee [48–50], Rao and Mukerjee [55,56], Ghosh and Mukerjee [27], Bravo [8,9], Chen and Cui [14] and the references therein (there is a vast literature on these topics for various statistical settings).

The present paper investigates, as a continuation of Kakizawa and Iwashita [37], the problem of testing a general linear hypothesis in multivariate linear regression model (Part I) and GMANOVA model (Part II) with nonparametric error distribution, where "nonparametric" means that the error distribution is not specified by a finite dimensional parameter. Although our setting is different from Dufour and Khalaf [19], an important feature in the statistical inference on the mean structure of the multivariate model is the fact that several test statistics derived under normality (including the likelihood ratio (LR) criterion, Lawley–Hotelling's trace and Bartlett–Nanda–Pillai's trace) are all functions of the eigenvalues of a characteristic determinantal equation which involves the restricted and unrestricted residual sum of squares matrices, hence they are easily shown to have the same limiting noncentral chi-square distribution under a local alternative. It may,

E-mail address: kakizawa@econ.hokudai.ac.jp.

0047-259X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2008.06.002

therefore, be of interest to study higher-order asymptotic comparison of the classical multivariate tests (actually, little is known about their local power properties under nonnormality).

1.1. Multivariate linear regression model

We suppose that an $N \times p$ matrix $\mathbf{Y}$ consists of N independent observations $\mathbf{y}_1, \ldots, \mathbf{y}_N$ on p variables, where $\mathbf{Y}' = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$. The multivariate linear regression model is

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{U},$$

where $\mathbf{X}$ is an $N \times q$ non-random design matrix of rank $q\ (< N)$, $\boldsymbol{\Theta}$ is a $q \times p$ unknown regression matrix, and $\mathbf{U}$ is an $N \times p$ unobservable random matrix with $\mathbf{U}' = [\mathbf{u}_1, \ldots, \mathbf{u}_N]$. It is assumed that each $p \times 1$ vector $\mathbf{u}_i$ is independently and identically distributed with mean vector $\mathbf{0}$ and positive definite covariance matrix $\boldsymbol{\Sigma}$. For testing a linear hypothesis $H : \mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$ ($r \times p$ zero matrix), where $\mathbf{B}$ is an $r \times q$ known matrix of rank $r\ (\le q)$, let $\mathbf{H}_Y$ and $\mathbf{E}_Y$ be the variation matrices due to the hypothesis and the error

$$\mathbf{H}_Y = \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\{\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\}^{-1}\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y},$$
$$\mathbf{E}_Y = \mathbf{Y}'\{\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}\mathbf{Y},$$

respectively. Then, the following three criteria have been used under normality:

(i) Likelihood ratio (or Wilks' Λ) $T_{LR} = -(N-q)\log(|\mathbf{E}_Y|/|\mathbf{E}_Y + \mathbf{H}_Y|)$,
(ii) Lawley–Hotelling's trace (T_0^2) $T_{LH} = (N-q)\mathrm{tr}(\mathbf{H}_Y\mathbf{E}_Y^{-1})$, and
(iii) Bartlett–Nanda–Pillai's trace $T_{BNP} = (N-q)\mathrm{tr}[\mathbf{H}_Y(\mathbf{E}_Y + \mathbf{H}_Y)^{-1}]$.

We notice that the problem of testing the hypothesis $\mathbf{B}_1\boldsymbol{\Theta}\mathbf{B}_2 = \mathbf{B}_0$, where $\mathbf{B}_1$ is a $r \times q$ known matrix of rank $r\ (\le q)$, $\mathbf{B}_2$ is a $p \times s$ known matrix of rank $s\ (\le p)$ and $\mathbf{B}_0$ is an $r \times s$ known matrix, reduces to that of the null hypothesis $\mathbf{B}_1\widetilde{\boldsymbol{\Theta}} = \mathbf{O}_{r,s}$ in the model $\widetilde{\mathbf{Y}} = \mathbf{X}\widetilde{\boldsymbol{\Theta}} + \widetilde{\mathbf{U}}$, by letting $\widetilde{\mathbf{Y}} = \mathbf{Y}\mathbf{B}_2 - \mathbf{X}\mathbf{B}_1^{-}\mathbf{B}_0$, $\widetilde{\mathbf{U}} = \mathbf{U}\mathbf{B}_2$ ($N \times s$ matrix) and $\widetilde{\boldsymbol{\Theta}} = \boldsymbol{\Theta}\mathbf{B}_2$ ($q \times s$ matrix), where $\mathbf{B}_1^{-}$ is a generalized inverse matrix of $\mathbf{B}_1$ (e.g. [30, Chapter 9]). Berndt and Savin [3] showed that Lawley–Hotelling's trace and Bartlett–Nanda–Pillai's trace are viewed as the Wald test and the Lagrange multiplier (LM) test, respectively, under normality. Further, the three criteria T_{LR}, T_{LH} and T_{BNP} are special cases of the generalized test statistic

$$T_\psi = (N-q)\sum_{j=1}^{p}\psi(\lambda_{Y,j})$$

(see [23]) by letting $\psi(x) = \log(1+x)$, x, $1 - 1/(1+x)$, respectively, where $\lambda_{Y,1}, \ldots, \lambda_{Y,p} \ge 0$ are eigenvalues of $\mathbf{H}_Y\mathbf{E}_Y^{-1}$. These three functions are usually generalized as $\psi(x;\gamma) = \{(1+x)^\gamma - 1\}/\gamma$ for any $\gamma \in \mathbf{R}$.

Asymptotic expansions under normality have been extensively studied (see e.g. [47, Chapter 10], [1, Chapter 8] and references theirin). Under nonnormality, Wakaki et al. [62] gave an asymptotic expansion for the null distribution of $T = T_{LR}, T_{LH}, T_{BNP}$, following Kano [39] and Fujikoshi [24–26]. The purpose of this paper is, as an extended study of Kakizawa and Iwashita [37], to obtain an asymptotic expansion for the nonnull distribution of T_ψ up to order N^{-1} and then compare their local powers after either Bartlett's type adjustment or Cornish–Fisher's type size adjustment under nonnormality. Unlike Fujikoshi [26] and Wakaki et al. [62], our derivation of asymptotic expansion is based on the differential operator method developed by Kakizawa and Iwashita [36,37] and Kakizawa [33–35].

In what follows, for any positive definite matrix $\mathbf{T}$ of $n \times n$, $\mathbf{T}^{-1/2}$ denotes the symmetric square root of $\mathbf{T}^{-1}$. That is, given a spectral decomposition $\mathbf{O}_T\boldsymbol{\Lambda}_T\mathbf{O}_T'$ of $\mathbf{T}$, where $\mathbf{O}_T$ is an orthogonal matrix of $n \times n$ and $\boldsymbol{\Lambda}_T = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_1, \ldots, \lambda_n > 0$ being eigenvalues of $\mathbf{T}$, we always set $\mathbf{T}^{-1/2} = \mathbf{O}_T\boldsymbol{\Lambda}_T^{-1/2}\mathbf{O}_T'$, which is the inverse matrix of $\mathbf{T}^{1/2} = \mathbf{O}_T\boldsymbol{\Lambda}_T^{1/2}\mathbf{O}_T'$, where $\boldsymbol{\Lambda}_T^{1/2} = \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})$ and $\boldsymbol{\Lambda}_T^{-1/2} = \mathrm{diag}(\lambda_1^{-1/2}, \ldots, \lambda_n^{-1/2})$. The Kronecker delta is denoted by $\delta_{a_1a_2}$, that is, $\delta_{a_1a_2} = 1$ iff $a_1 = a_2$, and 0 otherwise.

2. Preliminary results

Throughout this paper we set down the following assumptions:

(C_1) The $\mathbf{u}_i$'s are independently distributed according to a common p-variate distribution of $\mathbf{u} = (u_1, \ldots, u_p)'$ with mean vector $\mathbf{0}$, positive definite covariance matrix $\boldsymbol{\Sigma} = (\sigma_{jk})$ and υth order cumulant $Cum(u_{j_1}, \ldots, u_{j_\upsilon}) = \kappa_{j_1,\ldots,j_\upsilon}$ $(\upsilon \ge 3)$. Here and subsequently we use j, k, without or with suffixes, to denote indices, each such index running from 1 to p unless explicitly stated otherwise.

(C_2) (i) The bounded design matrix $\mathbf{X}$ of $N \times q$ satisfies $N^{-1}\mathbf{X}'\mathbf{X} = \mathbf{Q} + N^{-1}\mathbf{Q}_1 + o(N^{-1})$, where $\mathbf{Q}$ is a positive definite matrix of $q \times q$ and $\mathbf{Q}_1$ is a symmetric matrix of $q \times q$ (hence, there exists an integer $N_0 > q$, such that the smallest eigenvalue of $N^{-1}\mathbf{X}'\mathbf{X}$ is strictly positive for all $N \ge N_0$).
(ii) In addition to (i), writing

$$(\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}' = [\mathfrak{X}^{(1)}, \ldots, \mathfrak{X}^{(N)}] = (\mathfrak{X}_a^{(i)}) \quad (q \times N \text{ matrix}),$$

$\overline{\mathcal{X}}_{a_1\ldots a_s} = N^{-1}\sum_{i=1}^{N} \mathcal{X}_{a_1}^{(i)}\cdots\mathcal{X}_{a_s}^{(i)}$ ($s \in \mathbf{N}$; $a_1, \ldots, a_s \in \{1, \ldots, q\}$) has asymptotic representation $N^{1/2}\overline{\mathcal{X}}_{a_1} = w_{a_1} + \mathrm{o}(N^{-1/2})$, $N^{3/2}\overline{\mathcal{X}}_{a_1a_2a_3} = w_{a_1a_2a_3} + \mathrm{o}(N^{-1/2})$ and $N^2\overline{\mathcal{X}}_{a_1a_2a_3a_4} = w_{a_1a_2a_3a_4} + \mathrm{o}(1)$ (we have $N\overline{\mathcal{X}}_{a_1a_2} \equiv \sum_{i=1}^{N}\mathcal{X}_{a_1}^{(i)}\mathcal{X}_{a_2}^{(i)} = \delta_{a_1a_2}$ exactly and observe that $\max_{i=1,\ldots,N} N^{1/2}|\mathcal{X}_a^{(i)}|$ is uniformly bounded in $N \geq N_0$).

(C_3) The class of distributions of $\mathbf{u} = (u_1, \ldots, u_p)'$ is restricted to the distributions such that $\widetilde{\mathbf{u}} = (\mathbf{u}', \{\mathrm{vech}(\mathbf{uu}' - \boldsymbol{\Sigma})\}')'$ satisfies Cramér's condition

$$\limsup_{\|\xi\|\to\infty} |E[\exp(\mathrm{i}\boldsymbol{\xi}'\widetilde{\mathbf{u}})]| < 1 \quad (\text{write } \mathrm{i} = \sqrt{-1};\ \boldsymbol{\xi} \in \mathbf{R}^{p+p(p+1)/2}) \tag{1}$$

with a finite 8th absolute moment $E(\|\mathbf{u}\|^8) < \infty$.

(C_4) The third order derivative $\psi'''(x)$ of a nonnegative function $\psi(x)$ is continuous in a neighborhood $\mathcal{N}$ of $x = 0$ (we assume $\psi(0) = 0$ and $\psi'(0) = 1$). Write $\psi'' = \psi''(0)$.

Given an $r \times q$ matrix $\mathbf{B}$ of rank r, the solution space $\mathcal{V}_{\mathrm{B}}$ of a homogeneous linear system $\mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$ is a subspace of the linear space $\mathcal{R}^{q\times p}$ of all $q \times p$ matrices. We know $\dim(\mathcal{V}_{\mathrm{B}}) = p(q-r)$ (see [30, p143]). In what follows, fix $\boldsymbol{\Theta}_0 \in \mathcal{V}_{\mathrm{B}}$. We are now interested in deriving an asymptotic expansion of the nonnull distribution of T_ψ when the regression matrix is given by

$$\boldsymbol{\Theta} = \boldsymbol{\Theta}_0 + (\mathbf{X}'\mathbf{X})^{-1/2}\boldsymbol{\Theta}_\varepsilon \quad \text{with } (\mathbf{X}'\mathbf{X})^{-1/2}\boldsymbol{\Theta}_\varepsilon \notin \mathcal{V}_{\mathrm{B}} \tag{2}$$

(we always assume that $\boldsymbol{\Theta}_\varepsilon$ is a $q \times p$ matrix, independent of N). In that case, we have

$$\mathbf{H}_Y = \{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + \boldsymbol{\Theta}_\varepsilon'\}\dot{\mathbf{M}}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + \boldsymbol{\Theta}_\varepsilon'\}' = \widetilde{\mathbf{H}}_U \quad (\text{say}),$$

where

$$\dot{\mathbf{M}} = (\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{B}'\{\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\}^{-1}\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1/2} \tag{3}$$

($q \times q$ idempotent matrix of rank r). Note that $\mathbf{E}_Y/(N-q) = \widehat{\boldsymbol{\Sigma}}_Y$ (say) is the unbiased estimator of the covariance matrix $\boldsymbol{\Sigma}$, where

$$\widehat{\boldsymbol{\Sigma}}_Y = \frac{\mathbf{E}_Y}{N-q} = \frac{1}{N-q}\{\mathbf{U}'\mathbf{U} - \mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}\} = \widehat{\boldsymbol{\Sigma}}_U$$

may be positive definite with probability one if $N - q \geq p$ (see [20]), provided that under the distribution of $\mathbf{u}_1 \in \mathbf{R}^p$, every flat of dimension $p-1$ has probability zero (such a non-asymptotic result can be replaced by a higher-order one [6, Theorem 17.11]). Using these expressions, Lawley–Hotelling's trace $T_{LH} = \mathrm{tr}(\mathbf{H}_Y\widehat{\boldsymbol{\Sigma}}_Y^{-1})$ is written as

$$T_{LH} = [\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + \boldsymbol{\Theta}_\varepsilon'\}]'(\dot{\mathbf{M}} \otimes \widehat{\boldsymbol{\Sigma}}_U^{-1})\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + \boldsymbol{\Theta}_\varepsilon'\},$$

since $\mathrm{tr}(\mathbf{A}'\mathbf{BCD}') = \{\mathrm{vec}(\mathbf{A})\}'(\mathbf{D}\otimes\mathbf{B})\mathrm{vec}(\mathbf{C})$ (e.g. [30, p342]).

2.1. Asymptotic distribution

With

$$\mathbf{M} = (M_{a_1a_2}) = \lim_{N\to\infty}\dot{\mathbf{M}} = \mathbf{Q}^{-1/2}\mathbf{B}'(\mathbf{BQ}^{-1}\mathbf{B}')^{-1}\mathbf{BQ}^{-1/2}$$

($q \times q$ idempotent matrix of rank r), it is easy to see that the limiting nonnull distribution of T_ψ is the same as that of

$$[\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + \boldsymbol{\Theta}_\varepsilon'\}]'(\mathbf{M}\otimes\boldsymbol{\Sigma}^{-1})\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + \boldsymbol{\Theta}_\varepsilon'\},$$

which is asymptotically distributed as the noncentral chi-square distribution with $f = pr$ degrees of freedom and noncentrality parameter $\omega^2 = \{\mathrm{vec}(\boldsymbol{\Theta}_\varepsilon')\}'(\mathbf{M}\otimes\boldsymbol{\Sigma}^{-1})\mathrm{vec}(\boldsymbol{\Theta}_\varepsilon')$. This is the large sample argument via Slutsky's theorem and the central limit theorem $\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2}\} \xrightarrow{d} N_{pq}(\mathbf{0}, \mathbf{I}_q\otimes\boldsymbol{\Sigma})$.

Remark 1. Define the $p \times q$ matrix $\boldsymbol{\Delta} = (\Delta_{ab}) = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Theta}_\varepsilon'\mathbf{M}$. Using the fact that $\mathbf{M}$ is idempotent, we have $\omega^2 = \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Theta}_\varepsilon'\mathbf{M}\boldsymbol{\Theta}_\varepsilon) = \mathrm{tr}(\boldsymbol{\Omega})$, where $\boldsymbol{\Omega} = \boldsymbol{\Delta}\boldsymbol{\Delta}'\boldsymbol{\Sigma}$ ($p \times p$ matrix).

2.2. Asymptotic expansion

Our result in this paper is an asymptotic expansion for the nonnull distribution of T_ψ. We require additional notation. Writing

$$\begin{aligned}&[\boldsymbol{\mathcal{X}}^{(1)}, \ldots, \boldsymbol{\mathcal{X}}^{(N)}]'\dot{\mathbf{M}}[\boldsymbol{\mathcal{X}}^{(1)}, \ldots, \boldsymbol{\mathcal{X}}^{(N)}]\\ &\quad = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\{\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\}^{-1}\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \dot{\boldsymbol{\Psi}} = (\dot{\Psi}_{ii'})_{i,i'=1,\ldots,N} \quad (\text{say}),\end{aligned}$$

we first define

$$\dot{A}_0 = N\sum_{i=1}^{N}\dot{\Psi}_{ii}^2 - r(r+2) \to \sum_{a_1a_2a_3a_4=1}^{q} w_{a_1a_2a_3a_4}M_{a_1a_2}M_{a_3a_4} - r(r+2) = A_0,$$

$$\dot{A}_1 = N\sum_{i_1i_2=1}^{N}\dot{\Psi}_{i_1i_2}^3 \to \sum_{a_1a_2a_3a_4a_5a_6=1}^{q} w_{a_1a_2a_3}w_{a_4a_5a_6}M_{a_1a_4}M_{a_2a_5}M_{a_3a_5} = A_1,$$

$$\dot{A}_2 = N\sum_{i_1i_2=1}^{N}\dot{\Psi}_{i_1i_1}\dot{\Psi}_{i_1i_2}\dot{\Psi}_{i_2i_2} \to \sum_{a_1a_2a_3a_4a_5a_6=1}^{q} w_{a_1a_2a_3}w_{a_4a_5a_6}M_{a_1a_2}M_{a_3a_4}M_{a_5a_6} = A_2,$$

$$\dot{A}_3^\dagger = \frac{1}{N}\sum_{i_1i_2=1}^{N}\dot{\Psi}_{i_1i_2} \to \sum_{a_1a_2=1}^{q} w_{a_1}w_{a_2}M_{a_1a_2} = A_3^\dagger,$$

$$\dot{A}_4^\dagger = \sum_{i_1i_2=1}^{N}\dot{\Psi}_{i_1i_2}\dot{\Psi}_{i_2i_2} \to \sum_{aa_1a_2a_3=1}^{q} w_a w_{a_1a_2a_3}M_{aa_3}M_{a_1a_2} = A_4^\dagger$$

as $N\to\infty$.

It was shown in Wakaki et al. [62] that the two-term asymptotic expansion for the null distribution of $T = T_{LR}, T_{LH}, T_{BNP}$ depends on three summarized cumulants

$$K_4 = \sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4}\sigma^{j_1j_2}\sigma^{j_3j_4}, \tag{4}$$

$$K_{33,1} = \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}\sigma^{j_1j_4}\sigma^{j_2j_5}\sigma^{j_3j_6}, \tag{5}$$

$$K_{33,2} = \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}\sigma^{j_1j_2}\sigma^{j_3j_4}\sigma^{j_5j_6} \tag{6}$$

(σ^{jk} is the (j,k)th element of $\boldsymbol{\Sigma}^{-1}$), together with five quantities $\dot{A}_0$, $\dot{A}_1$, $\dot{A}_2$, $\dot{A}_3^\dagger$ and $\dot{A}_4^\dagger$. Unfortunately, we find that the coefficients of a_5 $(=\dot{A}_3^\dagger/8)$ in [62, (2.1)] are incorrect. Anyway, under our conditions (C_1)–(C_4), the null distribution in [62, (2.1)] should read as

$$\Pr[\{1-2c_1/(N-q)\}T_\psi \le x|H] = G_f(x) + \frac{1}{N}\sum_{\ell=0}^{3}\pi_{2,\ell}^{[0]}G_{f+2\ell}(x) + \frac{c_1}{N}f\{G_f(x) - G_{f+2}(x)\}$$
$$- \frac{\psi''}{4N}f(p+r+1)\{G_{f+2}(x) - G_{f+4}(x)\} + o(N^{-1})$$

(we set $c_1 = (p-r+1)/4, (p+1)/2, -r/2$ according to $T_\psi = T_{LR}, T_{LH}, T_{BNP}$), where

$$\pi_{2,0}^{[0]} = -\frac{f}{4}(p-r+1) + \frac{A_0}{8}K_4 + \left(-\frac{A_1}{12} + \frac{A_3^\dagger}{4}r\right)K_{33,1} + \left(-\frac{A_2}{8} - \frac{A_3^\dagger}{8}r^2 + \frac{A_4^\dagger}{4}r\right)K_{33,2},$$

$$\pi_{2,1}^{[0]} = -\frac{f}{2}r - \frac{A_0}{4}K_4 + \left(\frac{A_1}{4} - \frac{A_3^\dagger}{4}r - \frac{A_4^\dagger}{2}\right)K_{33,1} + \left\{\frac{3A_2}{8} + \frac{A_3^\dagger}{8}r(3r+4) - \frac{A_4^\dagger}{4}(3r+2)\right\}K_{33,2},$$

$$\pi_{2,2}^{[0]} = \frac{f}{4}(p+r+1) + \frac{A_0}{8}K_4 + \left\{-\frac{A_1}{4} - \frac{A_3^\dagger}{4}(r+2) + A_4^\dagger\right\}K_{33,1}$$
$$+ \left\{-\frac{3A_2}{8} - \frac{A_3^\dagger}{8}(r+2)(3r+2) + \frac{A_4^\dagger}{4}(3r+4)\right\}K_{33,2},$$

$$\pi_{2,3}^{[0]} = \left\{\frac{A_1}{12} + \frac{A_3^\dagger}{4}(r+2) - \frac{A_4^\dagger}{2}\right\}K_{33,1} + \left\{\frac{A_2}{8} + \frac{A_3^\dagger}{8}(r+2)^2 - \frac{A_4^\dagger}{4}(r+2)\right\}K_{33,2}$$

and we denote by $G_\nu(x)$ the distribution function of the central chi-square distribution with ν degrees of freedom.

We next prepare the following seven homogeneous polynomials of degrees 1, 2, 3 and 4 in $\boldsymbol{\Theta}_\varepsilon \in \mathcal{R}^{q\times p}$ through $\boldsymbol{\Delta}$ (see Remark 1):

$$
\begin{aligned}
K_3^{[a_3]} &= \sum_{j_1j_2j_3=1}^{p} \kappa_{j_1,j_2,j_3}\sigma^{j_1j_2}\Delta_{j_3a_3},\\
K_3^{[a_1a_2a_3]} &= \sum_{j_1j_2j_3=1}^{p} \kappa_{j_1,j_2,j_3}\Delta_{j_1a_1}\Delta_{j_2a_2}\Delta_{j_3a_3},\\
K_4^{[a_3a_4]} &= \sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4}\sigma^{j_1j_2}\Delta_{j_3a_3}\Delta_{j_4a_4},\\
K_4^{[a_1a_2a_3a_4]} &= \sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4}\Delta_{j_1a_1}\Delta_{j_2a_2}\Delta_{j_3a_3}\Delta_{j_4a_4},\\
K_{33,1}^{[a_3a_6]} &= \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}\sigma^{j_1j_4}\sigma^{j_2j_5}\Delta_{j_3a_3}\Delta_{j_6a_6},\\
K_{33,2}^{[a_5a_6]} &= \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}\sigma^{j_1j_2}\sigma^{j_3j_4}\Delta_{j_5a_5}\Delta_{j_6a_6},\\
K_{33}^{[a_2a_3a_5a_6]} &= \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}\sigma^{j_1j_4}\Delta_{j_2a_2}\Delta_{j_3a_3}\Delta_{j_5a_5}\Delta_{j_6a_6}
\end{aligned}
$$

for $a_1,\ldots,a_6 \in \{1,\ldots,q\}$ (we can see that these quantities are obtained from (4)–(6) when at least one $\sigma^{jj'}$ is replaced by $\Delta_{ja}\Delta_{j'a'}$).

Finally, the formula of $\lim_{N\to\infty} N(\dot{\mathbf{M}}-\mathbf{M})$ is given as follows. Since (C_2)(i) implies $(N^{-1}\mathbf{X}'\mathbf{X})^{-1} = \mathbf{Q}^{-1} - N^{-1}\mathbf{Q}^{-1}\mathbf{Q}_1\mathbf{Q}^{-1} + \mathrm{o}(N^{-1})$, there exists a symmetric $q\times q$ matrix $\widetilde{\mathbf{Q}}_1$, such that

$$(N^{-1}\mathbf{X}'\mathbf{X})^{-1/2} = \mathbf{Q}^{-1/2} + N^{-1}\widetilde{\mathbf{Q}}_1 + \mathrm{o}(N^{-1}).$$

Actually, $\widetilde{\mathbf{Q}}_1$ must satisfy $\mathbf{Q}^{-1/2}\widetilde{\mathbf{Q}}_1 + \widetilde{\mathbf{Q}}_1\mathbf{Q}^{-1/2} = -\mathbf{Q}^{-1}\mathbf{Q}_1\mathbf{Q}^{-1}$ (this matrix equation has a unique solution, which can be written in terms of the spectral decomposition of $\mathbf{Q}$). It then follows that

$$
\begin{aligned}
\dot{\mathbf{M}} &= \mathbf{M} + \frac{1}{N}(\mathbf{M}\mathbf{Q}^{-1/2}\mathbf{Q}_1\mathbf{Q}^{-1/2}\mathbf{M} + \widetilde{\mathbf{Q}}_1\mathbf{Q}^{1/2}\mathbf{M} + \mathbf{M}\mathbf{Q}^{1/2}\widetilde{\mathbf{Q}}_1) + \mathrm{o}(N^{-1})\\
&= \mathbf{M} + \frac{1}{N}\mathbf{M}_1 + \mathrm{o}(N^{-1}) \quad \text{(say)}.
\end{aligned} \tag{7}
$$

The following asymptotic expansion provides the foundation for comparing the local powers of tests for the linear hypothesis $H : \mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$ under general distributions.

Theorem 1. *For any $c_1, c_2, c_3 \in \mathbf{R}$, let $B_c(x) = x\{1 - (2/N)\sum_{j=1}^{3} c_j x^{j-1}\}$. Suppose that (C_1)–(C_4) hold. Under the local alternative (2), $B_c(T_\psi)$ admits an asymptotic expansion*

$$
\begin{aligned}
\Pr[B_c(T_\psi) \le x] = {} & G_f(x;\omega^2) + \frac{\mathrm{tr}(\boldsymbol{\Omega}_1)}{2N}\{G_{f+2}(x;\omega^2) - G_f(x;\omega^2)\}\\
& + \sum_{k=1}^{2}\frac{1}{N^{k/2}}\sum_{\ell=0}^{3k}\pi_{k,\ell}\,G_{f+2\ell}(x;\omega^2) + \frac{\psi''}{4N}\sum_{\ell=1}^{4}\pi_\ell\,G_{f+2\ell}(x;\omega^2)\\
& + \frac{1}{N}\sum_{\ell=0}^{6}\pi_\ell^c\,G_{f+2\ell}(x;\omega^2) + \mathrm{o}(N^{-1}),
\end{aligned}
$$

where $G_\nu(x;\omega^2)$ denotes the distribution function of the noncentral chi-square distribution with ν degrees of freedom and noncentrality parameter $\omega^2 = \mathrm{tr}(\boldsymbol{\Omega})$,

$$
\begin{aligned}
&\pi_1 = -f(p+r+1), \qquad \pi_2 = f(p+r+1) - 2(p+r+1)\mathrm{tr}(\boldsymbol{\Omega}),\\
&\pi_3 = 2(p+r+1)\mathrm{tr}(\boldsymbol{\Omega}) - \mathrm{tr}(\boldsymbol{\Omega}^2), \qquad \pi_4 = \mathrm{tr}(\boldsymbol{\Omega}^2),\\
&\pi_0^c = fc_1, \qquad \pi_1^c = -fc_1 + f(f+2)c_2 + c_1\omega^2,\\
&\pi_2^c = -f(f+2)c_2 + f(f+2)(f+4)c_3 + \{-c_1 + 2(f+2)c_2\}\omega^2,
\end{aligned}
$$

$$\pi_3^c = -f(f+2)(f+4)c_3 + \{-2(f+2)c_2 + 3(f+2)(f+4)c_3\}\omega^2 + c_2\omega^4,$$
$$\pi_4^c = -3(f+2)(f+4)c_3\omega^2 + \{-c_2 + 3(f+4)c_3\}\omega^4,$$
$$\pi_5^c = -3(f+4)c_3\omega^4 + c_3\omega^6, \qquad \pi_6^c = -c_3\omega^6,$$
$$\boldsymbol{\Omega}_1 = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Theta}_\varepsilon'\mathbf{M}_1\boldsymbol{\Theta}_\varepsilon \quad (p \times p \text{ matrix})$$

(the pattern of π_ℓ^c ($\ell = 0, \ldots, 6$) is identical to the coefficients which appeared in [36,37,34]). Here, the coefficients $\pi_{k,\ell}$'s, independent of ψ, are the sums of homogeneous polynomials of degrees 0, 1, 2, 3, 4 and 6 in $\boldsymbol{\Theta}_\varepsilon$; that is, $\pi_{1,\ell} = \pi_{1,\ell}^{[1]} + \pi_{1,\ell}^{[3]}$ ($\ell = 0, 1, 2$), $\pi_{1,3} = \pi_{1,3}^{[3]}$, $\pi_{2,\ell} = \pi_{2,\ell}^{[0]} + \pi_{2,\ell}^{[2]} + \pi_{2,\ell}^{[4]} + \pi_{2,\ell}^{[6]}$ ($\ell = 0, 1, 2, 3$), $\pi_{2,4} = \pi_{2,4}^{[2]} + \pi_{2,4}^{[4]} + \pi_{2,4}^{[6]}$, $\pi_{2,5} = \pi_{2,5}^{[4]} + \pi_{2,5}^{[6]}$, $\pi_{2,6} = \pi_{2,6}^{[6]}$ (the explicit expressions, except for $\pi_{2,\ell}^{[0]}$, where the details can be obtained from author on request, are not reported here, to preserve space, since they are complicated and long formulae depending on $K_3^{[a_3]}, \ldots, K_{33}^{[a_2a_3a_5a_6]}$; but it is remarkable that they are sometimes simplified drastically as in Section 3.2).

Letting $\boldsymbol{\vartheta}(\psi'') = (\vartheta_1, \vartheta_2(\psi''), \vartheta_3)$, where $\vartheta_1 = -\pi_{2,0}^{[0]}/f$,

$$\vartheta_2(\psi'') = -\frac{\pi_{2,0}^{[0]} + \pi_{2,1}^{[0]}}{f(f+2)} + \frac{(p+r+1)\psi''}{4(f+2)}, \qquad \vartheta_3 = -\frac{\pi_{2,0}^{[0]} + \pi_{2,1}^{[0]} + \pi_{2,2}^{[0]}}{f(f+2)(f+4)}$$

(see [16]), we have $\Pr[B_{\vartheta(\psi'')}(T_\psi) \le x|H] = G_f(x) + \mathrm{o}(N^{-1})$, which improves the large sample approximation $\Pr[T_\psi \le x|H] = G_f(x) + \mathrm{o}(1)$.

Corollary 2 (*Bartlett's Type Adjustment*). *Under the local alternative (2),*

$$\Pr[B_{\vartheta(\psi'')}(T_\psi) \le x] = \sum_{k=0}^{2} \frac{1}{N^{k/2}} \sum_{\ell=0}^{3k} \mathcal{P}_{k,\ell}\, G_{f+2\ell}(x; \omega^2) + \mathrm{o}(N^{-1}),$$

where $\mathcal{P}_{0,0} = 1$, $\mathcal{P}_{1,\ell} = \pi_{1,\ell}^{[1]} + \pi_{1,\ell}^{[3]}$ ($\ell = 0, 1, 2$), $\mathcal{P}_{1,3} = \pi_{1,3}^{[3]}$,

$$\mathcal{P}_{2,0} = \left[\pi_{2,0}^{[2]} - \frac{\mathrm{tr}(\boldsymbol{\Omega}_1)}{2}\right] + \pi_{2,0}^{[4]} + \pi_{2,0}^{[6]},$$
$$\mathcal{P}_{2,1} = \left[\pi_{2,1}^{[2]} + \vartheta_1 \mathrm{tr}(\boldsymbol{\Omega}) + \frac{\mathrm{tr}(\boldsymbol{\Omega}_1)}{2}\right] + \pi_{2,1}^{[4]} + \pi_{2,1}^{[6]},$$
$$\mathcal{P}_{2,2} = \left[\pi_{2,2}^{[2]} + \{-\vartheta_1 + 2(f+2)\vartheta_2(0)\}\mathrm{tr}(\boldsymbol{\Omega})\right] + \pi_{2,2}^{[4]} + \pi_{2,2}^{[6]},$$
$$\begin{aligned}\mathcal{P}_{2,3} &= \left[\pi_{2,3}^{[2]} + \{-2(f+2)\vartheta_2(0) + 3(f+2)(f+4)\vartheta_3\}\mathrm{tr}(\boldsymbol{\Omega})\right] \\ &\quad + \left[\pi_{2,3}^{[4]} - \frac{\psi''}{4}\mathrm{tr}(\boldsymbol{\Omega}^2) + \vartheta_2(\psi'')\{\mathrm{tr}(\boldsymbol{\Omega})\}^2\right] + \pi_{2,3}^{[6]},\end{aligned}$$
$$\begin{aligned}\mathcal{P}_{2,4} &= \left[\pi_{2,4}^{[2]} - 3(f+2)(f+4)\vartheta_3 \mathrm{tr}(\boldsymbol{\Omega})\right] \\ &\quad + \left[\pi_{2,4}^{[4]} + \frac{\psi''}{4}\mathrm{tr}(\boldsymbol{\Omega}^2) + \{-\vartheta_2(\psi'') + 3(f+4)\vartheta_3\}\{\mathrm{tr}(\boldsymbol{\Omega})\}^2\right] + \pi_{2,4}^{[6]},\end{aligned}$$
$$\mathcal{P}_{2,5} = \left[\pi_{2,5}^{[4]} - 3(f+4)\vartheta_3\{\mathrm{tr}(\boldsymbol{\Omega})\}^2\right] + \left[\pi_{2,5}^{[6]} + \vartheta_3\{\mathrm{tr}(\boldsymbol{\Omega})\}^3\right],$$
$$\mathcal{P}_{2,6} = \left[\pi_{2,6}^{[6]} - \vartheta_3\{\mathrm{tr}(\boldsymbol{\Omega})\}^3\right].$$

It is well known that for the chi-square type asymptotic expansion, making use of the Cornish–Fisher type expansion $CF_{\vartheta(\psi'')}(\chi^2_{f,\alpha})$, where $CF_c(x) = x\{1 + (2/N)\sum_{j=1}^3 c_j x^{j-1}\}$, yields the improvement $\Pr[T_\psi \le CF_{\vartheta(\psi'')}(\chi^2_{f,\alpha})|H] = 1 - \alpha + \mathrm{o}(N^{-1})$ over the large sample approximation $\Pr[T_\psi \le \chi^2_{f,\alpha}|H] = 1 - \alpha + \mathrm{o}(1)$, where $\chi^2_{f,\alpha}$ is the upper α percentile of the central chi-square distribution with f degrees of freedom. Using the relations $xg_f(x; \omega^2) = fg_{f+2}(x; \omega^2) + \omega^2 g_{f+4}(x; \omega^2)$ and $G_f(x; \omega^2) - G_{f+2}(x; \omega^2) = 2g_{f+2}(x; \omega^2)$, where $g_\nu(x; \omega^2)$ is the probability density function of the noncentral chi-square distribution with ν degrees of freedom and noncentrality parameter ω^2, we have

Corollary 3 (*Cornish–Fisher's Type Adjustment*). *Under the local alternative (2),*

$$\Pr[T_\psi \le CF_{\vartheta(\psi'')}(\chi^2_{f,\alpha})] = \sum_{k=0}^{2} \frac{1}{N^{k/2}} \sum_{\ell=0}^{3k} \mathcal{P}_{k,\ell}\, G_{f+2\ell}(\chi^2_{f,\alpha}; \omega^2) + \mathrm{o}(N^{-1}).$$

3. Implication of asymptotic expansion

3.1. Third-order local power comparison

Recall that $\boldsymbol{\vartheta}(\psi'')$ linearly depends on summarized cumulants $K_4, K_{33,1}, K_{33,2}$, given by (4)–(6). We take appropriate location invariant estimators $\widehat{K}_4, \widehat{K}_{33,1}, \widehat{K}_{33,2}$ and replace $\boldsymbol{\vartheta}(\psi'')$ by $\widehat{\boldsymbol{\vartheta}}(\psi'') = (\widehat{\vartheta}_1, \widehat{\vartheta}_2(\psi''), \widehat{\vartheta}_3)$. For example, we can construct an estimator

$$\widehat{K}_4 = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\mathbf{u}}_i'\widehat{\boldsymbol{\Sigma}}_Y^{-1}\widehat{\mathbf{u}}_i)^2 - p(p+2),$$

$$\widehat{K}_{33,1} = \frac{1}{N^2}\sum_{ii'=1}^{N}(\widehat{\mathbf{u}}_i'\widehat{\boldsymbol{\Sigma}}_Y^{-1}\widehat{\mathbf{u}}_{i'})^3,$$

$$\widehat{K}_{33,2} = \frac{1}{N^2}\sum_{ii'=1}^{N}(\widehat{\mathbf{u}}_i'\widehat{\boldsymbol{\Sigma}}_Y^{-1}\widehat{\mathbf{u}}_i)(\widehat{\mathbf{u}}_i'\widehat{\boldsymbol{\Sigma}}_Y^{-1}\widehat{\mathbf{u}}_{i'})(\widehat{\mathbf{u}}_{i'}'\widehat{\boldsymbol{\Sigma}}_Y^{-1}\widehat{\mathbf{u}}_{i'})$$

based on the residual matrix $[\widehat{\mathbf{u}}_1, \ldots, \widehat{\mathbf{u}}_N] = \mathbf{Y}'\{\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}$. Some asymptotic properties of estimators for $K_4, K_{33,1}, K_{33,2}$ are found in Mardia [45] and McCullagh [46, p107–109]. Yanagihara [65] proposed a family of estimators for Mardia's multivariate kurtosis K_4.

Corollary 4. *Suppose that (* C$_1$*)–(* C$_4$*) hold.*
(i) *If location invariant estimators $\widehat{K}_4, \widehat{K}_{33,1}, \widehat{K}_{33,2}$ satisfy*

$$(\widehat{K}_4, \widehat{K}_{33,1}, \widehat{K}_{33,2}) = (K_4, K_{33,1}, K_{33,2}) + O_p(N^{-\tau}) \tag{8}$$

for some $\tau \in (0, 1/2]$, we then have under the local alternative (2)

$$\Pr[B_{\widehat{\vartheta}(\psi'')}(T_\psi) \le x] = \Pr[B_{\vartheta(\psi'')}(T_\psi) \le x] + o(N^{-1}). \tag{9}$$

(ii) *If location invariant estimators $\widehat{K}_4, \widehat{K}_{33,1}, \widehat{K}_{33,2}$ satisfy*

$$\Pr[|\widehat{K}_4 - K_4| + |\widehat{K}_{33,1} - K_{33,1}| + |\widehat{K}_{33,2} - K_{33,2}| \ge \rho_N] = o(N^{-1}) \tag{10}$$

for some sequence $\rho_N \to 0$, we then have under the local alternative (2)

$$\Pr[T_\psi \le CF_{\widehat{\vartheta}(\psi'')}(\chi^2_{f,\alpha})] = \Pr[T_\psi \le CF_{\vartheta(\psi'')}(\chi^2_{f,\alpha})] + o(N^{-1}). \tag{11}$$

Proof. As in Kakizawa and Iwashita [37], (8) or (10) supports Chibisov's lemma [15] (see also [44]) to conclude (9) or (11). □

Our higher-order result shows that the local power function of tests for the linear hypothesis $H : \mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$ depends on the factor

$$D_{p,r}(\boldsymbol{\Omega}) = \operatorname{tr}(\boldsymbol{\Omega}^2) - \frac{p+r+1}{pr+2}\{\operatorname{tr}(\boldsymbol{\Omega})\}^2 = p\left\{\sigma_\lambda^2 - \frac{(p-1)(p+2)}{pr+2}\bar{\lambda}^2\right\},$$

where $\bar{\lambda} = \sum_{i=1}^p \lambda_i/p$ and $\sigma_\lambda = \{\sum_{i=1}^p(\lambda_i - \bar{\lambda})^2/p\}^{1/2}$, with $\lambda_1 \ge \cdots \ge \lambda_p \ge 0$ being eigenvalues of $\boldsymbol{\Omega}$ (for $p > r$, $\lambda_{r+1} = \cdots = \lambda_p = 0$). Following Kakizawa and Iwashita [37] with $r = q - 1$, who considered the equality of q mean vectors in the one-way MANOVA model, we have two equivalent procedures: One is the test procedure that if Bartlett's type adjusted criterion $B_{\widehat{\vartheta}(\psi'')}(T_\psi)$ exceeds $\chi^2_{f,\alpha}$, then we reject the null hypothesis H (we call B_ψ test). The other is the test based on Cornish–Fisher's type expansion that if T_ψ exceeds the size corrected critical value $CF_{\widehat{\vartheta}(\psi'')}(\chi^2_{f,\alpha})$, then we reject the null hypothesis H (we call CF_ψ test). In this way, given two ψ-functions ψ_1 and ψ_2, the difference of the powers for ψ_1 and ψ_2 tests (more precisely, B_{ψ_j} or CF_{ψ_j} tests ($j = 1, 2$)) under the local alternative (2) is given by

$$\begin{aligned}
&\lim_{N\to\infty}\left\{N\left(\Pr[B_{\widehat{\vartheta}(\psi_2'')}(T_{\psi_2}) > \chi^2_{f,\alpha}] - \Pr[B_{\widehat{\vartheta}(\psi_1'')}(T_{\psi_1}) > \chi^2_{f,\alpha}]\right)\right\} \\
&= \lim_{N\to\infty}\left\{N\left(\Pr[T_{\psi_2} > CF_{\widehat{\vartheta}(\psi_2'')}(\chi^2_{f,\alpha})] - \Pr[T_{\psi_1} > CF_{\widehat{\vartheta}(\psi_1'')}(\chi^2_{f,\alpha})]\right)\right\} \\
&= \frac{\psi_2'' - \psi_1''}{2} D_{p,r}(\boldsymbol{\Omega}) g_{f+8}(\chi^2_{f,\alpha}; \omega^2).
\end{aligned}$$

The conclusion stated in Anderson ([1], p336; he cited an unpublished working paper [57] and then made a power comparison among T_{LR}, T_{LH} and T_{BNP}) for the normal case (see also [21,23]) can be extended even for the general distributions as follows: For the case $\min(p, r) > 1$ (otherwise, $D_{p,r}(\boldsymbol{\Omega}) \equiv 0$), as long as $\psi_2'' - \psi_1''$ is positive, ψ_2 test is superior (inferior) to ψ_1 test if $\sigma_\lambda/\overline{\lambda}$ is greater (less) than or equal to $[(p-1)(p+2)/(pr+2)]^{1/2}$ (when $\psi_2'' - \psi_1''$ is negative, the ordering of power is reversed). We emphasize that our analysis, including the one-way MANOVA model (see [37]), is done without assumption of normality in the multivariate linear regression model and even in the GMANOVA model (see Section 5 below).

3.2. Special case

In some cases, the design matrix $\mathbf{X}$ and the constraint $r \times q$ matrix $\mathbf{B}$ (of rank $r \le q$) in a given linear hypothesis $H : \mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$ may satisfy $\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{1}_N = \mathbf{0}$, equivalently

$$(\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{1}_N = \sum_{i=1}^{N} \mathcal{X}^{(i)} \in \{\mathbf{h} \in \mathbf{R}^q : \dot{\mathbf{M}}\mathbf{h} = \mathbf{0}\}, \tag{12}$$

where $\mathbf{1}_N = (1, \ldots, 1)' \in \mathbf{R}^N$. It is worthwhile to conclude that using (C$_2$)(ii) and $\dot{\mathbf{M}} \to \mathbf{M}$, (12) implies $M_{aa'}w_{a'} = 0$ for $a = 1, \ldots, q$, hence, $A_3^\dagger = A_4^\dagger = 0$ (at the same time, Lemma 5 indicates that there are many other quantities in the coefficients $\pi_{k,\ell}$'s of Theorem 1, which also have the same property as $(A_3^\dagger, A_4^\dagger)$). Then,

$$\pi_{1,0} = \frac{1}{2}K_3^{*[1]} - \frac{1}{6}K_3^{*[3]} \equiv \pi_{1,0}^{[1]} + \pi_{1,0}^{[3]}, \qquad \pi_{1,1} = -K_3^{*[1]} + \frac{1}{2}K_3^{*[3]} \equiv \pi_{1,1}^{[1]} + \pi_{1,1}^{[3]},$$
$$\pi_{1,2} = \frac{1}{2}K_3^{*[1]} - \frac{1}{2}K_3^{*[3]} \equiv \pi_{1,2}^{[1]} + \pi_{1,2}^{[3]}, \qquad \pi_{1,3} = \frac{1}{6}K_3^{*[3]} \equiv \pi_{1,3}^{[3]}$$

and the resulting coefficients $\pi_{2,\ell}^{[n]}$'s are drastically simplified as follows:

$$\pi_{2,0}^{[0]} = -\frac{f}{4}(p-r+1) + \frac{A_0}{8}K_4 - \frac{A_1}{12}K_{33,1} - \frac{A_2}{8}K_{33,2},$$
$$\pi_{2,0}^{[2]} = \frac{r}{4}K_4^{[2]} - \frac{1}{4}K_4^{*[2]} + \frac{1}{4}K_{33}^{**[2]} + \frac{1}{8}(K_3^{*[1]})^2,$$
$$\pi_{2,0}^{[4]} = \frac{1}{24}K_4^{*[4]} - \frac{1}{8}K_{33}^{**[4]} - \frac{1}{12}K_3^{*[1]}K_3^{*[3]}, \qquad \pi_{2,0}^{[6]} = \frac{1}{72}(K_3^{*[3]})^2,$$
$$\pi_{2,1}^{[0]} = -\frac{f}{2}r - \frac{A_0}{4}K_4 + \frac{A_1}{4}K_{33,1} + \frac{3A_2}{8}K_{33,2},$$
$$\pi_{2,1}^{[2]} = \frac{r}{2}\mathrm{tr}(\boldsymbol{\Omega}) - \frac{1}{4}(3r+4)K_4^{[2]} + \frac{3}{4}K_4^{*[2]} - K_{33}^{**[2]} - \frac{1}{2}(K_3^{*[1]})^2,$$
$$\pi_{2,1}^{[4]} = \frac{1}{4}K_4^{[4]} - \frac{1}{6}K_4^{*[4]} + \frac{5}{8}K_{33}^{**[4]} + \frac{5}{12}K_3^{*[1]}K_3^{*[3]}, \qquad \pi_{2,1}^{[6]} = -\frac{1}{12}(K_3^{*[3]})^2,$$
$$\pi_{2,2}^{[0]} = \frac{f}{4}(p+r+1) + \frac{A_0}{8}K_4 - \frac{A_1}{4}K_{33,1} - \frac{3A_2}{8}K_{33,2},$$
$$\pi_{2,2}^{[2]} = -\frac{1}{2}(p+2r+1)\mathrm{tr}(\boldsymbol{\Omega}) + \frac{3}{4}(r+2)K_4^{[2]} - \frac{3}{4}K_4^{*[2]} + \frac{3}{2}K_{33}^{**[2]} + \frac{3}{4}(K_3^{*[1]})^2,$$
$$\pi_{2,2}^{[4]} = \frac{1}{4}\mathrm{tr}(\boldsymbol{\Omega}^2) - \frac{5}{8}K_4^{[4]} + \frac{1}{4}K_4^{*[4]} - \frac{5}{4}K_{33}^{**[4]} - \frac{5}{6}K_3^{*[1]}K_3^{*[3]}, \qquad \pi_{2,2}^{[6]} = \frac{5}{24}(K_3^{*[3]})^2,$$
$$\pi_{2,3}^{[0]} = \frac{A_1}{12}K_{33,1} + \frac{A_2}{8}K_{33,2},$$
$$\pi_{2,3}^{[2]} = \frac{1}{2}(p+r+1)\mathrm{tr}(\boldsymbol{\Omega}) - \frac{1}{4}(r+2)K_4^{[2]} + \frac{1}{4}K_4^{*[2]} - K_{33}^{**[2]} - \frac{1}{2}(K_3^{*[1]})^2,$$
$$\pi_{2,3}^{[4]} = -\frac{1}{2}\mathrm{tr}(\boldsymbol{\Omega}^2) + \frac{1}{2}K_4^{[4]} - \frac{1}{6}K_4^{*[4]} + \frac{5}{4}K_{33}^{**[4]} + \frac{5}{6}K_3^{*[1]}K_3^{*[3]}, \qquad \pi_{2,3}^{[6]} = -\frac{5}{18}(K_3^{*[3]})^2,$$
$$\pi_{2,4}^{[2]} = \frac{1}{4}K_{33}^{**[2]} + \frac{1}{8}(K_3^{*[1]})^2,$$
$$\pi_{2,4}^{[4]} = \frac{1}{4}\mathrm{tr}(\boldsymbol{\Omega}^2) - \frac{1}{8}K_4^{[4]} + \frac{1}{24}K_4^{*[4]} - \frac{5}{8}K_{33}^{**[4]} - \frac{5}{12}K_3^{*[1]}K_3^{*[3]}, \qquad \pi_{2,4}^{[6]} = \frac{5}{24}(K_3^{*[3]})^2,$$
$$\pi_{2,5}^{[4]} = \frac{1}{8}K_{33}^{**[4]} + \frac{1}{12}K_3^{*[1]}K_3^{*[3]}, \qquad \pi_{2,5}^{[6]} = -\frac{1}{12}(K_3^{*[3]})^2, \qquad \pi_{2,6}^{[6]} = \frac{1}{72}(K_3^{*[3]})^2,$$

where

$$K_3^{*[1]} = \sum_{a_1a_2a_3=1}^{q} w_{a_1a_2a_3}M_{a_1a_2}K_3^{[a_3]}, \qquad K_3^{*[3]} = \sum_{a_1a_2a_3=1}^{q} w_{a_1a_2a_3}K_3^{[a_1a_2a_3]},$$

$$K_4^{[2]} = \sum_{a=1}^{q} K_4^{[aa]}, \qquad K_4^{[4]} = \sum_{aa'=1}^{q} K_4^{[aaa'a']},$$

$$K_4^{*[2]} = \sum_{a_1a_2a_3a_4=1}^{q} w_{a_1a_2a_3a_4} M_{a_1a_2} K_4^{[a_3a_4]}, \qquad K_4^{*[4]} = \sum_{a_1a_2a_3a_4=1}^{q} w_{a_1a_2a_3a_4} K_4^{[a_1a_2a_3a_4]},$$

$$K_{33}^{**[2]} = \sum_{a_1a_2a_3a_4a_5a_6=1}^{q} w_{a_1a_2a_3} w_{a_4a_5a_6} (M_{a_1a_4} M_{a_2a_5} K_{33,1}^{[a_3a_6]} + M_{a_1a_2} M_{a_3a_4} K_{33,2}^{[a_5a_6]}),$$

$$K_{33}^{**[4]} = \sum_{a_1a_2a_3a_4a_5a_6=1}^{q} w_{a_1a_2a_3} w_{a_4a_5a_6} M_{a_1a_4} K_{33}^{[a_2a_3a_5a_6]}$$

(we remark that the notation with the superscript '*' or '**' is a sum of a single or double w-function; $w_{a_1a_2a_3}$ or $w_{a_1a_2a_3a_4}$ or $w_{a_1a_2a_3} w_{a_4a_5a_6}$).

Example 1. Given an integer $q \geq 2$, consider a multivariate one-way classification model $\mathbf{y}_i^{(a)} = \boldsymbol{\mu}^{(a)} + \mathbf{u}_i^{(a)}$ ($a = 1, \ldots, q$; $i = 1, \ldots, N_a$). This model can be written as the form $\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{U}$ with $\boldsymbol{\Theta}' = [\boldsymbol{\mu}^{(1)}, \ldots, \boldsymbol{\mu}^{(q)}]$ ($p \times q$ matrix), in which the design matrix is given by

$$\mathbf{X} = \begin{pmatrix} \mathbf{1}_{N_1} & & & \\ & \mathbf{1}_{N_2} & & \\ & & \ddots & \\ & & & \mathbf{1}_{N_q} \end{pmatrix} \quad \left(\text{write } N = \sum_{a=1}^{q} N_a \right).$$

As in Kakizawa and Iwashita [37], we assume that $(N_a/N)^{1/2} = \dot{\rho}_a$ ($a = 1, \ldots, q$) have the form $\dot{\rho}_a^2 = \rho_a^2 + N^{-1}\tau_a$, with $(\rho_1, \ldots, \rho_q)'$ and $(\tau_1, \ldots, \tau_q)'$ being, respectively, a vector of positive real numbers and a vector of integers, independent of N, satisfying $\rho_1^2 + \cdots + \rho_q^2 = 1$ and $\tau_1 + \cdots + \tau_q = 0$ (of course, $N_a = N\rho_a^2 + \tau_a \in \mathbf{N}$). It is easy to see that (C_2) holds, since we have $N^{-1}\mathbf{X}'\mathbf{X} = \mathrm{diag}(\dot{\rho}_1^2, \ldots, \dot{\rho}_q^2)$ and for any $s \in \mathbf{N}$, $N^{s/2}\overline{\mathfrak{X}}_{a_1 \ldots a_s} = \dot{\rho}_a^{-(s-2)}$ if $a_1 = \cdots = a_s = a \in \{1, \ldots, q\}$ and 0 otherwise. Then, (12) holds whenever the $r \times q$ matrix $\mathbf{B}$ of rank $r \leq q$ satisfies $\mathbf{B}\mathbf{1}_q = \mathbf{0}$ (use $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{1}_N = \mathbf{1}_q$ for the one-way classification model). Such a typical example is given by $\mathbf{B} = [\mathbf{I}_{q-1}, -\mathbf{1}_{q-1}]$ ($(q-1) \times q$ matrix), which corresponds to the problem of testing $\boldsymbol{\mu}^{(1)} = \cdots = \boldsymbol{\mu}^{(q)}$.

Example 2. In the usual multiple regression model, the first column of the design $N \times q$ matrix $\mathbf{X}$ of rank q is assumed to be $\mathbf{1}_N$, that is, $\mathbf{X} = [\mathbf{1}_N, \mathbf{X}_{(-1)}]$. In this case, since $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{1}_N = (1, 0, \ldots, 0)' \in \mathbf{R}^q$, (12) holds whenever the $r \times q$ matrix $\mathbf{B}$ of rank r has the form $\mathbf{B} = [\mathbf{0}_r, \mathbf{B}_{(-1)}]$ for some $r \in \{1, \ldots, q-1\}$.

4. Derivation of asymptotic expansion of Theorem 1

4.1. Differential operator approach for characteristic function

The differential operator approach under normality was first used by Welch [63] and James [31], who made an important contribution to the derivation of asymptotic expansions in multivariate statistical analysis under normality, in such a way that the independence of the standardized sample mean vectors (independent normal distributions) and the sample covariance matrices (independent Wishart distributions) enables us to use the conditional approach and then evaluate the expectation of a function of the Wishart distribution via Welch-James's technique (see e.g. [58]). In the normal GMANOVA model descrived later, Fujikoshi [22] used the canonical reduction method due to Gleser and Olkin [28] and then apply Welch-James's technique on the basis of the conditional set-up, which is slightly different from Gleser and Olkin [28].

As pointed out in Fujikoshi [26] for multivariate test statistics, it is crucial to find a convenient device (especially under nonnormality) for giving an asymptotic expansion of the characteristic function according to situations under consideration. Unlike Kano [39], Fujikoshi [24–26], Wakaki et al. [62] and Yanagihara [64], this subsection is devoted to extension of the differential operator approach developed by Kakizawa and Iwashita [36,37] to the multivariate linear regression model.

Now, recall that (3) is idempotent of rank r, having spectral decomposition

$$\dot{\mathbf{M}} = \dot{\mathbf{V}}\mathrm{diag}(\underbrace{1, \ldots, 1}_{r \text{ times}}, 0, \ldots, 0)\dot{\mathbf{V}}' = \dot{\mathbf{V}}^{(1:r)}(\dot{\mathbf{V}}^{(1:r)})', \tag{13}$$

where $\dot{\mathbf{V}}^{(1:r)} = [\dot{\mathbf{v}}^{(1)}, \ldots, \dot{\mathbf{v}}^{(r)}]$, with $\dot{\mathbf{V}} = [\dot{\mathbf{v}}^{(1)}, \ldots, \dot{\mathbf{v}}^{(q)}]$ being orthogonal matrix of $q \times q$ (for any matrix $\mathbf{A}$, $\mathbf{A}^{(1:r)}$ is a matrix which consists of the first r columns of $\mathbf{A}$; we shall sometimes use the notation $\mathbf{a}^{(1:r)} = [\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(r)}]$ for simplicity).

A technical observation perhaps worth emphasizing here is that the following lemma is concerned with the expectation of a certain analytic function of

$$\mathbf{z}_U^{(1:r)} = [\mathbf{z}_U^{(1)}, \ldots, \mathbf{z}_U^{(r)}] = \mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2}\dot{\mathbf{V}}^{(1:r)} \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_U,$$

which is applicable even in the GMANOVA model (Part II) of this paper.

Lemma 5. *Let $\boldsymbol{\gamma}^{(b)} = (\gamma_j^{(b)})$ $(b = 1, \ldots, r)$ be $p \times 1$ vectors of variables and $\boldsymbol{\Gamma} = (\gamma_{jk})$ be a $p \times p$ symmetric matrix of variables. Let $h(\boldsymbol{\gamma}^{(1:r)}, \boldsymbol{\Gamma})$ be an arbitrary multivariate polynomial of finite degree with coefficients in $\mathbf{R}$, which may depend on N but are of order $\mathrm{O}(1)$. Define vectors of differential operators by*

$$\boldsymbol{\partial}^{(b)} = (\partial_j^{(b)}) = \left(\frac{\partial}{\partial \gamma_j^{(b)}}\right) \quad (b = 1, \ldots, r)$$

and a matrix of differential operators by

$$\boldsymbol{\partial} = (\partial_{jk}) = \left(\frac{1}{2}(1 + \delta_{jk})\frac{\partial}{\partial \gamma_{jk}}\right)$$

applied to the function $\exp\{\mathrm{i}h(\boldsymbol{\gamma}^{(1:r)}, \boldsymbol{\Gamma})\}$, where $\mathrm{i} = \sqrt{-1}$. In addition to (C_1) and (C_2), if $E(\|\mathbf{u}\|^4) < \infty$, then

$$E \exp\{\mathrm{i}h(\mathbf{z}_U^{(1:r)}, \widehat{\boldsymbol{\Sigma}}_U)\} = \Xi \exp\{\mathrm{i}h(\boldsymbol{\gamma}^{(1:r)}, \boldsymbol{\Gamma})\}|_{\gamma^{(1:r)} = 0_{p,r}, \Gamma = \Sigma} + \mathrm{o}(N^{-1}),$$

where

$$\Xi = \Xi_0 \left[1 + \frac{\Xi_1}{N^{1/2}} + \frac{1}{N}\left\{\mathrm{tr}(\boldsymbol{\Sigma}\boldsymbol{\partial}\boldsymbol{\Sigma}\boldsymbol{\partial}) + \Xi_2 + \frac{\Xi_1^2}{2}\right\}\right]$$

is the differential operator with

$$\Xi_0 = \exp\left(\frac{1}{2}\sum_{b=1}^{r} \boldsymbol{\partial}^{(b)\prime}\boldsymbol{\Sigma}\boldsymbol{\partial}^{(b)}\right),$$

$$\Xi_1 = \sum_{j_1 j_2 j_3 = 1}^{p} \kappa_{j_1, j_2, j_3}\left(\sum_{b_1=1}^{r} w^{(b_1)}\partial_{j_1}^{(b_1)}\partial_{j_2 j_3} + \frac{1}{6}\sum_{b_1 b_2 b_3 = 1}^{r} w^{(b_1 b_2 b_3)}\partial_{j_1}^{(b_1)}\partial_{j_2}^{(b_2)}\partial_{j_3}^{(b_3)}\right),$$

$$\Xi_2 = \sum_{j_1 j_2 j_3 j_4 = 1}^{p} \kappa_{j_1, j_2, j_3, j_4}\frac{1}{2}\left(\partial_{j_1 j_2}\partial_{j_3 j_4} + \sum_{b=1}^{r}\partial_{j_1}^{(b)}\partial_{j_2}^{(b)}\partial_{j_3 j_4} + \frac{1}{12}\sum_{b_1 b_2 b_3 b_4 = 1}^{r} w^{(b_1 b_2 b_3 b_4)}\partial_{j_1}^{(b_1)}\partial_{j_2}^{(b_2)}\partial_{j_3}^{(b_3)}\partial_{j_4}^{(b_4)}\right)$$

and

$$w^{(b_1 \ldots b_s)} = \sum_{a_1 \ldots a_s = 1}^{q} w_{a_1 \ldots a_s} v_{a_1}^{(b_1)} \cdots v_{a_s}^{(b_s)}$$

for $s = 1, 3, 4$ (we denote by $v_a^{(b)}$ the ath element of $\mathbf{v}^{(b)}$, given in Lemma A.1).

Especially, if the design matrix $\mathbf{X}$ satisfies (12), then $w^{(b)} = 0$ for $b = 1, \ldots, r$, hence Ξ_1 simplifies

$$\Xi_1 = \frac{1}{6}\sum_{j_1 j_2 j_3 = 1}^{p} \kappa_{j_1, j_2, j_3} \sum_{b_1 b_2 b_3 = 1}^{r} w^{(b_1 b_2 b_3)}\partial_{j_1}^{(b_1)}\partial_{j_2}^{(b_2)}\partial_{j_3}^{(b_3)},$$

as illustrated in the top of Section 3.2 (see also [37, Remark 4]).

Proof. Let $\dot{v}_a^{(b)}$ be the (a, b)th element of $\dot{\mathbf{V}}$ ($q \times q$ orthogonal matrix), which satisfies

$$\sum_{a=1}^{q} \dot{v}_a^{(b_1)}\dot{v}_a^{(b_2)} = \delta_{b_1 b_2}. \tag{14}$$

We know from Lemma A.1

$$\dot{v}_a^{(b)} = v_a^{(b)} + \mathrm{O}(N^{-1}). \tag{15}$$

With $\mathbf{u}_x^{(1:q)} = [\mathbf{u}_x^{(1)}, \ldots, \mathbf{u}_x^{(q)}] = \mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2}$ ($p \times q$ matrix), we have

$$\mathbf{u}_x^{(a)} = \frac{1}{N^{1/2}}\sum_{i=1}^{N} \mathbf{u}_i (N^{1/2}\mathcal{X}_a^{(i)})$$

and

$$\widehat{\boldsymbol{\Sigma}}_U - \boldsymbol{\Sigma} = \frac{N}{N-q}\left\{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{u}_i\mathbf{u}_i' - \boldsymbol{\Sigma})\right\} - \frac{1}{N-q}\sum_{a=1}^{q}(\mathbf{u}_x^{(a)}\mathbf{u}_x^{(a)\prime} - \boldsymbol{\Sigma}).$$

Also, $\mathbf{z}_U^{(b)} = \sum_{a=1}^{q}\dot{v}_a^{(b)}\mathbf{u}_x^{(a)}$ $(b = 1, \ldots, r)$. In line with Kakizawa and Iwashita [36], we obtain

$$E\exp\{\mathrm{i}h(\mathbf{z}_U^{(1:r)}, \widehat{\boldsymbol{\Sigma}}_U)\} = E\exp\{\mathrm{i}h(\mathbf{z}_{U^\dagger}^{(1:r)}, \widehat{\boldsymbol{\Sigma}}_{U^\dagger})\} + \mathrm{o}(N^{-1}),$$

where

$$\mathbf{u}_i^\dagger = \begin{cases}\mathbf{u}_i, & \|\mathbf{u}_i\| \le N^{1/2}\\ \mathbf{0}, & \|\mathbf{u}_i\| > N^{1/2}\end{cases} \quad (i = 1, \ldots, N)$$

are truncated random vectors. Since

$$\exp\{\mathrm{i}h(\mathbf{z}_{U^\dagger}^{(1:r)}, \widehat{\boldsymbol{\Sigma}}_{U^\dagger})\} = \exp\left[\sum_{b=1}^{r}\mathbf{z}_{U^\dagger}^{(b)\prime}\boldsymbol{\partial}^{(b)} + \mathrm{tr}\{(\widehat{\boldsymbol{\Sigma}}_{U^\dagger} - \boldsymbol{\Sigma})\boldsymbol{\partial}\}\right]\exp\{\mathrm{i}h(\boldsymbol{\gamma}^{(1:r)}, \boldsymbol{\Gamma})\}|_{\gamma^{(1:r)}=0_{p,r},\Gamma=\Sigma},$$

we obtain

$$E\exp\{\mathrm{i}h(\mathbf{z}_{U^\dagger}^{(1:r)}, \widehat{\boldsymbol{\Sigma}}_{U^\dagger})\} = \Xi(\boldsymbol{\partial}^{(1:r)}, \boldsymbol{\partial})\exp\{\mathrm{i}h(\boldsymbol{\gamma}^{(1:r)}, \boldsymbol{\Gamma})\}|_{\gamma^{(1:r)}=0_{p,r},\Gamma=\Sigma}$$

(due to the truncation, the removal of the evaluation at $\boldsymbol{\gamma}^{(1:r)} = \mathbf{0}_{p,r}$ and $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}$ to the outside of the expectation regarding the vectors $\boldsymbol{\partial}^{(b)}$'s and the matrix $\boldsymbol{\partial}$ as constants is guaranteed), where

$$\begin{aligned}\Xi(\boldsymbol{\partial}^{(1:r)}, \boldsymbol{\partial}) &= E\exp\left[\sum_{b=1}^{r}\mathbf{z}_{U^\dagger}^{(b)\prime}\boldsymbol{\partial}^{(b)} + \mathrm{tr}\{(\widehat{\boldsymbol{\Sigma}}_{U^\dagger} - \boldsymbol{\Sigma})\boldsymbol{\partial}\}\right]\\ &= E\exp\left[\sum_{a=1}^{q}\mathbf{u}_x^{\dagger(a)\prime}\boldsymbol{\partial}^{[a]} + \frac{1}{N}\sum_{i=1}^{N}\mathrm{tr}\left\{(\mathbf{u}_i^\dagger\mathbf{u}_i^{\dagger\prime} - \boldsymbol{\Sigma})\frac{\boldsymbol{\partial}}{1-q/N}\right\} - \sum_{a=1}^{q}\mathrm{tr}\left\{(\mathbf{u}_x^{\dagger(a)}\mathbf{u}_x^{\dagger(a)\prime} - \boldsymbol{\Sigma})\frac{\boldsymbol{\partial}}{N-q}\right\}\right]\end{aligned}$$

with $\boldsymbol{\partial}^{[a]} = (\partial_j^{[a]}) = \sum_{b=1}^{r}\dot{v}_a^{(b)}\boldsymbol{\partial}^{(b)}$ $(a = 1, \ldots, q)$. Now, let us consider

$$\begin{aligned}M(\mathbf{t}^{[1:q]}, \mathbf{T}) &= E\exp\left[\sum_{a=1}^{q}\mathbf{u}_x^{\dagger(a)\prime}\mathbf{t}^{[a]} + \frac{1}{N}\sum_{i=1}^{N}\mathrm{tr}\left\{(\mathbf{u}_i^\dagger\mathbf{u}_i^{\dagger\prime} - \boldsymbol{\Sigma})\frac{\mathbf{T}}{1-q/N}\right\}\right]\\ &= \prod_{i=1}^{N}E\exp\left[\frac{\mathbf{u}_i^{\dagger\prime}}{N^{1/2}}\sum_{a=1}^{q}(N^{1/2}\mathcal{X}_a^{(i)})\mathbf{t}^{[a]} + \frac{1}{N}\mathrm{tr}\left\{(\mathbf{u}_i^\dagger\mathbf{u}_i^{\dagger\prime} - \boldsymbol{\Sigma})\frac{\mathbf{T}}{1-q/N}\right\}\right],\end{aligned}$$

where $\mathbf{t}^{[1:q]} = [\mathbf{t}^{[1]}, \ldots, \mathbf{t}^{[q]}]$. Then, we have

$$\Xi(\boldsymbol{\partial}^{(1:r)}, \boldsymbol{\partial}) = \exp\left[-\sum_{a=1}^{q}\mathrm{tr}\left\{\left(\frac{\partial^2}{\partial\mathbf{t}^{[a]}\partial\mathbf{t}^{[a]\prime}} - \boldsymbol{\Sigma}\right)\frac{\mathbf{T}}{N-q}\right\}\right]M(\mathbf{t}^{[1:q]}, \mathbf{T})|_{t^{[1:q]}=\partial^{[1:q]},\mathrm{T}=\partial}, \tag{16}$$

where we take the expectation regarding the vectors $\boldsymbol{\partial}^{[a]}$'s and the matrix $\boldsymbol{\partial}$ as constants $\mathbf{t}^{[a]} = (t_j^{[a]})$'s and $\mathbf{T} = (\mathrm{T}_{jk}) = ((1/2)(1+\delta_{jk})t_{jk})$, with $t_{jk} = t_{kj}$, and interchange the order of differentiation and integration (this is also allowable due to the truncation).

In order to obtain the expansion of (16), we define

$$G(\mathbf{t}, \mathbf{T}) = \log E\exp\left[\mathbf{u}^{\dagger\prime}\frac{\mathbf{t}}{N^{1/2}} + \mathrm{tr}\left\{(\mathbf{u}^\dagger\mathbf{u}^{\dagger\prime} - \boldsymbol{\Sigma})\frac{\mathbf{T}}{N}\right\}\right].$$

Now, for each $\mathbf{t}$ and $\mathbf{T}$, $NG(\mathbf{t}, \mathbf{T})$ can be arranged according to powers of $N^{-1/2}$ by using the joint cumulants of $\mathbf{u}^{\dagger\prime}\mathbf{t}$ and $\mathrm{tr}\{(\mathbf{u}^\dagger\mathbf{u}^{\dagger\prime} - \boldsymbol{\Sigma})\mathbf{T}\}$ (see [36]). Letting

$$\mathbf{t}^{(i)} = (t_j^{(i)}) = \sum_{a=1}^{q}(N^{1/2}\mathcal{X}_a^{(i)})\mathbf{t}^{[a]},$$

we have $N^{-1}\sum_{i=1}^{N}t_{j_1}^{(i)}\cdots t_{j_s}^{(i)} = \sum_{a_1\ldots a_s=1}^{q}N^{s/2}\overline{\mathcal{X}}_{a_1\ldots a_s}t_{j_1}^{[a_1]}\cdots t_{j_s}^{[a_s]}$ $(s \in \mathbf{N})$. Recalling (C_2)(ii) and using Lemmas 4 and 5 of Kakizawa and Iwashita [36], it is straightforward to see that

$$
\begin{aligned}
\log M(\mathbf{t}^{[1:q]}, \mathbf{T}) &= \sum_{i=1}^{N} G\left(\mathbf{t}^{(i)}, \frac{\mathbf{T}}{1-q/N}\right) \\
&= \frac{1}{2}\sum_{a=1}^{q} \mathbf{t}^{[a]\prime}\boldsymbol{\Sigma}\mathbf{t}^{[a]} + \frac{1}{N^{1/2}} \sum_{j_1j_2j_3=1}^{p} \kappa_{j_1,j_2,j_3}\left(\sum_{a_1=1}^{q} w_{a_1} t_{j_1}^{[a_1]}\mathrm{T}_{j_2j_3} + \frac{1}{6}\sum_{a_1a_2a_3=1}^{q} w_{a_1a_2a_3} t_{j_1}^{[a_1]}t_{j_2}^{[a_2]}t_{j_3}^{[a_3]}\right) \\
&\quad + \frac{1}{N}\left[\sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4}\frac{1}{2}\left(\mathrm{T}_{j_1j_2}\mathrm{T}_{j_3j_4} + \sum_{a=1}^{q} t_{j_1}^{[a]}t_{j_2}^{[a]}\mathrm{T}_{j_3j_4} + \frac{1}{12}\sum_{a_1a_2a_3a_4=1}^{q} w_{a_1a_2a_3a_4} t_{j_1}^{[a_1]}t_{j_2}^{[a_2]}t_{j_3}^{[a_3]}t_{j_4}^{[a_4]}\right)\right. \\
&\quad \left. + \operatorname{tr}(\boldsymbol{\Sigma}\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}) + \sum_{a=1}^{q}\mathbf{t}^{[a]\prime}\boldsymbol{\Sigma}\mathbf{T}\boldsymbol{\Sigma}\mathbf{t}^{[a]}\right] + \mathrm{o}(N^{-1}).
\end{aligned}
$$

The differential operator Ξ follows from (16), together with (14) and (15). □

4.2. *Outline of proof of* Theorem 1

We notice that with $\mathbf{z}_U^{(1:q)} = [\mathbf{z}_U^{(1)}, \ldots, \mathbf{z}_U^{(q)}] = \mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2}\dot{\mathbf{V}}$,

$$
\begin{aligned}
N^{1/2}\operatorname{vec}(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}_U^{(1:q)}) &= \sum_{i=1}^{N}\operatorname{vec}\{\boldsymbol{\Sigma}^{-1/2}\mathbf{u}_i(N^{1/2}\boldsymbol{\mathcal{X}}^{(i)\prime}\dot{\mathbf{V}})\} \\
&= \sum_{i=1}^{N}\{\dot{\mathbf{V}}(N^{1/2}\boldsymbol{\mathcal{X}}^{(i)}) \otimes \boldsymbol{\Sigma}^{-1/2}\}\mathbf{u}_i
\end{aligned}
$$

and

$$
\widehat{\boldsymbol{\Sigma}}_U - \boldsymbol{\Sigma} = \frac{N}{N-q}\left\{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{u}_i\mathbf{u}_i' - \boldsymbol{\Sigma})\right\} - \frac{1}{N-q}\sum_{a=1}^{q}(\mathbf{z}_U^{(a)}\mathbf{z}_U^{(a)\prime} - \boldsymbol{\Sigma}).
$$

If the smallest eigenvalue of the averaged covariance matrix $\mathrm{Cov}_N^x \equiv N^{-1}\sum_{i=1}^{N}\mathrm{Cov}(\widetilde{\mathbf{u}}_i^x)$ of the sum of the independent random vectors

$$
\widetilde{\mathbf{u}}_i^x = \begin{pmatrix} \{\dot{\mathbf{V}}(N^{1/2}\boldsymbol{\mathcal{X}}^{(i)}) \otimes \boldsymbol{\Sigma}^{-1/2}\}\mathbf{u}_i \\ \operatorname{vech}(\mathbf{u}_i\mathbf{u}_i' - \boldsymbol{\Sigma}) \end{pmatrix} \quad (i = 1, \ldots, N)
$$

is bounded away from zero, we can apply Bhattacharya and Rao's [6, Theorem 20.6] theory (their conditions can be verified in line with [54,62]), provided that (C_1)–(C_3) hold. Then, by following the usual BG transformation argument (e.g. [5,10, 4]), several smooth functions of $N^{-1}\sum_{i=1}^{N}\widetilde{\mathbf{u}}_i^x$ admit valid Edgeworth expansions up to order N^{-1} by means of appropriate transformations of a valid Edgeworth expansion of the distribution of $N^{-1/2}\sum_{i=1}^{N}\widetilde{\mathbf{u}}_i^x$. To complete the proof of Theorem 1, we need to (i) derive the stochastic expansion of $B_c(T_\psi) \approx \operatorname{tr}(\mathbf{s}_{\psi,c,\varepsilon}^{(1:r)}\mathbf{s}_{\psi,c,\varepsilon}^{(1:r)\prime})$, where

$$
\mathbf{s}_{\psi,c,\varepsilon}^{(1:r)} = \mathbf{s}_{\varepsilon}^{(1:r)} + \frac{\psi''}{4N}(\mathbf{s}_{\varepsilon}^{(1:r)}\mathbf{s}_{\varepsilon}^{(1:r)\prime})\mathbf{s}_{\varepsilon}^{(1:r)} - \frac{1}{N}\left[\sum_{j=1}^{3} c_j\{\operatorname{tr}(\mathbf{s}_{\varepsilon}^{(1:r)}\mathbf{s}_{\varepsilon}^{(1:r)\prime})\}^{j-1}\right]\mathbf{s}_{\varepsilon}^{(1:r)}
$$

with

$$
\begin{aligned}
\mathbf{s}_{\varepsilon}^{(1:r)} &= [\mathbf{s}_{\varepsilon}^{(1)}, \ldots, \mathbf{s}_{\varepsilon}^{(r)}] = \left(\mathbf{I}_p - \frac{1}{2}\widetilde{\boldsymbol{\Delta}}_U + \frac{3}{8}\widetilde{\boldsymbol{\Delta}}_U^2\right)\boldsymbol{\Sigma}^{-1/2}(\mathbf{z}_U^{(1:r)} + \dot{\boldsymbol{\delta}}_{\varepsilon}^{(1:r)}), \\
\widetilde{\boldsymbol{\Delta}}_U &= \boldsymbol{\Sigma}^{-1/2}(\widehat{\boldsymbol{\Sigma}}_U - \boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1/2} \quad \text{and} \quad \dot{\boldsymbol{\delta}}_{\varepsilon}^{(1:q)} = [\dot{\boldsymbol{\delta}}_{\varepsilon}^{(1)}, \ldots, \dot{\boldsymbol{\delta}}_{\varepsilon}^{(q)}] = \boldsymbol{\Theta}_{\varepsilon}'\dot{\mathbf{V}},
\end{aligned}
$$

(ii) check Chibisov's lemma (see [15]) and (iii) evaluate the asymptotic expansion for the characteristic function of $\operatorname{tr}(\mathbf{s}_{\psi,c,\varepsilon}^{(1:r)}\mathbf{s}_{\psi,c,\varepsilon}^{(1:r)\prime})$ via the differential operator given in Lemma 5 (the details are omitted, since these steps can be carried out similarly as in [37] for the one-way MANOVA model with $r = q - 1$).

Now, it is easy to see that Cramér's condition (1) in (C_3) implies the covariance matrix of $\widetilde{\mathbf{u}} = (\mathbf{u}', \{\operatorname{vech}(\mathbf{u}\mathbf{u}' - \boldsymbol{\Sigma})\}')'$;

$$
\mathrm{Cov}(\widetilde{\mathbf{u}}) = \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}
$$

is positive definite. So, with $\mathbf{L}_{11,i} = \dot{\mathbf{V}}(N^{1/2}\mathcal{X}^{(i)}) \otimes \boldsymbol{\Sigma}^{-1/2}$, we have

$$\begin{aligned}
\mathrm{Cov}_N^x &= \frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} \mathbf{L}_{11,i}\boldsymbol{\Sigma}\mathbf{L}_{11,i}' & \mathbf{L}_{11,i}\boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21}\mathbf{L}_{11,i}' & \boldsymbol{\Sigma}_{22}\end{pmatrix} \\
&= \sum_{i=1}^{N}\begin{pmatrix} \dot{\mathbf{V}}\mathcal{X}^{(i)}\mathcal{X}^{(i)\prime}\dot{\mathbf{V}}' \otimes \mathbf{I}_p & N^{-1/2}\dot{\mathbf{V}}\mathcal{X}^{(i)} \otimes \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}_{12} \\ N^{-1/2}\mathcal{X}^{(i)\prime}\dot{\mathbf{V}}' \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}^{-1/2} & N^{-1}\boldsymbol{\Sigma}_{22}\end{pmatrix} \\
&= \begin{pmatrix} \mathbf{I}_q \otimes \mathbf{I}_p & \text{symmetric} \\ N^{-1/2}\mathbf{1}_N'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2}\dot{\mathbf{V}}' \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}^{-1/2} & \boldsymbol{\Sigma}_{22}\end{pmatrix},
\end{aligned}$$

whose Schur complement of $\boldsymbol{\Sigma}_{22}$ (see [30, p100]) is equal to

$$\begin{aligned}
&\boldsymbol{\Sigma}_{22} - N^{-1}\mathbf{1}_N'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{1}_N\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{12} \\
&\quad = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{12} + N^{-1}\mathbf{1}_N'\{\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}\mathbf{1}_N\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{12}.
\end{aligned}$$

Hence, the smallest eigenvalue of Cov_N^x is bounded away from zero.

5. Extension to GMANOVA model

5.1. Introduction

We suppose that an $N \times p$ matrix $\mathbf{Y}$ consists of N independent observations $\mathbf{y}_1, \ldots, \mathbf{y}_N$ on p variables, where $\mathbf{Y}' = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$. The GMANOVA model (there are other names such as Generalized Multivariate Analysis of Variance model, Generalized MANOVA model, generalized linear model, Growth Curve Model (GCM) and Potthoff and Roy model [53]) is defined by

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Xi}\mathbf{A} + \mathbf{U},$$

where $\mathbf{X}$ is an $N \times q$ non-random between-individuals design matrix of rank q $(< N)$, $\mathbf{A}$ is an $m \times p$ non-random within-individuals design matrix of rank m $(\le p)$, $\boldsymbol{\Xi}$ is a $q \times m$ unknown parameter matrix, and $\mathbf{U}$ is an $N \times p$ unobservable random matrix with $\mathbf{U}' = [\mathbf{u}_1, \ldots, \mathbf{u}_N]$. It is assumed that each $p \times 1$ vector $\mathbf{u}_i$ is independently and identically distributed with mean vector $\mathbf{0}$ and unknown positive definite covariance matrix $\boldsymbol{\Sigma}$.

Gleser and Olkin [28] used the canonical reduction method (there are many other derivations; we also refer to [60], [43, subsection 4.1.2]) to obtain the maximum likelihood (ML) estimator of $\boldsymbol{\Xi}$ and $\boldsymbol{\Sigma}$ in the normal GMANOVA model (that is, $\mathbf{u}_1, \ldots, \mathbf{u}_N$ are independent p-variate normals), given by

$$\begin{aligned}
\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\mathbf{E}_Y^{-1}\mathbf{A}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}, \\
\widehat{\boldsymbol{\Sigma}}_{Y,\mathrm{ML}} &= \frac{1}{N}(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{A})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{A}) = \frac{1}{N}(\mathbf{E}_Y + \mathbf{E}_{1Y})
\end{aligned}$$

with

$$\mathbf{E}_{1Y} = \{\mathbf{I}_p - \mathbf{E}_Y^{-1}\mathbf{A}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}\mathbf{A}\}'\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\{\mathbf{I}_p - \mathbf{E}_Y^{-1}\mathbf{A}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}\mathbf{A}\}.$$

For testing a general linear hypothesis $H : \mathbf{B}\boldsymbol{\Xi}\mathbf{C} = \mathbf{D}$, where $\mathbf{B}$ is an $r \times q$ known matrix of rank r $(\le q)$, $\mathbf{C}$ is an $m \times s$ known matrix of rank s $(\le m)$ and $\mathbf{D}$ is an $r \times s$ known matrix, let $\mathbf{H}_Y^\circ$, $\mathbf{E}_Y^\circ$ and $\mathbf{H}_Y^\bullet$ be $s \times s$ matrices defined by

$$\begin{aligned}
\mathbf{H}_Y^\circ &= (\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D})'(\mathbf{B}\mathbf{R}_Y\mathbf{B}')^{-1}(\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D}), \\
\mathbf{E}_Y^\circ &= \mathbf{C}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}\mathbf{C}, \\
\mathbf{H}_Y^\bullet &= (\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D})'\{\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\}^{-1}(\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D}),
\end{aligned}$$

respectively, where

$$\mathbf{R}_Y = (\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\{\mathbf{E}_Y^{-1} - \mathbf{E}_Y^{-1}\mathbf{A}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}\mathbf{A}\mathbf{E}_Y^{-1}\}\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}.$$

Then, the following four criteria have been used under normality (e.g. [41,28,42]):

(i) Likelihood ratio (LR) $T_{LR}^\circ = -(N-q)\log|\mathbf{E}_Y^\circ|/|\mathbf{E}_Y^\circ + \mathbf{H}_Y^\circ|$,

(ii) Lawley–Hotelling's trace (T_0^2) $T_{LH}^\circ = (N-q)\mathrm{tr}[\mathbf{H}_Y^\circ(\mathbf{E}_Y^\circ)^{-1}]$,

(iii) Bartlett–Nanda–Pillai's trace $T_{BNP}^\circ = (N-q)\mathrm{tr}[\mathbf{H}_Y^\circ(\mathbf{E}_Y^\circ + \mathbf{H}_Y^\circ)^{-1}]$, and

(iv) Kleinbaum's Wald statistic $T_W^\bullet = (N-q)\mathrm{tr}[\mathbf{H}_Y^\bullet(\mathbf{E}_Y^\circ)^{-1}]$ (it is worth noting that using the identity $\mathbf{C}'\{\mathbf{A}(N\widehat{\boldsymbol{\Sigma}}_{Y,\mathrm{ML}})^{-1}\mathbf{A}'\}^{-1}\mathbf{C} = \mathbf{E}_Y^\circ$, his statistic $T_W^\bullet$ is proportional to the (original) Wald statistic $W = \mathrm{tr}[\mathbf{H}_Y^\bullet\{\mathbf{C}'(\mathbf{A}\widehat{\boldsymbol{\Sigma}}_{Y,\mathrm{ML}}^{-1}\mathbf{A}')^{-1}\mathbf{C}\}^{-1}]$ in the normal GMANOVA model, based on the unrestricted maximum likelihood estimator $\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}$ and $\widehat{\boldsymbol{\Sigma}}_{Y,\mathrm{ML}}$, that is, $T_W^\bullet = (1 - q/N)W$).

Given an $r \times q$ matrix $\mathbf{B}$ of rank r and an $m \times s$ matrix $\mathbf{C}$ of rank s, the solution space $\mathcal{V}_{\mathrm{B,C}}$ of a linear system $\mathbf{B}\boldsymbol{\Xi}\mathbf{C} = \mathbf{O}_{r,s}$ is a subspace of the linear space $\mathcal{R}^{q\times m}$ of all $q \times m$ matrices. In what follows, let $\boldsymbol{\Xi}_0 = \mathbf{B}^{-}\mathbf{D}\mathbf{C}^{-}$ be any particular solution to a system $\mathbf{B}\boldsymbol{\Xi}\mathbf{C} = \mathbf{D}$ in $\boldsymbol{\Xi}$ (e.g. [30, p157]). We are now interested in studying the nonnull distribution of the above test statistics when the parameter matrix is given by

$$\boldsymbol{\Xi} = \boldsymbol{\Xi}_0 + (\mathbf{X}'\mathbf{X})^{-1/2}\boldsymbol{\Xi}_\varepsilon \quad \text{with } (\mathbf{X}'\mathbf{X})^{-1/2}\boldsymbol{\Xi}_\varepsilon \notin \mathcal{V}_{\mathrm{B,C}} \tag{17}$$

(we always assume that $\boldsymbol{\Xi}_\varepsilon$ is a $q \times m$ matrix, independent of N). In that case, it is easy to see that $\mathbf{E}_Y^\circ/(N-q) = \mathbf{C}'(\mathbf{A}\widehat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{A}')^{-1}\mathbf{C}$,

$$\begin{aligned}
\mathbf{R}_Y^{(d)} &\equiv (\mathbf{X}'\mathbf{X})^{-1} + d(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\{\mathbf{E}_Y^{-1} - \mathbf{E}_Y^{-1}\mathbf{A}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}\mathbf{A}\mathbf{E}_Y^{-1}\}\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\\
&= (\mathbf{X}'\mathbf{X})^{-1} + \frac{d}{N-q}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}\{\widehat{\boldsymbol{\Sigma}}_U^{-1} - \widehat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{A}'(\mathbf{A}\widehat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{A}')^{-1}\mathbf{A}\widehat{\boldsymbol{\Sigma}}_U^{-1}\}\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\\
&= \widetilde{\mathbf{R}}_U^{(d)} \quad \text{(say)}
\end{aligned}$$

for a nonnegative constant d, and

$$\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D} = \mathbf{B}(\mathbf{X}'\mathbf{X})^{-1/2}\{(\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{X}'\mathbf{U} + \boldsymbol{\Xi}_\varepsilon\mathbf{A}\}\widehat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{A}'(\mathbf{A}\widehat{\boldsymbol{\Sigma}}_U^{-1}\mathbf{A}')^{-1}\mathbf{C}.$$

Then, Lawley–Hotelling's trace and Kleinbaum's Wald statistic are written as

$$\begin{aligned}
T_{LH}^\circ &= \mathrm{tr}\left[\mathbf{H}_Y\left(\frac{\mathbf{E}_Y^\circ}{N-q}\right)^{-1}\right]\\
&= [\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + (\boldsymbol{\Theta}_\varepsilon^\circ)'\}]'(\ddot{\mathbf{M}}_U \otimes \mathbf{Q}^{\widehat{\Sigma}_U})\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + (\boldsymbol{\Theta}_\varepsilon^\circ)'\},\\
T_W^\bullet &= \mathrm{tr}\left[\mathbf{H}_Y^\bullet\left(\frac{\mathbf{E}_Y^\circ}{N-q}\right)^{-1}\right]\\
&= [\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + (\boldsymbol{\Theta}_\varepsilon^\circ)'\}]'(\dot{\mathbf{M}} \otimes \mathbf{Q}^{\widehat{\Sigma}_U})\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + (\boldsymbol{\Theta}_\varepsilon^\circ)'\},
\end{aligned}$$

where we define

$$\begin{aligned}
&\boldsymbol{\Theta}_\varepsilon^\circ = \boldsymbol{\Xi}_\varepsilon\mathbf{A}\,(q \times p \text{ matrix}), \qquad \ddot{\mathbf{M}}_U = (\mathbf{X}'\mathbf{X})^{-1/2}\mathbf{B}'(\mathbf{B}\widetilde{\mathbf{R}}_U^{(1)}\mathbf{B}')^{-1}\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1/2},\\
&\mathbf{Q}^{\Sigma} = \boldsymbol{\Sigma}^{-1}\mathbf{A}'(\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{A}')^{-1}\mathbf{C}\{\mathbf{C}'(\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{A}')^{-1}\mathbf{C}\}^{-1}\mathbf{C}'(\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{A}')^{-1}\mathbf{A}\boldsymbol{\Sigma}^{-1}.
\end{aligned}$$

It is easy to see that the limiting nonnull distribution of generalized test statistic

$$T_\psi^\circ = (N-q)\sum_{j=1}^{s}\psi(\lambda_{Y,j}^\circ) \quad \text{or} \quad T_\psi^\bullet = (N-q)\sum_{j=1}^{s}\psi(\lambda_{Y,j}^\bullet)$$

($\lambda_{Y,1}^\circ, \ldots, \lambda_{Y,s}^\circ \geq 0$ are eigenvalues of $\mathbf{H}_Y^\circ(\mathbf{E}_Y^\circ)^{-1}$ and $\lambda_{Y,1}^\bullet, \ldots, \lambda_{Y,s}^\bullet \geq 0$ are eigenvalues of $\mathbf{H}_Y^\bullet(\mathbf{E}_Y^\circ)^{-1}$) is the same as that of

$$[\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + (\boldsymbol{\Theta}_\varepsilon^\circ)'\}]'(\mathbf{M} \otimes \mathbf{Q}^{\Sigma})\mathrm{vec}\{\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2} + (\boldsymbol{\Theta}_\varepsilon^\circ)'\},$$

which is asymptotically distributed as the noncentral chi-square distribution with $f_\circ = rs$ degrees of freedom and noncentrality parameter $\omega_\circ^2 = [\mathrm{vec}\{(\boldsymbol{\Theta}_\varepsilon^\circ)'\}]'(\mathbf{M} \otimes \mathbf{Q}^{\Sigma})\mathrm{vec}\{(\boldsymbol{\Theta}_\varepsilon^\circ)'\}$, by noting that $\mathbf{Q}^{\Sigma}$ satisfies $\mathbf{Q}^{\Sigma}\boldsymbol{\Sigma}\mathbf{Q}^{\Sigma} = \mathbf{Q}^{\Sigma}$ and $\mathrm{tr}(\mathbf{Q}^{\Sigma}\boldsymbol{\Sigma}) = s$.

Remark 2. Define the $p \times q$ matrix $\boldsymbol{\Delta}^\circ = (\Delta_{ab}^\circ) = \mathbf{Q}^{\Sigma}(\boldsymbol{\Theta}_\varepsilon^\circ)'\mathbf{M}$. Then, using the fact that $\mathbf{M}$ is idempotent and noting $\mathbf{Q}^{\Sigma}\boldsymbol{\Sigma}\mathbf{Q}^{\Sigma} = \mathbf{Q}^{\Sigma}$, we have $\omega_\circ^2 = \mathrm{tr}\{\mathbf{Q}^{\Sigma}(\boldsymbol{\Theta}_\varepsilon^\circ)'\mathbf{M}\boldsymbol{\Theta}_\varepsilon^\circ\} = \mathrm{tr}(\boldsymbol{\Omega}_\circ)$, where $\boldsymbol{\Omega}_\circ = \boldsymbol{\Delta}^\circ(\boldsymbol{\Delta}^\circ)'\boldsymbol{\Sigma}$ ($p \times p$ matrix).

It is of interest to compare these tests by means of higher-order expansion. To avoid two separate analyses according to $T = T_\psi^\circ$ or $T = T_\psi^\bullet$, we will introduce a class of tests

$$T_{(\psi,d)} = (N-q)\sum_{j=1}^{s}\psi(\lambda_{Y,j}^{(d)}),$$

where $\lambda_{Y,1}^{(d)}, \ldots, \lambda_{Y,s}^{(d)} \geq 0$ are eigenvalues of $\mathbf{H}_Y^{(d)}(\mathbf{E}_Y^\circ)^{-1}$, with

$$\mathbf{H}_Y^{(d)} = (\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D})'(\mathbf{B}\mathbf{R}_Y^{(d)}\mathbf{B}')^{-1}(\mathbf{B}\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}}\mathbf{C} - \mathbf{D})$$

for a nonnegative constant d (obviously, we have $T_{(\psi,1)} = T_\psi^\circ$ and $T_{(\psi,0)} = T_\psi^\bullet$).

5.2. Asymptotic expansion in a nonnormal GMANOVA model

An asymptotic expansion for the nonnull distribution of $T^\circ = T_{LR}^\circ, T_{LH}^\circ, T_{BNP}^\circ$ was first studied by Fujikoshi [22] under normality (we refer to [38] for the null distribution of $T_W^\bullet$). Under nonnormality, Yanagihara [64,66] gave an asymptotic expansion for the null distribution of $T^\circ = T^\circ_{LR}, T^\circ_{LH}, T^\circ_{BNP}$, by following Wakaki et al. [62]. Writing

$$\mathbf{P}^{\Sigma} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{A}'(\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{A}')^{-1}\mathbf{A}\boldsymbol{\Sigma}^{-1}$$

(we note $\mathbf{P}^{\Sigma} = \mathbf{O}_{p,p}$ whenever $\mathbf{A}$ is nonsingular), his null result was described in terms of

$$\begin{aligned}
K_4^{\mathrm{QQ}} &= \sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4} Q^{\Sigma}_{j_1j_2} Q^{\Sigma}_{j_3j_4},\\
K_{33,J}^{\mathrm{QQQ}} &= K_{33,J}(\mathbf{Q}^{\Sigma}, \mathbf{Q}^{\Sigma}, \mathbf{Q}^{\Sigma}) \quad (J = 1, 2),\\
K_{33,J}^{\mathrm{QQP}} &= K_{33,J}(\mathbf{Q}^{\Sigma}, \mathbf{Q}^{\Sigma}, \mathbf{P}^{\Sigma}) \quad (J = 1, 2), \qquad K_{33,2}^{\mathrm{QPQ}} = K_{33,2}(\mathbf{Q}^{\Sigma}, \mathbf{P}^{\Sigma}, \mathbf{Q}^{\Sigma}),\\
K_{33,2}^{\mathrm{PQP}} &= K_{33,2}(\mathbf{P}^{\Sigma}, \mathbf{Q}^{\Sigma}, \mathbf{P}^{\Sigma}), \qquad K_{33,J}^{\mathrm{QPP}} = K_{33,J}(\mathbf{Q}^{\Sigma}, \mathbf{P}^{\Sigma}, \mathbf{P}^{\Sigma}) \quad (J = 1, 2),
\end{aligned} \tag{18}$$

where we introduce the notation

$$\begin{aligned}
K_{33,1}(\mathbf{P}^{(1)}, \mathbf{P}^{(2)}, \mathbf{P}^{(3)}) &= \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6} P^{(1)}_{j_1j_4} P^{(2)}_{j_2j_5} P^{(3)}_{j_3j_6},\\
K_{33,2}(\mathbf{P}^{(1)}, \mathbf{P}^{(2)}, \mathbf{P}^{(3)}) &= \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6} P^{(1)}_{j_1j_2} P^{(2)}_{j_3j_4} P^{(3)}_{j_5j_6}
\end{aligned}$$

for any $p \times p$ symmetric matrix $\mathbf{P}^{(i)} = (P^{(i)}_{j_1j_2})$ $(i = 1, 2, 3)$. Our contribution is to obtain an asymptotic expansion for the nonnull distribution of $T_{(\psi,d)}$ up to order N^{-1} and then study the local power properties of tests after Bartlett's type adjustment or Cornish–Fisher's type size adjustment.

Recall that $s \times s$ matrices $\mathbf{H}_Y^{(d)}$ and $\mathbf{E}_Y^\circ$ are functions of $\mathbf{U}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1/2}$ and $\widehat{\boldsymbol{\Sigma}}_U$. As in Section 4.2, our routine is to (i) derive the stochastic expansion $B_c(T_{(\psi,d)}) \approx \mathrm{tr}\{(\bar{\mathbf{s}}^{(1:r)}_{(\psi,d),c,\mathcal{E}^\circ})'\widetilde{\mathbf{Q}}^{\Sigma}\bar{\mathbf{s}}^{(1:r)}_{(\psi,d),c,\mathcal{E}^\circ}\}$ as a functional of $\mathbf{z}_U^{(1:r)}$ and $\widehat{\boldsymbol{\Sigma}}_U$, (ii) check Chibisov's lemma and (iii) evaluate the asymptotic expansion for the characteristic function of (i) as an application of the differential operator of Lemma 5 (for the GMANOVA case, Lemmas A.2 and A.3 in Appendix A.2 provide basic tools for several patterned derivatives), where

$$\begin{aligned}
\bar{\mathbf{s}}^{(1:r)}_{(\psi,d),c,\mathcal{E}^\circ} = \Bigg[&\mathbf{I}_p - \widetilde{\boldsymbol{\Delta}}_U\left(\widetilde{\mathbf{P}}^{\Sigma} + \frac{1}{2}\widetilde{\mathbf{Q}}^{\Sigma}\right) + \left\{\widetilde{\boldsymbol{\Delta}}_U\left(\widetilde{\mathbf{P}}^{\Sigma} + \frac{1}{2}\widetilde{\mathbf{Q}}^{\Sigma}\right)\right\}^2 + \frac{1}{8}(\widetilde{\boldsymbol{\Delta}}_U\widetilde{\mathbf{Q}}^{\Sigma})^2\\
&- \frac{d}{2N}(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ})(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ})'\widetilde{\mathbf{P}}^{\Sigma} + \frac{\psi''}{4N}(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ})(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ})'\widetilde{\mathbf{Q}}^{\Sigma}\\
&- \frac{1}{N}\sum_{j=1}^{3} c_j[\mathrm{tr}\{(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ})'\widetilde{\mathbf{Q}}^{\Sigma}(\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ})\}]^{j-1}\mathbf{I}_p\Bigg](\boldsymbol{\Sigma}^{-1/2}\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ}),
\end{aligned}$$

with $\widetilde{\mathbf{P}}^{\Sigma} = \boldsymbol{\Sigma}^{1/2}\mathbf{P}^{\Sigma}\boldsymbol{\Sigma}^{1/2}$, $\widetilde{\mathbf{Q}}^{\Sigma} = \boldsymbol{\Sigma}^{1/2}\mathbf{Q}^{\Sigma}\boldsymbol{\Sigma}^{1/2}$ and $\mathbf{z}^{(1:r)}_{U,\dot{\mathcal{E}}^\circ} = \mathbf{z}_U^{(1:r)} + \dot{\boldsymbol{\delta}}_{\mathcal{E}}^{\circ(1:r)}$ (we now set $\dot{\boldsymbol{\delta}}_{\mathcal{E}}^{\circ(1:q)} = [\dot{\boldsymbol{\delta}}_{\mathcal{E}}^{\circ(1)}, \ldots, \dot{\boldsymbol{\delta}}_{\mathcal{E}}^{\circ(q)}] = (\boldsymbol{\Theta}_{\mathcal{E}}^\circ)'\dot{\mathbf{V}}$ as compared with the notation in Section 4.2).

We are now in a position to state one of the main results of Part II: Suppose that (C_1)–(C_4) hold. Then, the distribution function of $B_c(T_{(\psi,d)})$ under the local alternative (17) admits an asymptotic expansion

$$\begin{aligned}
\Pr[B_c(T_{(\psi,d)}) \le x] = {}& G_{f_\circ}(x; \omega_\circ^2) + \frac{\mathrm{tr}(\boldsymbol{\Omega}_{\circ 1})}{2N}\{G_{f_\circ+2}(x; \omega_\circ^2) - G_{f_\circ}(x; \omega_\circ^2)\}\\
&+ \sum_{k=1}^{2}\frac{1}{N^{k/2}}\sum_{\ell=0}^{3k}\pi^\circ_{k,\ell}\, G_{f_\circ+2\ell}(x; \omega_\circ^2) + \frac{\psi''}{4N}\sum_{\ell=1}^{4}\pi^\circ_{\ell}\, G_{f_\circ+2\ell}(x; \omega_\circ^2)\\
&+ \frac{d}{2N}\sum_{\ell=0}^{2}\widetilde{\pi}^\circ_{\ell}\, G_{f_\circ+2\ell}(x; \omega_\circ^2) + \frac{1}{N}\sum_{\ell=0}^{6}\pi^{\circ c}_{\ell}\, G_{f_\circ+2\ell}(x; \omega_\circ^2) + \mathrm{o}(N^{-1}),
\end{aligned} \tag{19}$$

where

$$\begin{aligned}
&\pi_1^\circ = -f_\circ(s + r + 1), \qquad \pi_2^\circ = f_\circ(s + r + 1) - 2(s + r + 1)\mathrm{tr}(\boldsymbol{\Omega}_\circ),\\
&\pi_3^\circ = 2(s + r + 1)\mathrm{tr}(\boldsymbol{\Omega}_\circ) - \mathrm{tr}(\boldsymbol{\Omega}_\circ^2), \qquad \pi_4^\circ = \mathrm{tr}(\boldsymbol{\Omega}_\circ^2),\\
&\widetilde{\pi}_0^\circ = f_\circ(p - m), \qquad \widetilde{\pi}_1^\circ = -f_\circ(p - m) + (p - m)\mathrm{tr}(\boldsymbol{\Omega}_\circ), \qquad \widetilde{\pi}_2^\circ = -(p - m)\mathrm{tr}(\boldsymbol{\Omega}_\circ),\\
&\pi_0^{\circ c} = f_\circ c_1, \qquad \pi_1^{\circ c} = -f_\circ c_1 + f_\circ(f_\circ + 2)c_2 + c_1\omega_\circ^2,\\
&\pi_2^{\circ c} = -f_\circ(f_\circ + 2)c_2 + f_\circ(f_\circ + 2)(f_\circ + 4)c_3 + \{-c_1 + 2(f_\circ + 2)c_2\}\omega_\circ^2,
\end{aligned}$$

$$\pi_3^{\circ c} = -f_\circ(f_\circ+2)(f_\circ+4)c_3 + \{-2(f_\circ+2)c_2 + 3(f_\circ+2)(f_\circ+4)c_3\}\omega_\circ^2 + c_2\omega_\circ^4,$$
$$\pi_4^{\circ c} = -3(f_\circ+2)(f_\circ+4)c_3\omega_\circ^2 + \{-c_2 + 3(f_\circ+4)c_3\}\omega_\circ^4,$$
$$\pi_5^{\circ c} = -3(f_\circ+4)c_3\omega_\circ^4 + c_3\omega_\circ^6, \qquad \pi_6^{\circ c} = -c_3\omega_\circ^6,$$
$$\boldsymbol{\Omega}_{\circ 1} = \mathbf{Q}^{\Sigma}(\boldsymbol{\Theta}_\varepsilon^\circ)'\mathbf{M}_1\boldsymbol{\Theta}_\varepsilon^\circ \quad (p\times p \text{ matrix}).$$

Here, as in the case of Theorem 1, the coefficients $\pi_{k,\ell}^\circ$'s, independent of ψ, d, are the sums of homogeneous polynomials of degrees 0, 1, 2, 3, 4 and 6 in $\boldsymbol{\Xi}_\varepsilon$ (the explicit expressions are omitted to preserve space, but the details can be obtained from the author on request), depending on (18) and

$$K_{\circ 3}^{\mathrm{Q}[a_3]} = K_{\circ 3}^{[a_3]}(\mathbf{Q}^{\Sigma}), \qquad K_{\circ 3}^{\mathrm{P}[a_3]} = K_{\circ 3}^{[a_3]}(\mathbf{P}^{\Sigma}),$$
$$K_{\circ 3}^{[a_1a_2a_3]} = \sum_{j_1j_2j_3=1}^{p} \kappa_{j_1,j_2,j_3}\Delta_{j_1a_1}^\circ\Delta_{j_2a_2}^\circ\Delta_{j_3a_3}^\circ,$$
$$K_{\circ 4}^{\mathrm{Q}[a_3a_4]} = K_{\circ 4}^{[a_3a_4]}(\mathbf{Q}^{\Sigma}), \qquad K_{\circ 4}^{\mathrm{P}[a_3a_4]} = K_{\circ 4}^{[a_3a_4]}(\mathbf{P}^{\Sigma}),$$
$$K_{\circ 4}^{[a_1a_2a_3a_4]} = \sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4}\Delta_{j_1a_1}^\circ\Delta_{j_2a_2}^\circ\Delta_{j_3a_3}^\circ\Delta_{j_4a_4}^\circ,$$
$$K_{\circ 33,1}^{\mathrm{QQ}[a_3a_6]} = K_{\circ 33,1}^{[a_3a_6]}(\mathbf{Q}^{\Sigma},\mathbf{Q}^{\Sigma}), \qquad K_{\circ 33,1}^{\mathrm{PP}[a_3a_6]} = K_{\circ 33,1}^{[a_3a_6]}(\mathbf{P}^{\Sigma},\mathbf{P}^{\Sigma}),$$
$$K_{\circ 33,2}^{\mathrm{QQ}[a_5a_6]} = K_{\circ 33,2}^{[a_5a_6]}(\mathbf{Q}^{\Sigma},\mathbf{Q}^{\Sigma}), \qquad K_{\circ 33,2}^{\mathrm{PP}[a_5a_6]} = K_{\circ 33,2}^{[a_5a_6]}(\mathbf{P}^{\Sigma},\mathbf{P}^{\Sigma}),$$
$$K_{\circ 33}^{\mathrm{Q}[a_2a_3a_5a_6]} = K_{\circ 33}^{[a_2a_3a_5a_6]}(\mathbf{Q}^{\Sigma}), \qquad K_{\circ 33}^{\mathrm{P}[a_2a_3a_5a_6]} = K_{\circ 33}^{[a_2a_3a_5a_6]}(\mathbf{P}^{\Sigma}),$$
$$K_{\circ 33,1}^{\mathrm{QP}[a_3a_6]} = K_{\circ 33,1}^{[a_3a_6]}(\mathbf{Q}^{\Sigma},\mathbf{P}^{\Sigma}),$$
$$K_{\circ 33,2}^{\mathrm{QP}[a_5a_6]} = K_{\circ 33,2}^{[a_5a_6]}(\mathbf{Q}^{\Sigma},\mathbf{P}^{\Sigma}), \qquad K_{\circ 33,2}^{\mathrm{PQ}[a_5a_6]} = K_{\circ 33,2}^{[a_5a_6]}(\mathbf{P}^{\Sigma},\mathbf{Q}^{\Sigma})$$

for $a_1,\ldots,a_6 \in \{1,\ldots,q\}$ (these quantities except $K_{\circ 4}^{\mathrm{P}[a_3a_4]}$ are obtained from (18) when at least one $P_{jj'}^{\Sigma}$ or $Q_{jj'}^{\Sigma}$ is replaced by $\Delta_{ja}^\circ\Delta_{j'a'}^\circ$), where we introduce the notation

$$K_{\circ 3}^{[a_3]}(\mathbf{P}^{(1)}) = \sum_{j_1j_2j_3=1}^{p} \kappa_{j_1,j_2,j_3}P_{j_1j_2}^{(1)}\Delta_{j_3a_3}^\circ,$$
$$K_{\circ 4}^{[a_3a_4]}(\mathbf{P}^{(1)}) = \sum_{j_1j_2j_3j_4=1}^{p} \kappa_{j_1,j_2,j_3,j_4}P_{j_1j_2}^{(1)}\Delta_{j_3a_3}^\circ\Delta_{j_4a_4}^\circ,$$
$$K_{\circ 33,1}^{[a_3a_6]}(\mathbf{P}^{(1)},\mathbf{P}^{(2)}) = \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}P_{j_1j_4}^{(1)}P_{j_2j_5}^{(2)}\Delta_{j_3a_3}^\circ\Delta_{j_6a_6}^\circ,$$
$$K_{\circ 33,2}^{[a_5a_6]}(\mathbf{P}^{(1)},\mathbf{P}^{(2)}) = \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}P_{j_1j_2}^{(1)}P_{j_3j_4}^{(2)}\Delta_{j_5a_5}^\circ\Delta_{j_6a_6}^\circ,$$
$$K_{\circ 33}^{[a_2a_3a_5a_6]}(\mathbf{P}^{(1)}) = \sum_{j_1j_2j_3j_4j_5j_6=1}^{p} \kappa_{j_1,j_2,j_3}\kappa_{j_4,j_5,j_6}P_{j_1j_4}^{(1)}\Delta_{j_2a_2}^\circ\Delta_{j_3a_3}^\circ\Delta_{j_5a_5}^\circ\Delta_{j_6a_6}^\circ$$

for any $p\times p$ symmetric matrix $\mathbf{P}^{(i)} = (P_{j_1j_2}^{(i)})$ $(i=1,2)$. Interestingly, we find that the local power of Bartlett's type or Cornish–Fisher's type adjusted $T_{(\psi,d_1)}$ test (shortly $\mathrm{B}_{(\psi,d_1)}$ or $\mathrm{CF}_{(\psi,d_1)}$ test) up to order N^{-1} is the same as that of $\mathrm{B}_{(\psi,d_2)}$ or $\mathrm{CF}_{(\psi,d_2)}$ test; that is, letting

$$D_{s,r}^\circ(\boldsymbol{\Omega}_\circ) = \mathrm{tr}(\boldsymbol{\Omega}_\circ^2) - \frac{s+r+1}{sr+2}\{\mathrm{tr}(\boldsymbol{\Omega}_\circ)\}^2,$$

we have

$$\begin{aligned}
&\lim_{N\to\infty}\left\{N\left(\Pr[B_{\widehat{\vartheta}^\circ(\psi_2'',d_2)}(T_{(\psi_2,d_2)}) > \chi^2_{f_\circ,\alpha}] - \Pr[B_{\widehat{\vartheta}^\circ(\psi_1'',d_1)}(T_{(\psi_1,d_1)}) > \chi^2_{f_\circ,\alpha}]\right)\right\}\\
&= \lim_{N\to\infty}\left\{N\left(\Pr[T_{(\psi_2,d_2)} > CF_{\widehat{\vartheta}^\circ(\psi_2'',d_2)}(\chi^2_{f_\circ,\alpha})] - \Pr[T_{(\psi_1,d_1)} > CF_{\widehat{\vartheta}^\circ(\psi_1'',d_1)}(\chi^2_{f_\circ,\alpha})]\right)\right\}\\
&= \frac{\psi_2''-\psi_1''}{2}D_{s,r}^\circ(\boldsymbol{\Omega}_\circ)g_{f_\circ+8}(\chi^2_{f_\circ,\alpha};\omega_\circ^2)
\end{aligned}$$

(these results essentially follow along the line of Corollaries 2–4, on the basis of the asymptotic expansion (19)), where $\widehat{\boldsymbol{\vartheta}}^{\circ}(\psi'', d) = (\widehat{\vartheta}_1^{\circ}(d), \widehat{\vartheta}_2^{\circ}(\psi''), \widehat{\vartheta}_3^{\circ})$ is a location invariant estimator of $\boldsymbol{\vartheta}^{\circ}(\psi'', d) = (\vartheta_1^{\circ}(d), \vartheta_2^{\circ}(\psi''), \vartheta_3^{\circ})$, with $\vartheta_1^{\circ}(d) = -\pi_{2,0}^{\circ[0]}/f_{\circ} - d(p-m)/2$,

$$\vartheta_2^{\circ}(\psi'') = -\frac{\pi_{2,0}^{\circ[0]} + \pi_{2,1}^{\circ[0]}}{f_{\circ}(f_{\circ}+2)} + \frac{(s+r+1)\psi''}{4(f_{\circ}+2)}, \qquad \vartheta_3^{\circ} = -\frac{\pi_{2,0}^{\circ[0]} + \pi_{2,1}^{\circ[0]} + \pi_{2,2}^{\circ[0]}}{f_{\circ}(f_{\circ}+2)(f_{\circ}+4)}$$

(as pointed out in Section 2.2, the coefficients of a_5 $(= \dot{A}_3^{\dagger}/8)$ in [62, (2.1)], hence [66, (2.5)] for the null distribution of $\{1-(p-m)/(N-q)\}T_{\psi}^{\circ}$ with $T_{\psi}^{\circ} = T_{LR}^{\circ}, T_{LH}^{\circ}, T_{BNP}^{\circ}$, are incorrect). We observe that $\pi_{2,0}^{\circ[0]}$, $\pi_{2,0}^{\circ[0]}+\pi_{2,1}^{\circ[0]}$ and $\sum_{\ell=0}^{2}\pi_{2,\ell}^{\circ[0]} = -\pi_{2,3}^{\circ[0]}$ are given by

$$\begin{aligned}
\pi_{2,0}^{\circ[0]} &= -\frac{f_{\circ}}{4}\{s-r+1+4(p-m)\} + \frac{A_0}{8}K_4^{QQ} \\
&\quad + \left(-\frac{A_1}{12} + \frac{A_3^{\dagger}}{4}r\right)K_{33,1}^{QQQ} + \left(-\frac{A_2}{8} - \frac{A_3^{\dagger}}{8}r^2 + \frac{A_4^{\dagger}}{4}r\right)K_{33,2}^{QQQ} \\
&\quad + \left\{\frac{A_3^{\dagger}}{2}(2r+1) - A_4^{\dagger}\right\}K_{33,1}^{QQP} + \left\{\frac{A_3^{\dagger}}{2}(r+1) - \frac{A_4^{\dagger}}{2}\right\}K_{33,2}^{QPQ} \\
&\quad + \left(\frac{A_3^{\dagger}}{2}r - \frac{A_4^{\dagger}}{2}\right)K_{33,2}^{QQP} - \frac{A_3^{\dagger}}{2}(3K_{33,1}^{QPP} + 2K_{33,2}^{QPP} + K_{33,2}^{PQP}), \\
\pi_{2,0}^{\circ[0]} + \pi_{2,1}^{\circ[0]} &= -\frac{f_{\circ}}{4}(s+r+1) - \frac{A_0}{8}K_4^{QQ} + \left(\frac{A_1}{6} - \frac{A_4^{\dagger}}{2}\right)K_{33,1}^{QQQ} + \left\{\frac{A_2}{4} + \frac{A_3^{\dagger}}{4}r(r+2) - \frac{A_4^{\dagger}}{2}(r+1)\right\}K_{33,2}^{QQQ} \\
&\quad + \{-A_3^{\dagger}(r+2) + A_4^{\dagger}\}\left(K_{33,1}^{QQP} + \frac{1}{2}K_{33,2}^{QPQ} + \frac{1}{2}K_{33,2}^{QQP}\right), \\
\sum_{\ell=0}^{2}\pi_{2,\ell}^{\circ[0]} &= -\pi_{2,3}^{\circ[0]} \\
&= \left\{-\frac{A_1}{12} - \frac{A_3^{\dagger}}{4}(r+2) + \frac{A_4^{\dagger}}{2}\right\}K_{33,1}^{QQQ} + \left\{-\frac{A_2}{8} - \frac{A_3^{\dagger}}{8}(r+2)^2 + \frac{A_4^{\dagger}}{4}(r+2)\right\}K_{33,2}^{QQQ}.
\end{aligned}$$

6. Simulation study

In this section, we present some results of simulation studies examining finite sample performance of original or Bartlett's type adjusted or Cornish–Fisher's type adjusted LR, LH, BNP tests for $H : \mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$ with $\mathbf{B} = [\mathbf{O}_{r,q-r}, \mathbf{I}_r]$ for some $r \in \{1, \dots, q-1\}$, assuming the multivariate linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{U}$ with $\mathbf{X} = [\mathbf{1}_N, \mathbf{X}_{(-1)}]$.

For the case of $(p, q, r) = (4, 3, 2)$, the polynomial regressor $\mathbf{x}_i = \mathbf{w}(i/N)$ $(i = 1, \dots, N)$ with $\mathbf{w}(t) = (1, t, \dots, t^{q-1})'$ (we now set $\mathbf{X}' = [\mathbf{x}_1, \dots, \mathbf{x}_N]$) and $N = 30, 60, 90, 120$, we generated $\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{U}$ with $\mathbf{U}' = [\mathbf{u}_1, \dots, \mathbf{u}_N]$, where $\mathbf{u}_1, \dots, \mathbf{u}_N$ were assumed to be independent and identically distributed according to

(CN) the contaminated normal $0.8N_p(\mathbf{0}, \mathbf{I}_p) + 0.2N_p(\mathbf{0}, 4\mathbf{I}_p)$,

(LN) the log-normal $\mathbf{u}_{\mathrm{LN}} = (u_{\mathrm{LN},1}, \dots, u_{\mathrm{LN},p})'$ with $u_{\mathrm{LN},i} = (\mathrm{e}^{z_i} - \mathrm{e}^{1/2})/\{\mathrm{e}^{1/2}(\mathrm{e}-1)\}^{1/2}$, where $\mathbf{z} = (z_1, \dots, z_p)' \sim N_p(\mathbf{0}, \mathbf{I}_p)$.

The parameter $\boldsymbol{\Theta}$ were chosen as follows: Under the null hypothesis $H : \mathbf{B}\boldsymbol{\Theta} = \mathbf{O}_{r,p}$, we took $\boldsymbol{\Theta} = \mathbf{O}_{q,p}$ due to the location invariance of the proposed tests. Under the alternative hypothesis, we considered the following two different types of the parameter $c\boldsymbol{\Theta}$, such that the $p \times p$ matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}(c\boldsymbol{\Theta})'\mathbf{B}'\{\mathbf{B}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{B}'\}^{-1}\mathbf{B}(c\boldsymbol{\Theta})$, with $\boldsymbol{\Sigma}$ being the population covariance matrix for (CN) or (LN), has eigenvalues $c^2\lambda_1, \dots, c^2\lambda_p$:

(A) $\lambda_1 = 1, \lambda_2 = \cdots = \lambda_p = 0$ $(D_{p,r}(\boldsymbol{\Omega}) > 0$ if $\min(p, r) > 1)$

(B) $\lambda_1 = \cdots = \lambda_{\min(p,r)} = 1, \lambda_{\min(p,r)+1} = \cdots = \lambda_p = 0$ $(D_{p,r}(\boldsymbol{\Omega}) < 0$ if $\min(p, r) > 1)$,

where $c > 0$ was determined so that the noncentrality parameter $\mathrm{tr}(\boldsymbol{\Omega}) = c^2\sum_{i=1}^{p}\lambda_i$ is equal to $5, 10, 15, 20$, respectively. The number of repetitions was set to be 1000,000.

With $f = pr$, the original large sample LR, LH, BNP tests have the rejection region $T_{LR} > \chi^2_{f,\alpha}$, $T_{LH} > \chi^2_{f,\alpha}$ and $T_{BNP} > \chi^2_{f,\alpha}$, respectively. Furthermore, using asymptotic expansions with $A_3^{\dagger} = A_4^{\dagger} = 0$, as in Example 2, Bartlett's type adjusted LR, LH, BNP tests (for shortly, B-adjusted (B-ad.)) and Cornish–Fisher's type adjusted LR, LH, BNP tests (for shortly, CF-adjusted (CF-ad.)) have the rejection region $B_{\widehat{\vartheta}(-1)}(T_{LR}) > \chi^2_{f,\alpha}$, $B_{\widehat{\vartheta}(0)}(T_{LH}) > \chi^2_{f,\alpha}$, $B_{\widehat{\vartheta}(-2)}(T_{BNP}) > \chi^2_{f,\alpha}$, $T_{LR} > CF_{\widehat{\vartheta}(-1)}(\chi^2_{f,\alpha})$, $T_{LH} > CF_{\widehat{\vartheta}(0)}(\chi^2_{f,\alpha})$ and $T_{BNP} > CF_{\widehat{\vartheta}(-2)}(\chi^2_{f,\alpha})$, respectively. Following Kakizawa [32] and Cribari-Neto and Ferrari [17], Bartlett's type monotone adjusted LR, LH, BNP tests (for shortly, B-monotone adjusted (B-mon.ad.)) or Cornish–Fisher's type

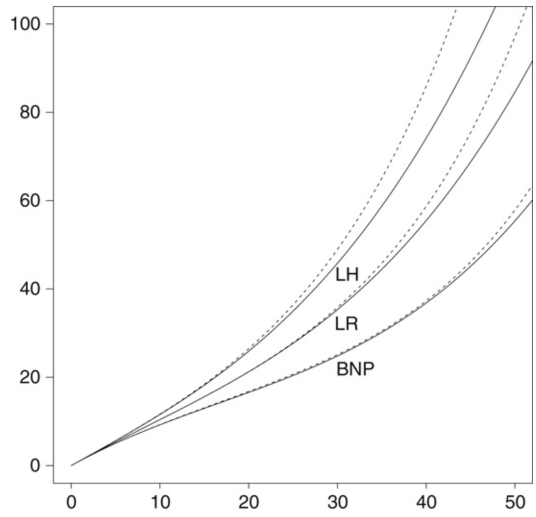

Fig. 1. Averaged estimated Cornish–Fisher's type transformations for $T = T_{LR}, T_{LH}, T_{BNP}$; $CF_{\widehat{\vartheta}(-1)}(x)$, $CF_{\widehat{\vartheta}(0)}(x)$, $CF_{\widehat{\vartheta}(-2)}(x)$ [solid] and $CF^{\mathrm{mon}}_{\widehat{\vartheta}(-1)}(x)$, $CF^{\mathrm{mon}}_{\widehat{\vartheta}(0)}(x)$, $CF^{\mathrm{mon}}_{\widehat{\vartheta}(-2)}(x)$ [dashed] (LN, $N = 30$, repetitions = 10,000).

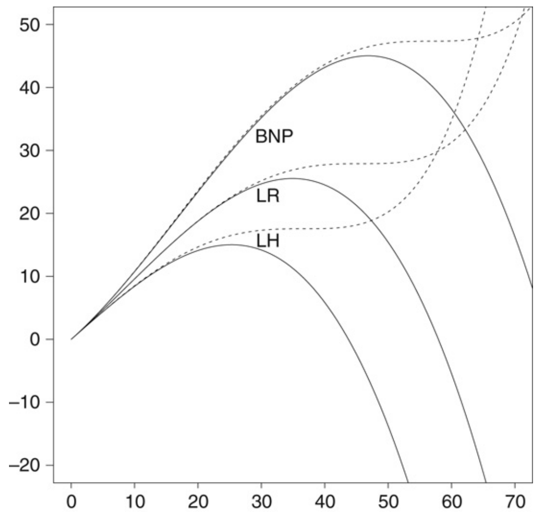

Fig. 2. Averaged estimated Bartlett's type transformations for $T = T_{LR}, T_{LH}, T_{BNP}$; $B_{\widehat{\vartheta}(-1)}(x)$, $B_{\widehat{\vartheta}(0)}(x)$, $B_{\widehat{\vartheta}(-2)}(x)$ [solid] and $B^{\mathrm{mon}}_{\widehat{\vartheta}(-1)}(x)$, $B^{\mathrm{mon}}_{\widehat{\vartheta}(0)}(x)$, $B^{\mathrm{mon}}_{\widehat{\vartheta}(-2)}(x)$ [dashed] (LN, $N = 30$, repetitions = 10,000).

monotone adjusted LR, LH, BNP tests (for shortly, CF-monotone adjusted (CF-mon.ad.)) are also examined, whose rejection region are given by $B^{\mathrm{mon}}_{\widehat{\vartheta}(-1)}(T_{LR}) > \chi^2_{f,\alpha}$, $B^{\mathrm{mon}}_{\widehat{\vartheta}(0)}(T_{LH}) > \chi^2_{f,\alpha}$, $B^{\mathrm{mon}}_{\widehat{\vartheta}(-2)}(T_{BNP}) > \chi^2_{f,\alpha}$, $T_{LR} > CF^{\mathrm{mon}}_{\widehat{\vartheta}(-1)}(\chi^2_{f,\alpha})$, $T_{LH} > CF^{\mathrm{mon}}_{\widehat{\vartheta}(0)}(\chi^2_{f,\alpha})$ and $T_{BNP} > CF^{\mathrm{mon}}_{\widehat{\vartheta}(-2)}(\chi^2_{f,\alpha})$, respectively, where

$$B^{\mathrm{mon}}_c(x) = x\left(1 - \frac{2}{N}\sum_{j=1}^{3} c_j x^{j-1} + \frac{1}{N^2}\sum_{j_1 j_2=1}^{3} \frac{j_1 j_2 c_{j_1} c_{j_2}}{j_1 + j_2 - 1}\, x^{j_1+j_2-2}\right),$$

$$CF^{\mathrm{mon}}_c(x) = x\left(1 + \frac{2}{N}\sum_{j=1}^{3} c_j x^{j-1} + \frac{1}{N^2}\sum_{j_1 j_2=1}^{3} \frac{j_1 j_2 c_{j_1} c_{j_2}}{j_1 + j_2 - 1}\, x^{j_1+j_2-2}\right).$$

One may suspect that such an additional N^{-2} term is negligible compared with the error $o(N^{-1})$ term in Theorem 1, but we feel that it corrects some bad performance of the right tail in the cubic polynomial $B_c(x) = x\{1 - (2/N)\sum_{j=1}^{3} c_j x^{j-1}\}$ (in our case, the cubic polynomial $CF_c(x) = x\{1 + (2/N)\sum_{j=1}^{3} c_j x^{j-1}\}$ evaluated at $x = \chi^2_{f,\alpha}$ (see Fig. 1) seems to have good performance, unlike [17]). To illustrate it, Fig. 2 shows the averaged (repetitions 10,000) estimated polynomials $B_{\widehat{\vartheta}(-1)}(x)$, $B_{\widehat{\vartheta}(0)}(x)$, $B_{\widehat{\vartheta}(-2)}(x)$, together with $B^{\mathrm{mon}}_{\widehat{\vartheta}(-1)}(x)$, $B^{\mathrm{mon}}_{\widehat{\vartheta}(0)}(x)$, $B^{\mathrm{mon}}_{\widehat{\vartheta}(-2)}(x)$, where the error distribution is (LN) with $N = 30$.

Table 1a
Empirical sizes ($\times 100$) of tests (CN)

N	Original			B-ad.			B-mon.ad.			CF-ad.			CF-mon.ad.		
	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP
120	5.31	6.63	4.12	5.01	4.94	4.93	5.01	4.98	4.94	5.01	5.10	5.00	5.01	5.08	4.98
90	5.40	7.20	3.83	4.99	4.88	4.85	4.99	4.94	4.87	5.00	5.16	4.96	4.99	5.13	4.94
60	5.66	8.60	3.26	5.00	4.71	4.70	5.01	4.85	4.74	5.02	5.40	4.96	5.02	5.31	4.90
30	6.66	13.96	1.76	5.13	2.65	3.93	5.15	4.06	4.05	5.22	6.75	4.88	5.20	6.41	4.63

Significance level ($\times 100$) was $100\alpha = 5$.

Table 1b
Empirical sizes ($\times 100$) of tests (LN)

N	Original			B-ad.			B-mon.ad.			CF-ad.			CF-mon.ad.		
	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP
120	5.14	6.40	4.01	5.12	5.04	5.05	5.13	5.07	5.08	5.12	5.20	5.13	5.12	5.18	5.10
90	5.22	6.95	3.73	5.11	4.96	4.98	5.11	5.01	5.02	5.11	5.25	5.12	5.10	5.22	5.08
60	5.51	8.34	3.24	5.18	4.77	4.90	5.19	4.93	4.96	5.19	5.52	5.19	5.19	5.44	5.10
30	6.63	13.51	1.85	5.37	0.82	4.18	5.40	3.94	4.32	5.47	6.90	5.18	5.45	6.59	4.89

Significance level ($\times 100$) was $100\alpha = 5$.

Table 2a
Empirical sizes and powers ($\times 100$) of tests (CN)

tr($\boldsymbol{\Omega}$)	B-ad.			B-mon.ad.			CF-ad.			CF-mon.ad.		
	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP
0	4.99	4.88	4.85	4.99	4.94	4.87	5.00	5.16	4.96	4.99	5.13	4.94
Type A												
5	28.74	28.59	28.07	28.75	28.76	28.13	28.76	29.46	28.45	28.76	29.35	28.37
10	56.54	56.58	55.53	56.55	56.78	55.60	56.57	57.56	55.96	56.56	57.43	55.87
15	77.28	77.41	76.35	77.28	77.56	76.40	77.29	78.15	76.67	77.29	78.05	76.61
20	89.32	89.45	88.67	89.33	89.54	88.71	89.33	89.89	88.89	89.33	89.83	88.85
Type B												
5	29.19	28.66	28.91	29.19	28.83	28.97	29.21	29.53	29.28	29.20	29.41	29.21
10	57.74	57.00	57.58	57.74	57.19	57.65	57.76	57.97	58.00	57.76	57.84	57.91
15	78.58	77.95	78.53	78.58	78.10	78.58	78.59	78.67	78.83	78.59	78.58	78.77
20	90.42	90.00	90.43	90.42	90.10	90.46	90.43	90.43	90.59	90.43	90.38	90.56

Significance level ($\times 100$) was $100\alpha = 5$ and sample size was $N = 90$.

The inequalities $T_{LH} \geq T_{LR} \geq T_{BNP}$ (e.g. [1, p337]) indicate that whenever $T_{BNP} \geq \chi^2_{8,0.05} = 15.507$ (we notice that the values of T_ψ are likely to be large as the alternatives are far away from the null), a B-adjusted LH test $B_{\widehat{\vartheta}(0)}(T_{LH})$ is likely to be not rejected at 5% level (actually, $B_{\widehat{\vartheta}(0)}(T_{LH})$ can be negative for the large value T_{LH}). Clearly, such an unsuccessful output can be got rid of by means of a monotone Bartlett's type adjusted test $B^{\text{mon}}_{\widehat{\vartheta}(0)}(T_{LH})$.

We first study the empirical sizes of three tests (original or B-adjusted or CF-adjusted LR, LH, BNP tests) at significance level $\alpha = 0.05$. Table 1 show that (i) the empirical sizes improve with increasing the sample size N and that (ii) the empirical sizes ($\times 100$) of B(or CF)-adjusted tests are closer to 100α than those of original tests, which obviously supports the higher-order improvements $\Pr[B_{\widehat{\vartheta}(\psi'')}(T_\psi) > \chi^2_{f,\alpha}|H] = \alpha + o(N^{-1})$ and $\Pr[T_\psi > CF_{\widehat{\vartheta}(\psi'')}(\chi^2_{f,\alpha})|H] = \alpha + o(N^{-1})$ to $\Pr[T_\psi > \chi^2_{f,\alpha}|H] = \alpha + O(N^{-1})$. However, CF-(monotone) adjusted LH tests tend to have a slight size distortion as compared with CF-(monotone) adjusted LR and BNP tests. On the other hand, for a small sample size as $N = 30$, B-adjusted LH and BNP tests, especially, B-adjusted LH test for the case (LN), have substantial size distortion, while the monotone adjustment $B^{\text{mon}}_{\widehat{\vartheta}(0)}(T_{LH})$ works reasonably to improve its sizes. In what follows, we set $N = 90$.

We next study the empirical power of three tests (B-adjusted or CF-adjusted LR, LH, BNP tests), in which their empirical sizes ($\times 100$), rounded to the nearest integer, are identical. Strictly speaking, since their empirical sizes are slightly different, we cannot, of course, make a definitive conclusion on the power comparison among three tests. But, we can see from Table 2 that under the alternative hypothesis of type A (B), the power of CF-(monotone) adjusted LR tests is greater (less) than that of BNP tests (a fair power comparison between LH and LR(or BNP) is impossible, since the size ($\times 100$); 5.1 or 5.2 of LH is slightly greater than the size ($\times 100$); 5.0 or 5.1 of LR(or BNP)). We observe that a B-adjusted LH test (also a B-adjusted LR test) suffers a substantial loss of power (especially when the noncentrality parameter tr($\boldsymbol{\Omega}$) is large and when the population distribution (LN) is considered), as compared with a B-adjusted BNP test, which is caused by a more frequent occurrence of low values (or negative values) of the transformed statistics $B_{\widehat{\vartheta}(0)}(T_{LH})$ and $B_{\widehat{\vartheta}(-1)}(T_{LR})$ than those of $B_{\widehat{\vartheta}(-2)}(T_{BNP})$. Such a bad performance, especially for (LN) case, is effectively resolved by considering monotone adjustments $B^{\text{mon}}_{\widehat{\vartheta}(-1)}(T_{LR})$, $B^{\text{mon}}_{\widehat{\vartheta}(0)}(T_{LH})$ and $B^{\text{mon}}_{\widehat{\vartheta}(-2)}(T_{BNP})$, as shown in Fig. 2 (the same comment is made for Table 3b). Under the alternative hypothesis of type A, the

Table 2b
Empirical sizes and powers ($\times 100$) of tests (LN)

tr($\boldsymbol{\Omega}$)	B-ad.			B-mon.ad.			CF-ad.			CF-mon.ad.		
	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP
0	5.11	4.96	4.98	5.11	5.01	5.02	5.11	5.25	5.12	5.10	5.22	5.08
Type A												
5	39.73	39.13	39.00	39.78	39.66	39.14	39.76	40.44	39.46	39.74	40.34	39.31
10	67.08	63.27	66.80	67.59	67.60	66.93	67.57	68.22	67.19	67.55	68.15	67.06
15	79.50	68.38	81.21	81.87	81.91	81.39	81.85	82.34	81.57	81.84	82.28	81.47
20	82.88	63.22	88.37	89.10	89.13	88.77	89.09	89.40	88.89	89.08	89.37	88.82
Type B												
5	40.60	39.63	40.37	40.64	40.05	40.50	40.63	40.85	40.85	40.60	40.75	40.68
10	71.19	68.02	71.36	71.54	70.99	71.49	71.52	71.65	71.76	71.50	71.57	71.62
15	85.03	75.65	86.75	86.98	86.62	86.96	86.96	87.01	87.12	86.95	86.97	87.03
20	87.85	69.40	93.21	93.88	93.69	93.88	93.88	93.89	93.97	93.87	93.87	93.92

Significance level ($\times 100$) was $100\alpha = 5$ and sample size was $N = 90$.

Table 3a
Empirical powers ($\times 100$) of tests using their empirical 5% critical values (CN)

tr($\boldsymbol{\Omega}$)	Original			B-ad.			B-mon.ad.		
	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP
Type A									
5	28.78	28.94	28.57	28.77	28.95	28.56	28.77	28.95	28.56
10	56.57	56.98	56.08	56.58	56.99	56.08	56.58	56.98	56.08
15	77.30	77.71	76.77	77.31	77.72	76.77	77.31	77.72	76.77
20	89.33	89.62	88.94	89.34	89.63	88.95	89.34	89.63	88.95
Type B									
5	29.22	29.03	29.40	29.22	29.03	29.40	29.22	29.03	29.40
10	57.77	57.39	58.11	57.78	57.41	58.12	57.78	57.41	58.11
15	78.60	78.25	78.92	78.61	78.26	78.92	78.61	78.26	78.92
20	90.43	90.18	90.64	90.44	90.19	90.64	90.44	90.19	90.64

Sample size was $N = 90$.

Table 3b
Empirical powers ($\times 100$) of tests using their empirical 5% critical values (LN)

tr($\boldsymbol{\Omega}$)	Original			B-ad.			B-mon.ad.		
	LR	LH	BNP	LR	LH	BNP	LR	LH	BNP
Type A									
5	39.35	39.63	39.02	39.37	39.27	39.05	39.39	39.64	39.06
10	67.18	67.51	66.76	66.78	63.39	66.84	67.27	67.58	66.86
15	81.56	81.82	81.23	79.30	68.48	81.24	81.66	81.89	81.34
20	88.87	89.05	88.63	82.73	63.30	88.39	88.97	89.12	88.74
Type B									
5	40.20	40.00	40.39	40.24	39.78	40.42	40.25	40.02	40.43
10	71.12	70.89	71.31	70.90	68.15	71.40	71.23	70.97	71.43
15	86.67	86.52	86.80	84.85	75.75	86.78	86.79	86.61	86.93
20	93.69	93.60	93.76	87.73	69.48	93.23	93.79	93.68	93.86

Sample size was $N = 90$.

ordering of the power among B-monotone adjusted LR, LH, BNP tests is $LH \succ LR \succ BNP$ when the noncentrality parameter tr($\boldsymbol{\Omega}$) is large, while, under the alternative hypothesis of type B, a B-monotone adjusted BNP test is more powerful than a B-monotone adjusted LH test. Another interesting feature is that unlike Table 2b (LN; asymmetric case), Bartlett's type adjusted LR, LH, BNP tests in Table 2a (CN; symmetric case) show good performance, as pointed out in Kakizawa and Iwashita [37].

We finally study the empirical power of three tests (original or B-adjusted LR, LH, BNP tests) that exceed their empirical upper 5% critical values (note that the size corrected original LR, LH, BNP tests correspond to CF-(monotone) adjusted LR, LH, BNP tests). This enables us to make a fair power comparison of different tests, although such critical values cannot be evaluated exactly, in applications where the information on the error disturbances is absent. As expected in the higher-order power analysis of Section 3.1, Table 3 shows that (iii) the power of a B-monotone adjusted test is almost identical to that of the original test and that (iv) the ordering of the power depends on the sign of the factor $D_{p,r}(\boldsymbol{\Omega})$, that is, since $D_{p,r}(\boldsymbol{\Omega}) > 0$ for the type A, the power of a B-monotone adjusted LH test is greater than that of a B-monotone adjusted LR test, which in turn is greater than that of a B-monotone adjusted BNP test (this ordering is reversed for the type B in which $D_{p,r}(\boldsymbol{\Omega}) < 0$ holds).

Our simulation experiments show that both correction methods (B-monotone adjusted and CF-(monotone) adjusted tests) work well in finite sample sizes. Remarkably, making use of a monotone Bartlett's type adjustment [32] in place of a Bartlett's type adjustment [16] is important to resolve not only the size distortion for the small sample size N but also the loss of power at alternatives far from the null. Actually, the power problem turns out to be serious for the asymmetric and leptokurtic distribution, which will be due to the difficulty of estimating summarized cumulants (4)–(6), as well as the nature of the cubic transformation (because of the inequalities $T_{LH} \geq T_{LR} \geq T_{BNP}$, Bartlett's type adjusted LH test under the alternatives far from the null is likely to be smallest value among three tests or have frequently negative value, which necessarily leads to the loss of the power).

7. Conclusion and future works

In this paper, we established that for normal-based multivariate tests, the local power properties stated in Anderson [1, p. 336] remain valid even for the nonnormal GMANOVA model, after either Bartlett's type adjustment or Cornish–Fisher's type size adjustment under nonnormality. The essential point behind this paper is that the statistical inference on the mean structure of the multivariate model is often asymptotically robust. More precisely, since several test statistics derived under normality (including the likelihood ratio (LR) criterion, Lawley–Hotelling's trace and Bartlett–Nanda–Pillai's trace) are all functions of the eigenvalues of a characteristic determinantal equation which involves the restricted and unrestricted residual sum of squares matrices, the (non)null distributions of these multivariate tests admit the (non)central chi-square type asymptotic expansion. It has also an implication of the correctness up to order $o(N^{-1})$ of the coverage probability of the bootstrap confidence region $\mathcal{R}_W^{\bullet} = \{\mathbf{D} : T_W^{\bullet}(\mathbf{D}) \leq z_W^{*}(\alpha)\}$ of $\mathbf{D} \equiv \mathbf{B}\boldsymbol{\Xi}\mathbf{C}$ based on the Wald test statistic $T_W^{\bullet}(\mathbf{D}) \equiv T_W^{\bullet}$, where $z_W^{*}(\alpha)$ is the residual-based nonparametric bootstrap critical value for $T_W^{\bullet}$ (details are omitted to preserve space). See Hall [29] for asymptotically normal statistics.

One of the most important problems for regression analysis, where the underlying distribution of the error is unknown, is how to construct a reasonable test or confidence region. Since its introduction as a nonparametric likelihood alternative to traditional likelihood-based methods for inference, Owen's empirical likelihood has gained increasing popularity among statisticians and econometricians. For the one-way MANOVA setting, Owen [52, subsection 4.4] mentioned a EL-based solution, in which the observations do not need to be normal distributed, or to have a common covariance matrix, and the sample size need not be equal (for the univariate case, the Euclidian likelihood ratio test described in [52, subsection 4.10] was essentially James's test [31]). Thus, it would be of some interest to investigate a higher-order power comparison between the EL approach and the classical normal-based approach discussed in Kakizawa and Iwashita [36,37] and Kakizawa [34]. For the one-sample case, Chen [13] investigated the EL-based test as a counterpart of Hotelling's T^2 test, which can be viewed as the Euclidian likelihood ratio test in the EL framework (e.g. [52, p65]), and then showed that the second-order power of the EL-based test is not always more powerful than that of the bootstrap T^2 test. Similar noncomparable conclusions on the second-order powers for a class of tests from the empirical discrepancy approach (the ED-based Cressie-Read tests) is found in Bravo [8], unless one considers the averaged power criterion (e.g. [48]). It would also be of interest to extend the results for the univariate linear regression model (e.g. [51,52,12]) to the GMANOVA model. As in the univariate case [52, Section 4], the multivariate linear regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{U}$ will be expected to have the empirical likelihood approach through estimating equation associated with the ordinary least squares (LS) estimator $\widehat{\boldsymbol{\Theta}}_{Y,\mathrm{LS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ (which is the maximum likelihood estimator under normality), whereas, for the GMANOVA model $\mathbf{Y} = \mathbf{X}\boldsymbol{\Xi}\mathbf{A} + \mathbf{U}$, attention must be paid to the choice of the arbitrary estimating function. As pointed out in Kariya [40, p26], there is a problem of choice between the OLSE and the GLSE, if $m < p$ (generally, the MLE $\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{ML}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\mathbf{E}_Y^{-1}\mathbf{A}'(\mathbf{A}\mathbf{E}_Y^{-1}\mathbf{A}')^{-1}$ under normality, which can be viewed as the generalized (or weighted) least squares estimator, that is, the minimizer of the criterion $(1/2)\mathrm{tr}[\widehat{\boldsymbol{\Sigma}}_Y^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\Xi}\mathbf{A})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\Xi}\mathbf{A})]$, is not equal to the LSE $\widehat{\boldsymbol{\Xi}}_{Y,\mathrm{LS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}$, so the least squares approach may be inefficient).

Appendix

A.1. Technical lemma

Recall (3) and (13). We need an asymptotic representation of $\dot{\mathbf{V}}$ by means of the so-called perturbation method (e.g. [59, Section 4.6]).

Lemma A.1. *For a given spectral decomposition of the $q \times q$ idempotent matrix $\mathbf{M}$;*

$$\mathbf{M} = \mathbf{V}\mathrm{diag}(\underbrace{1, \ldots, 1}_{r \text{ times}}, 0, \ldots, 0)\mathbf{V}' = \mathbf{V}^{(1:r)}(\mathbf{V}^{(1:r)})',$$

with $\mathbf{V}^{(1:r)} = [\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(r)}]$ being the first r columns of $\mathbf{V}$ ($q \times q$ orthogonal matrix), there exist $q \times q$ orthogonal matrix $\dot{\mathbf{V}}$ and $q \times q$ matrix $\mathbf{V}_1$ such that $\dot{\mathbf{V}} = \mathbf{V} + N^{-1}\mathbf{V}_1 + o(N^{-1})$ and (13).

Proof. With $\boldsymbol{\Lambda} = \text{diag}(\underbrace{1, \dots, 1}_{r \text{ times}}, 0, \dots, 0)$, we have from (7)

$$\mathbf{V}'\dot{\mathbf{M}}\mathbf{V} = \boldsymbol{\Lambda} + N^{-1}\mathbf{V}'\mathbf{M}_1\mathbf{V} + \text{o}(N^{-1})$$

and

$$\mathbf{V}'\mathbf{M}_1\mathbf{V} = \begin{pmatrix} \mathbf{O}_{r,r} & [\mathbf{V}'\widetilde{\mathbf{Q}}_1\mathbf{Q}^{1/2}\mathbf{V}]_{(12)} \\ [\mathbf{V}'\mathbf{Q}^{1/2}\widetilde{\mathbf{Q}}_1\mathbf{V}]_{(21)} & \mathbf{O}_{r-q,q-r} \end{pmatrix}, \tag{A.1}$$

where $[\mathbf{A}]_{(b_1b_2)}$ denotes the (b_1, b_2)th block of any matrix $\mathbf{A}$. Since the eigenvalues of $\mathbf{V}'\dot{\mathbf{M}}\mathbf{V}$ are one (with multiplicity r) and zero (with multiplicity $q-r$), there exists a $q \times q$ orthogonal matrix $\dot{\mathbf{W}}$, such that $\dot{\mathbf{W}} = \mathbf{I}_q + N^{-1}\mathbf{W}_1 + \text{o}(N^{-1})$ and $\mathbf{V}'\dot{\mathbf{M}}\mathbf{V} = \dot{\mathbf{W}}\boldsymbol{\Lambda}\dot{\mathbf{W}}'$. That is, by considering the N^{-1} terms of the equations $\dot{\mathbf{W}}\dot{\mathbf{W}}' = \mathbf{I}_q$ and $\mathbf{V}'\dot{\mathbf{M}}\mathbf{V} = \dot{\mathbf{W}}\boldsymbol{\Lambda}\dot{\mathbf{W}}'$, such a $q \times q$ matrix $\mathbf{W}_1$ must satisfy the relation $\mathbf{W}_1 = -\mathbf{W}_1'$ ($\mathbf{W}_1$ is skew-symmetric, hence all diagonal elements of $\mathbf{W}_1$ are zero) and $\mathbf{V}'\mathbf{M}_1\mathbf{V} = \mathbf{W}_1\boldsymbol{\Lambda} + \boldsymbol{\Lambda}\mathbf{W}_1'$. It follows from (A.1) that

$$\begin{pmatrix} \mathbf{O}_{r,r} & [\mathbf{V}'\widetilde{\mathbf{Q}}_1\mathbf{Q}^{1/2}\mathbf{V}]_{(12)} \\ [\mathbf{V}'\mathbf{Q}^{1/2}\widetilde{\mathbf{Q}}_1\mathbf{V}]_{(21)} & \mathbf{O}_{r-q,q-r} \end{pmatrix} = \begin{pmatrix} \mathbf{O}_{r,r} & -[\mathbf{W}_1]_{(12)} \\ [\mathbf{W}_1]_{(21)} & \mathbf{O}_{q-r,q-r} \end{pmatrix}$$

(we set $[\mathbf{W}_1]_{(11)} = \mathbf{O}_{r,r}$ and $[\mathbf{W}_1]_{(22)} = \mathbf{O}_{q-r,q-r}$ here, since the off-diagonal elements of $[\mathbf{W}_1]_{(11)}$ and $[\mathbf{W}_1]_{(22)}$ can be chosen arbitrarily). Thus, we can take $\mathbf{V}_1 = \mathbf{V}\mathbf{W}_1$. □

A.2. Additional lemmas for GMANOVA

Lemma A.2. *Let* $\boldsymbol{\delta}_{\mathcal{E}}^{\circ(1:r)} = [\boldsymbol{\delta}_{\mathcal{E}}^{\circ(1)}, \dots, \boldsymbol{\delta}_{\mathcal{E}}^{\circ(r)}] = (\boldsymbol{\Theta}_{\mathcal{E}}^{\circ})'\mathbf{V}^{(1:r)}$. *Let*

$$f(\boldsymbol{\gamma}^{(1:r)}) = \widetilde{\mathcal{P}}(\mathbf{Q}^{\Sigma}\boldsymbol{\gamma}^{(1:r)}) \exp\{\text{i}\widetilde{\mathcal{P}}_e(\mathbf{Q}^{\Sigma}\boldsymbol{\gamma}^{(1:r)})\}$$

and

$$\widetilde{f}(\boldsymbol{\gamma}^{(1:r)}) = \widetilde{\mathcal{P}}_{\perp}(\mathbf{P}^{\Sigma}\boldsymbol{\gamma}^{(1:r)})\widetilde{\mathcal{P}}(\mathbf{Q}^{\Sigma}\boldsymbol{\gamma}^{(1:r)}) \exp\{\text{i}\widetilde{\mathcal{P}}_e(\mathbf{Q}^{\Sigma}\boldsymbol{\gamma}^{(1:r)})\},$$

where $\widetilde{\mathcal{P}}_e(\boldsymbol{\gamma}^{(1:r)}) \in \mathbf{R}[\boldsymbol{\gamma}^{(1:r)}]$ *and* $\widetilde{\mathcal{P}}(\boldsymbol{\gamma}^{(1:r)}), \widetilde{\mathcal{P}}_{\perp}(\boldsymbol{\gamma}^{(1:r)}) \in \mathbf{C}[\boldsymbol{\gamma}^{(1:r)}]$. *Here,* $\mathbf{R}[\boldsymbol{\gamma}^{(1:r)}]$ ($\mathbf{C}[\boldsymbol{\gamma}^{(1:r)}]$) *is the set of polynomials of finite degree with coefficients in* $\mathbf{R}$ ($\mathbf{C}$). *Then,*

$$\Xi_0\widetilde{f}(\boldsymbol{\gamma}^{(1:r)} + \boldsymbol{\delta}_{\mathcal{E}}^{\circ(1:r)})|_{\gamma^{(1:r)}=0_{p,r}} = E_{Z_0}[\widetilde{\mathcal{P}}_{\perp}(\mathbf{z}_0^{(1:r)})]\, \Xi_0 f(\boldsymbol{\gamma}^{(1:r)} + \boldsymbol{\delta}_{\mathcal{E}}^{\circ(1:r)})|_{\gamma^{(1:r)}=0_{p,r}},$$

where $E_{Z_0}[\widetilde{\mathcal{P}}_{\perp}(\mathbf{z}_0^{(1:r)})]$ *denotes the expectation of* $\widetilde{\mathcal{P}}_{\perp}(\mathbf{z}_0^{(1:r)})$ *with respect to independent p-variate normal distributions* $\mathbf{z}_0^{(b)} \sim N_p(\mathbf{0}, \mathbf{P}^{\Sigma})$ ($b = 1, \dots, r$).

Proof. As in Kakizawa and Iwashita [36],

$$\Xi_0\widetilde{f}(\boldsymbol{\gamma}^{(1:r)} + \boldsymbol{\delta}_{\mathcal{E}}^{\circ(1:r)})|_{\gamma^{(1:r)}=0_{p,r}} = E_{Z^{\circ}}[\widetilde{\mathcal{P}}_{\perp}(\mathbf{P}^{\Sigma}\mathbf{z}^{\circ(1:r)})\widetilde{\mathcal{P}}(\mathbf{Q}^{\Sigma}\mathbf{z}^{\circ(1:r)}) \exp\{\text{i}\widetilde{\mathcal{P}}_e(\mathbf{Q}^{\Sigma}\mathbf{z}^{\circ(1:r)})\}],$$

where $E_{Z^{\circ}}[\ \ ]$ is the expectation with respect to independent p-variate normal distributions $\mathbf{z}^{\circ(b)} \sim N_p(\boldsymbol{\delta}_{\mathcal{E}}^{\circ(b)}, \boldsymbol{\Sigma})$ ($b = 1, \dots, r$). The assertion follows from the fact that

$$\text{vec}(\mathbf{P}^{\Sigma}\mathbf{z}^{\circ(1:r)}) = (\mathbf{I}_r \otimes \mathbf{P}^{\Sigma})\text{vec}(\mathbf{z}^{\circ(1:r)} - \boldsymbol{\delta}_{\mathcal{E}}^{\circ(1:r)}) \sim N_{pr}(\mathbf{0}, \mathbf{I}_r \otimes \mathbf{P}^{\Sigma})$$

is independent of $\text{vec}(\mathbf{Q}^{\Sigma}\mathbf{z}^{\circ(1:r)}) = (\mathbf{I}_r \otimes \mathbf{Q}^{\Sigma})\text{vec}(\mathbf{z}^{\circ(1:r)})$, since $\mathbf{Q}^{\Sigma}\boldsymbol{\Sigma}\mathbf{P}^{\Sigma} = \mathbf{O}_{p,p}$, $\mathbf{P}^{\Sigma}\boldsymbol{\Sigma}\mathbf{P}^{\Sigma} = \mathbf{P}^{\Sigma}$ and $\boldsymbol{\delta}_{\mathcal{E}}^{\circ(1:r)'}\mathbf{P}^{\Sigma} = \mathbf{O}_{r,p}$. □

We now define $\widetilde{\mathcal{H}}_{Q^{\Sigma}}(\boldsymbol{\gamma}^{(1:r)}) = \text{tr}[(\boldsymbol{\gamma}^{(1:r)})'\mathbf{Q}^{\Sigma}\boldsymbol{\gamma}^{(1:r)}] = \sum_{b=1}^{r}(\boldsymbol{\gamma}^{(b)})'\mathbf{Q}^{\Sigma}\boldsymbol{\gamma}^{(b)}$.

Lemma A.3. *For any* $\boldsymbol{\delta}^{\circ(1:r)} = [\boldsymbol{\delta}^{\circ(1)}, \dots, \boldsymbol{\delta}^{\circ(r)}] \in \mathcal{R}^{p\times r}$ *and* $t \in \mathbf{R}$, *one has*

$$\begin{aligned} &\exp\left(\frac{1}{2}\sum_{b=1}^{r}\boldsymbol{\partial}^{(b)'}\boldsymbol{\Sigma}\boldsymbol{\partial}^{(b)}\right)\exp\{\text{i}t\widetilde{\mathcal{H}}_{Q^{\Sigma}}(\boldsymbol{\gamma}^{(1:r)} + \boldsymbol{\delta}^{\circ(1:r)})\}|_{\gamma^{(1:r)}=0_{p,r}} \\ &\quad = (1-2\text{i}t)^{-rs/2}\exp\left[\frac{\text{i}t}{1-2\text{i}t}\,\text{tr}\{\mathbf{Q}^{\Sigma}\boldsymbol{\delta}^{\circ(1:r)}(\boldsymbol{\delta}^{\circ(1:r)})'\}\right] \end{aligned}$$

and

$$\begin{aligned} &\exp\left(\frac{1}{2}\sum_{b=1}^{r}\boldsymbol{\partial}^{(b)'}\boldsymbol{\Sigma}\boldsymbol{\partial}^{(b)}\right)\partial_{j_1}^{(b_1)}\cdots\partial_{j_v}^{(b_v)}\exp\{\text{i}t\widetilde{\mathcal{H}}_{Q^{\Sigma}}(\boldsymbol{\gamma}^{(1:r)} + \boldsymbol{\delta}^{\circ(1:r)})\}|_{\gamma^{(1:r)}=0_{p,r}} \\ &\quad = (1-2\text{i}t)^{-rs/2}\exp\left[\frac{\text{i}t}{1-2\text{i}t}\,\text{tr}\{\mathbf{Q}^{\Sigma}\boldsymbol{\delta}^{\circ(1:r)}(\boldsymbol{\delta}^{\circ(1:r)})'\}\right]\widetilde{Q}_{j_1\dots j_v}^{b_1\dots b_v}\left(\boldsymbol{\delta}^{\circ(1:r)}; \frac{\text{i}t}{1-2\text{i}t}\right) \end{aligned}$$

for $v \in \mathbf{N}$; $b_1, \ldots, b_v \in \{1, \ldots, r\}$, *where the general formula for* $\widetilde{Q}^{b_1\ldots b_v}_{j_1\ldots j_v}\{\delta^{\circ(1:r)}; \mathrm{i}t/(1-2\mathrm{i}t)\}$ *is given by*

$$\widetilde{Q}^{b_1\ldots b_v}_{j_1\ldots j_v}(\varphi) = \sum_{h=0}^{\nu} (\varphi - 1)^{v-h} \left\langle \frac{v!}{2^h h!(v-2h)!} \right\rangle_{2^h|1^{v-2h}} [Q^{\Sigma}]^{j_1 j_2}_{b_1 b_2} \cdots [Q^{\Sigma}]^{j_{2h-1} j_{2h}}_{b_{2h-1} b_{2h}} \widetilde{\delta}^{\circ(b_{2h+1})}_{j_{2h+1}} \cdots \widetilde{\delta}^{\circ(b_v)}_{j_v}$$

for $v = 2\nu (\neq 0)$ *or* $2\nu + 1$ ($\nu \in \mathbf{N}_0$; *the nonnegative integers* $\{0, 1, 2, \ldots, \}$) *with* $\varphi = (1 - 2\mathrm{i}t)^{-1}$ *and* $\mathbf{Q}^{\Sigma}\boldsymbol{\delta}^{\circ(b)} = (\widetilde{\delta}^{\circ(b)}_1, \ldots, \widetilde{\delta}^{\circ(b)}_p)'$ $(b = 1, \ldots, r)$. *Here,* $[Q^{\Sigma}]^{j_1 j_2}_{b_1 b_2} = [Q^{\Sigma}]_{j_1 j_2} \delta_{b_1 b_2}$ *is the* (j_1, j_2)*th element of the* (b_1, b_2)*th block of* $\mathbf{I}_r \otimes \mathbf{Q}^{\Sigma}$, *and* $\langle n \rangle_{2^h|1^{v-2h}}$ *before terms with indices is a sum of* n *similar terms obtained by the permutation of* $\left\{ \binom{b_1}{j_1}, \ldots, \binom{b_v}{j_v} \right\}$.

Proof. This lemma is an extension of Kakizawa and Iwashita [37, Lemma A1], who considered the case of $\mathbf{Q}^{\Sigma} = \boldsymbol{\Sigma}^{-1}$ (with $\mathbf{A} = \mathbf{C} = \mathbf{I}_p$). A proof of Lemma A.3 follows the arguments in proving Kakizawa and Iwashita [36, Proposition 2]. We omit the detail (the essential point here is that $\boldsymbol{\Sigma}^{1/2}\mathbf{Q}^{\Sigma}\boldsymbol{\Sigma}^{1/2}$ is idempotent with rank $s \leq p$). □

References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd ed., Wiley, New York, 2003.
[2] O.E. Barndorff-Nielsen, D.R. Cox, Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator, J. Roy. Statist. Soc. Ser. B 46 (1984) 483–495.
[3] E.R. Berndt, N.E. Savin, Conflict among criteria for testing hypotheses in the multivariate linear regression model, Econometrica 45 (1977) 1263–1277.
[4] R. Bhattacharya, M. Denker, Asymptotic Statistics, Birkhäuser-Verlag, Basel, 1990.
[5] R.N. Bhattacharya, J.K. Ghosh, On the validity of the formal Edgeworth expansion, Ann. Statist. 6 (1978) 434–451. Correction: 8 (1980) 1399.
[6] R.N. Bhattacharya, R.R. Rao, Normal Approximation and Asymptotic Expansions, Wiley, New York, 1976.
[7] P.J. Bickel, J.K. Ghosh, A decomposition for the likelihood ratio statistic and the Bartlett Correction — A Bayesian argument, Ann. Statist. 18 (1990) 1070–1090.
[8] F. Bravo, Second-order power comparisons for a class of nonparametric likelihood-based tests, Biometrika 90 (2003) 881–890.
[9] F. Bravo, Bartlett-type adjustments for empirical discrepancy test statistics, J. Stat. Plan. Inference 136 (2006) 537–554.
[10] T.K. Chandra, J.K. Ghosh, Valid asymptotic expansions for the likelihood ratio and other statistics under contiguous alternatives, Sankhyā 42 (1980) 170–184.
[11] T.K. Chandra, R. Mukerjee, Bartlett-type modification for Rao's efficient score statistic, J. Multivariate Anal. 36 (1991) 103–112.
[12] S.X. Chen, On the accuracy of empirical likelihood confidence regions for linear regression model, Ann. Inst. Statist. Math. 45 (1993) 621–637.
[13] S.X. Chen, Comparing empirical likelihood and bootstrap hypothesis tests, J. Multivariate Anal. 51 (1994) 277–293.
[14] S.X. Chen, H. Cui, On Bartlett correction of empirical likelihood in the presence of nuisance parameters, Biometrika 93 (2006) 215–220.
[15] D.M. Chibisov, An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion, Theory Probab. Appl. 17 (1972) 620–630.
[16] G.M. Cordeiro, S.L.P. Ferrari, A modified score test statistic having chi-squared distribution to order n^{-1}, Biometrika 78 (1991) 573–582.
[17] F. Cribari-Neto, S.L.P. Ferrari, Monotonic improved critical values for two χ^2 asymptotic criteria, Econom. Lett. 71 (2001) 307–316.
[18] T. DiCiccio, P. Hall, J. Romano, Empirical likelihood is Bartlett-correctable, Ann. Statist. 19 (1991) 1053–1061.
[19] J.-M. Dufour, L. Khalaf, Simulation based finite and large sample tests in multivariate regressions, J. Econometrics 111 (2002) 303–322.
[20] M.L. Eaton, M.D. Perlman, The non-singularity of generalized sample covariance matrices, Ann. Statist. 1 (1973) 710–717.
[21] Y. Fujikoshi, Asymptotic expansions of the distributions of test statistics in multivariate analysis, J. Sci. Hiroshima Univ. 34 (1970) 73–144.
[22] Y. Fujikoshi, Asymptotic expansions of the non-null distributions of three statistics in GMANOVA, Ann. Inst. Statist. Math. 26 (1974) 289–297.
[23] Y. Fujikoshi, Comparison of powers of a class of tests for multivariate linear hypothesis and independence, J. Multivariate Anal. 26 (1988) 48–58.
[24] Y. Fujikoshi, An asymptotic expansion for the distribution of Hotelling's T^2-statistic under nonnormality, J. Multivariate Anal. 61 (1997) 187–193.
[25] Y. Fujikoshi, Asymptotic expansions for the distributions of multivariate basic statistics and one-way MANOVA tests under nonnormality, J. Stat. Plan. Inference 108 (2002) 263–282.
[26] Y. Fujikoshi, Some recent results on asymptotic expansions of multivariate test statistics for mean vectors under nonnormality, Calcutta Statist. Assoc. Bulletin 52 (2002) 1–46.
[27] J.K. Ghosh, R. Mukerjee, Test statistics arising from quasi likelihood: Bartlett adjustment and higher-order power, J. Stat. Plan. Inference 97 (2001) 45–55.
[28] L.J. Gleser, I. Olkin, Linear models in multivariate analysis, in: R.C. Bose, I.M. Chakravarti, P.C. Mahalanobis, C.R. Rao, K.J.C. Simith (Eds.), Essays in Probability and Statistics, The University of North Carolina Press, North Carolina, 1970, pp. 267–292.
[29] P. Hall, The Bootstrap and Edgeworth Expansion, Springer, New York, 1992.
[30] D.A. Harville, Matrix Algebra From a Statistician's Perspective, Springer, New York, 1997.
[31] G.S. James, Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown, Biometrika 41 (1954) 19–43.
[32] Y. Kakizawa, Higher order monotone Bartlett-type adjustment for some multivariate test statistics, Biometrika 83 (1996) 923–927.
[33] Y. Kakizawa, Siotani's modified second approximation for multiple comparisons of mean vectors, SUT J. Math. 42 (2006) 59–96.
[34] Y. Kakizawa, A test of equality of mean vectors of several heteroscedastic multivariate populations, J. Japan Statist. Soc. 37 (2007) 253–283.
[35] Y. Kakizawa, Asymptotic expansions for the distributions of maximum and sum of quasi-independent Hotelling's T^2 statistics under nonnormality, Commun. Stat. Theory Methods 37 (2008) 97–120.
[36] Y. Kakizawa, T. Iwashita, 2008. Hotelling's one-sample and two-sample T^2 tests and the multivariate Behrens-Fisher problem under nonnormality. J. Stat. Plan. Inference, in press (doi:10.1016/j.jspi.2006.04.020).
[37] Y. Kakizawa, T. Iwashita, A comparison of higher-order local powers of a class of one-way MANOVA tests under general distributions, J. Multivariate Anal. 99 (2008) 1128–1153.
[38] T. Kanda, Growth curve model with covariance structures, Hiroshima Math. J. 24 (1994) 135–176.
[39] Y. Kano, An asymptotic expansion of the distribution of Hotelling's T^2-statistic under general distributions, Amer. J. Math. Management Sci. 15 (1995) 317–341.
[40] T. Kariya, Testing in the Multivariate General Linear Model, Kinokuniya, Tokyo, 1985.
[41] C.G. Khatri, A note on a MANOVA model applied to problems in growth curve, Ann. Inst. Statist. Math. 18 (1966) 75–86.
[42] D.G. Kleinbaum, A generalization of the growth curve model which allows missing data, J. Multivariate Anal. 3 (1973) 117–124.
[43] T. Kollo, D. von Rosen, Advanced Multivariate Statistics with Matrices, Springer, Netherlands, 2005.
[44] M.A. Magdalinos, Stochastic expansions and asymptotic approximations, Econom. Theory 8 (1992) 343–367.
[45] K.V. Mardia, Measures of multivariate skewness and kurtosis with applications, Biometrika 57 (1970) 519–530.
[46] P. McCullagh, Tensor Methods in Statistics, Chapman and Hall, London, 1987.
[47] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.

[48] R. Mukerjee, Rao's score test: Recent asymptotic results, in: G.S. Maddala, C.R. Rao, H.D. Vinod (Eds.), Handbook of Statistics, Vol.11, North-Holland, Amsterdam, 1993, pp. 363–379.
[49] R. Mukerjee, Comparison between the locally most mean power unbiased and Rao's tests in the multiparameter case, J. Multivariate Anal. 45 (1993) 9–24.
[50] R. Mukerjee, Comparison of tests in their original forms, Sankyhā A 56 (1994) 118–127.
[51] A.B. Owen, Empirical likelihood for linear models, Ann. Statist. 19 (1991) 1725–1747.
[52] A.B. Owen, Empirical Likelihood, Chapman and Hall, London, 2001.
[53] R.F. Potthoff, S.N. Roy, A generalized multivariate analysis of variance model useful especially for growth curve problems, Biometrika 51 (1964) 313–326.
[54] M.B. Qumsiyeh, Edgeworth expansion in regression models, J. Multivariate Anal. 35 (1990) 86–101.
[55] C.R. Rao, R. Mukerjee, Comparison of Bartlett-type adjustments for the efficient score statistic, J. Stat. Plan. Inference 46 (1995) 137–146.
[56] C.R. Rao, R. Mukerjee, Comparison of LR, score, and Wald tests in a non-iid setting, J. Multivariate Anal. 60 (1997) 99–110.
[57] T.J. Rothenberg, Edgeworth expansions for multivariate test statistics, IP–255, Center for Research in Management Science, University of California, Berkeley, 1977.
[58] M. Siotani, Distributions of Lawley–Hotelling's T_0^2 and related statistics: A review, Amer. J. Math. Management Sci. 9 (1989) 5–30.
[59] M. Siotani, T. Hayakawa, Y. Fujikoshi, Modern Multivariate Analysis: A Graduate Course and Handbook, American Sciences Press, Ohio, 1985.
[60] M.S. Srivastava, D. von Rosen, Growth curve models, in: S. Ghosh (Ed.), Multivariate Analysis, Design of Experiments, and Survey Sampling, Marcel Dekker, New York, 1999, pp. 547–578.
[61] M. Taniguchi, Third-order asymptotic properties of a class of test statistics under a local alternative, J. Multivariate Anal. 37 (1991) 223–238.
[62] H. Wakaki, H. Yanagihara, Y. Fujikoshi, Asymptotic expansions of the null distributions of test statistics for multivariate linear hypothesis under nonnormality, Hiroshima Math. J. 32 (2002) 17–50.
[63] B.L. Welch, The generalization of 'Student's' problem when several different population variances are involved, Biometrika 34 (1947) 28–35.
[64] H. Yanagihara, Asymptotic expansion of the null distribution of three test statistics in a nonnormal GMANOVA model, Hiroshima Math. J. 31 (2001) 213–262.
[65] H. Yanagihara, A family of estimators for multivariate kurtosis in a nonnormal linear regression model, J. Multivariate Anal. 98 (2007) 1–29.
[66] H. Yanagihara, Conditions for robustness to nonnormality on test statistics in a GMANOVA model, J. Japan Statist. Soc. 37 (2007) 135–155.