TEST
https://doi.org/10.1007/s11749-024-00922-0

ORIGINAL PAPER

Testing covariance structures belonging to a quadratic subspace under a doubly multivariate model

Katarzyna Filipiak[1] · **Mateusz John**[1] · **Yuli Liang**[2]

Received: 7 June 2023 / Accepted: 22 January 2024

Abstract
A hypothesis related to the block structure of a covariance matrix under the doubly multivariate normal model is studied. It is assumed that the block structure of the covariance matrix belongs to a quadratic subspace, and under the null hypothesis, each block of the covariance matrix also has a structure belonging to some quadratic subspace. The Rao score and the likelihood ratio test statistics are derived, and the exact distribution of the likelihood ratio test is determined. Simulation studies show the advantage of the Rao score test over the likelihood ratio test in terms of speed of convergence to the limiting chi-square distribution, while both proposed tests are competitive in terms of their power. The results are applied to both simulated and real-life example data.

Keywords Doubly multivariate model · Covariance structure · Quadratic subspace · Rao score test · Likelihood ratio test

Mathematics Subject Classification 62H15 · 62E20 · 62E15

1 Introduction

It is very common nowadays to collect multi-level multivariate data, which are hierarchical by nature. In particular, for each subject, there may be several variables measured

✉ Yuli Liang
yuli.liang@lnu.se

Katarzyna Filipiak
katarzyna.filipiak@put.poznan.pl

Mateusz John
mateusz.john@put.poznan.pl

[1] Institute of Mathematics, Poznań University of Technology, Piotrowo 3A, 60-965 Poznań, Poland

[2] Department of Economics and Statistics, Linnaeus University, Georg Lückligs väg 4, 352 52 Växjö, Sweden

Published online: 28 February 2024

at various occasions (sites, locations, time points, etc.), resulting in so-called doubly multivariate data, that is, multivariate on two levels. Thus, the variables of interest can be naturally subdivided into certain groups, which results in a covariance matrix having some block structure, also called a nested covariance structure. The particular dependence pattern for variables from different groups (interclass dependence) leads to the block structure of the covariance matrix, and the specific dependence pattern for variables within each group (intraclass dependence) results in certain pattered sub-blocks. For example, assuming equal dependence between any given pair of variables from the same group results in a compound symmetry (CS) structure for each block of the covariance matrix (cf. Wilks, 1946), while assuming equal dependence between any given pair of variables from different groups results in a block compound symmetry (BCS) covariance structure (for algebraic definitions of the structures see Sect. 2). Clearly, in this case CS characterizes the intraclass dependence pattern, while BCS describes the interclass dependence structure. In the literature (B)CS structures are also called (block) exchangeable structures. Study of the BCS covariance structure goes back to Arnold (1973, 1979), in connection with a general linear model where the error vectors are assumed to be normally distributed.

Another typical dependence pattern is circular dependence (cf. Olkin and Press 1969, Eaton 1983, Nahtman and von Rosen 2008), common in many applications, especially in medical (Louden and Roy 2010) and educational studies (Steinmetz et al. 2012). A circular dependence structure, known also as a circular Toeplitz (CT) structure, means that the correlation between two measurements depends only on the circular distance between them, in other words, on the number of positions between them on a circle. Circular dependence in blocks leading to a block circular Toeplitz (BCT) structure was considered for example by Olkin (1973).

It should be emphasized that under doubly multivariate models, the number of parameters in the covariance matrix increases quadratically with the dimensions (number of variables and occasions), which may cause problems with estimation. Imposing some nested structure reduces the number of unknown parameters significantly, especially if this structure is block diagonal, BCS or BCT, or, more generally, belongs to a quadratic subspace; cf. Seely (1971).

Before any statistical analysis is conducted, it is necessary to check the validity of the covariance structure using a relevant testing procedure. There has been much interest over the years in statistical testing problems in block covariance matrices, due to demand for more efficient modeling and hence for extracting the inherent dependence present in the data. For example, the Rao score test (RST) for a first-order autoregression structure was derived by Chi and Reinsel (1989). Computationally intensive procedures for testing covariance structures have also been developed, such as parametric bootstrap tests and permutation tests. Hypothesis testing for a parallel profile model with a CS random-effects covariance structure was considered by Srivastava and Singull (2012), who observed that only the distinct eigenvalues of the CS covariance matrix are required to be estimated, rather than the original (co)variance parameters. In Filipiak et al. (2016, 2017), the RST of a separable covariance structure in a doubly multivariate model was studied, and it was demonstrated that the RST is much less biased than the corresponding likelihood ratio test (LRT) and is suitable for small sample sizes. Similar results were obtained by Roy et al. (2018) for testing the

BCS structure. Tests for the mean structure under a model with the BCS covariance matrix were proposed in Zmyślony et al. (2018). Estimation of the BCS covariance structure with a CS or CT intraclass pattern was studied in Liang et al. (2015), while recently, the LRT for testing simultaneously the mean and circular structure of blocks of the BCS covariance matrix was derived by Liang et al. (2022). The RST and Wald test statistics, as well as the exact distribution of the LRT for testing independence of features between any two repeated measures under the BCS structure, were studied by Filipiak et al. (2023a).

It should be noted that all commonly used structures, such as diagonality (D), CS, CT, or their interclass (block) versions, belong to certain quadratic subspaces. Moreover, the intraclass dependence structure is often commutative, in contrast to the situation when the only requirement on the subblocks of the covariance matrix is their symmetry (note that two arbitrary symmetric matrices do not commute). Therefore, in this article, we propose tests of a general hypothesis related to patterned covariance matrices belonging to some (commutative) quadratic subspaces, formulated in (4). In particular, our considerations cover, among others, tests for a block diagonal (BD) with equal diagonal blocks having a CT structure versus BD with equal unstructured symmetric diagonal blocks, BCT with CS blocks versus BCT with unstructured symmetric blocks, or BCT with CS blocks versus BCT with CT blocks. Three particular hypotheses (see (16)) will later be used as illustrative examples of the results obtained, however, the methods presented can also be applied to covariance structures that are not yet named in the literature, but are common in real-life experiments.

The main goals of this paper are to derive the RST and LRT statistics for testing (4), to establish the exact distribution of the LRT, to compare the speed of convergence of the exact distribution of the LRT and the empirical distribution of the RST to the limiting chi-square distribution, and to compare the powers of the RST and LRT. The latter two aims will be verified via Monte Carlo simulation studies. The results will be illustrated with applications to both simulated data and real-life data, taken from Liang et al. (2015).

Finally, we note that in this paper we assume both of the dimensions (numbers of variables and occasions) to be fixed. Under the standard (one-dimensional) multivariate model some structures have also been considered under a high-dimensional setup; for example, the distribution of tests related to the likelihood ratio and/or Rao score test has been studied in Bai et al. (2009); Jiang et al. (2012) (for testing the identity matrix), John (1971); Jiang (2016) (for testing sphericity), and Kato et al. (2010); Yamada (2012); Yi and Xie (2018); Xie and Sun (2019); Tang et al. (2022); Klein et al. (2022) (for testing compound symmetry). Very recently (Lai et al. 2023) proposed a test for testing block-diagonality, based on U-statistics. Note however, that generalization of the known results to the common structures for doubly multivariate models is relatively challenging, and hence will be the topic of future research.

The structure of the paper is as follows. Section 2 introduces the statistical model of interest and the hypothesis. The main results are stated in Sect. 3, where the RST and LRT statistics are derived, and the null distribution of the LRT is determined. The results of simulation studies related to the asymptotic behavior of the distributions of the RST and LRT, as well as a comparison of their powers, are given in Sect. 4. The theory presented is illustrated with both simulated data and a real-life example in

Sect. 5. Finally, in Sect. 6, some conclusions are stated and further perspectives are discussed.

2 Model and hypothesis

Let us consider the following doubly multivariate model, in which p features are measured repeatedly q times on n independent subjects:

$$\boldsymbol{X} \sim N_{n,qp}(\mathbf{1}_n\boldsymbol{\mu}', \boldsymbol{I}_n, \boldsymbol{\Omega}), \tag{1}$$

such that $\boldsymbol{\mu}$ is a qp-dimensional general mean (the same for every subject), $\mathbf{1}_n$ is the n-dimensional vector of ones, $\boldsymbol{I}_n$ is the identity matrix of order n, and the $qp \times qp$ symmetric positive definite (p.d.) matrix $\boldsymbol{\Omega}$ belongs to a quadratic subspace $\mathcal{A}$. Recall that a subspace $\mathcal{A}$ of the linear space of real symmetric matrices is a quadratic subspace if $\boldsymbol{\Omega} \in \mathcal{A}$ implies $\boldsymbol{\Omega}^2 \in \mathcal{A}$; cf. Seely (1971).

Assume that $p = 1$. Then, using spectral decomposition, the covariance matrix $\boldsymbol{\Omega} = \boldsymbol{\Omega}_{p=1}$ can be represented as

$$\boldsymbol{\Omega}_{p=1} = \sum_{i=1}^{\nu} \lambda_i \boldsymbol{V}_i, \tag{2}$$

where λ_i are the eigenvalues of $\boldsymbol{\Omega}_{p=1}$ and $\boldsymbol{V}_i = \boldsymbol{H}_i\boldsymbol{H}_i'$, with $\boldsymbol{H}_i$ consisting of the set of eigenvectors corresponding to λ_i. Following Seely (1971, Lemma 6), the quadratic subspace is commutative if and only if there exist matrices $\boldsymbol{V}_i$ that are idempotent and pairwise orthogonal, that is, $\boldsymbol{V}_i^2 = \boldsymbol{V}_i$ and $\boldsymbol{V}_i\boldsymbol{V}_j = \mathbf{0}$ for each $i \neq j$, $i, j = 1 \ldots, \nu$. Throughout the paper we assume that $\boldsymbol{V}_i$, $i = 1, \ldots, \nu$, satisfy the aforementioned conditions, and we denote by $\mathcal{V}$ the commutative quadratic subspace generated by $\boldsymbol{V}_i$. Clearly, $\dim\mathcal{V} = \nu$.

Note that the most common examples of structures belonging to a commutative quadratic subspace $\mathcal{V}$ are: matrices proportional to the identity, $\boldsymbol{\Omega}_{\mathrm{I}} = \lambda\boldsymbol{I}_q$; diagonal (D), $\boldsymbol{\Omega}_{\mathrm{D}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_q)$; compound symmetric, $\boldsymbol{\Omega}_{\mathrm{CS}} = \sigma^2\left[(1-\rho)\boldsymbol{I}_q + \rho\mathbf{1}_q\mathbf{1}_q'\right]$, where $\sigma^2 \in \mathbb{R}^+$ and $\rho \in (-1/(q-1), 1)$, cf. Olkin and Press (1969); and circular Toeplitz,

$$\boldsymbol{\Omega}_{\mathrm{CT}} = \begin{cases} \sigma^2\left[\boldsymbol{I}_q + \sum_{i=1}^{\lfloor q/2 \rfloor} \rho_i(\boldsymbol{C}^i + \boldsymbol{C}'^i)\right] & \text{for odd } q \\ \sigma^2\left[\boldsymbol{I}_q + \sum_{i=1}^{q/2-1} \rho_i(\boldsymbol{C}^i + \boldsymbol{C}'^i) + \rho_{q/2}\boldsymbol{C}^{q/2}\right] & \text{for even } q, \end{cases}$$

where $\sigma^2 \in \mathbb{R}^+$, all ϱ_i are real and ensure the positive definiteness of $\boldsymbol{\Omega}_{\mathrm{CT}}$, and $\boldsymbol{C}$ is a permutation matrix with all zero elements except the first subdiagonal and the upper right corner, where the elements are equal to 1; cf. Basilevsky (1983); Olkin and Press (1969). It can be seen that a CS matrix has all diagonal entries equal and all off-diagonal entries equal, while a CT matrix is symmetric and each subsequent row is simply a circular shift of the previous one. Note that to obtain the representation (2), both $\mathbf{\Omega}_{\mathrm{CS}}$ and $\mathbf{\Omega}_{\mathrm{CT}}$ should be expressed in terms of an idempotent basis of orthogonal matrices; cf. Appendix C for the spectral decomposition of the structures in question.

For $p > 1$ each scalar entry of $\mathbf{\Omega}_{p=1}$ can be replaced by a $p \times p$ symmetric matrix representing intraclass dependence, that is,

$$\mathbf{\Omega} = \sum_{i=1}^{\nu} \boldsymbol{V}_i \otimes \mathbf{\Delta}_i =: \mathbf{\Omega}_1, \tag{3}$$

where $\mathbf{\Delta}_i$, $i = 1, \dots, \nu$, are symmetric p.d. covariance matrices of order p. Observe that $\mathbf{\Omega}_1$ still belongs to a quadratic subspace; however, commutativity no longer holds (as symmetric matrices do not commute). Denoting by $\mathcal{W}$ the space of all symmetric p.d. matrices of order p (which is in fact a quadratic subspace), we will write that $\mathbf{\Omega}_1 \in \mathcal{V} \boxtimes \mathcal{W}$. Note that in general $\mathcal{V} \boxtimes \mathcal{W}$ is a quadratic but not commutative subspace; however, if we assume a particular structure on $\mathbf{\Delta}_i$, $i = 1, \dots, \nu$, that implies commutativity, then the subspace $\mathcal{V} \boxtimes \mathcal{W}$ becomes commutative too. Observe moreover, that if $\mathcal{V}$ represents one of the structures given in previous paragraph, then $\mathbf{\Omega}_1$ indicates respectively a block structure with the same matrices on the diagonal (BI), block diagonal (BD), BCS, and BCT structures.

In this paper we are interested in testing the hypothesis that the intraclass dependence is also patterned, and it belongs to some commutative quadratic subspace $\mathcal{U} \subset \mathcal{W}$ of dimension ω. We may therefore formulate the following hypothesis:

$$H_0: \ \mathbf{\Omega} \in \mathcal{V} \boxtimes \mathcal{U} \qquad \text{vs} \qquad H_1: \ \mathbf{\Omega} \in \mathcal{V} \boxtimes (\mathcal{W} \backslash \mathcal{U}), \tag{4}$$

where $\mathcal{V}$ and $\mathcal{U}$ are commutative quadratic subspaces of dimension ν and ω, respectively, while $\mathcal{W}$ is an arbitrary subspace of a $p(p+1)/2$-dimensional space of symmetric matrices of order p, with $\mathcal{U} \subset \mathcal{W}$. The symbol $\mathcal{W} \backslash \mathcal{U}$ denotes the subspace $\mathcal{W}$ with the subspace $\mathcal{U}$ excluded, and is used to stress the mutual exclusiveness of the hypotheses.

Let $\{\boldsymbol{U}_1, \dots, \boldsymbol{U}_\omega\}$ be an orthogonal basis of the commutative quadratic subspace $\mathcal{U}$ of matrices of order p. Then the covariance matrix under the null hypothesis can be expressed as

$$\mathbf{\Omega} = \sum_{i=1}^{\nu} \boldsymbol{V}_i \otimes \sum_{j=1}^{\omega} \delta_{ij} \boldsymbol{U}_j = \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \delta_{ij} \left(\boldsymbol{V}_i \otimes \boldsymbol{U}_j \right) =: \mathbf{\Omega}_0, \tag{5}$$

where δ_{ij} are the eigenvalues associated with the eigenvectors forming $\mathcal{V} \boxtimes \mathcal{U}$. Note that $\mathbf{\Omega}_0$ covers many combinations of known covariance structures, for example:

- BD matrices with D, CS or CT blocks (denoted respectively as $\mathbf{\Omega}_{\mathrm{BD\_D}}$, $\mathbf{\Omega}_{\mathrm{BD\_CS}}$, and $\mathbf{\Omega}_{\mathrm{BD\_CT}}$);
- BCS matrices with the same diagonal blocks and the same off-diagonal blocks, having D, CS or CT structure (denoted respectively as $\mathbf{\Omega}_{\mathrm{BCS\_D}}$, $\mathbf{\Omega}_{\mathrm{BCS\_CS}}$, and $\mathbf{\Omega}_{\mathrm{BCS\_CT}}$);

- BCT matrices with the same diagonal blocks, the same blocks on the ith diagonal, $i = 1, \ldots, \lfloor q/2 \rfloor$, having D, CS or CT structure (denoted respectively as $\boldsymbol{\Omega}_{\mathrm{BCT\_D}}$, $\boldsymbol{\Omega}_{\mathrm{BCT\_CS}}$, and $\boldsymbol{\Omega}_{\mathrm{BCT\_CT}}$);
- Kronecker products of any combination of D, CS or CT structures, e.g. CT $\otimes$ CS, CT $\otimes$ CT.

We would also like to note that some of the structures considered here actually reflect the dependence structure of the same data, which arises, however, from different labeling of the two levels. Algebraically, such a change is equivalent to multiplication of the considered structures by relevant commutation matrices, as for each $i = 1, \ldots, \nu$, $j = 1, \ldots, \omega$,

$$\boldsymbol{K}_{p,q}(\boldsymbol{V}_i \otimes \boldsymbol{U}_j)\boldsymbol{K}_{q,p} = \boldsymbol{U}_j \otimes \boldsymbol{V}_i; \tag{6}$$

cf. Magnus and Neudecker (1986, formula (24)). Thus, one can observe that the above multiplication transforms, for example, $\boldsymbol{\Omega}_{\mathrm{BCT\_CS}}$ into $\boldsymbol{\Omega}_{\mathrm{BCS\_CT}}$ because $\boldsymbol{\Omega}_{\mathrm{BCS\_CT}} = \boldsymbol{K}_{p,q}\boldsymbol{\Omega}_{\mathrm{BCT\_CS}}\boldsymbol{K}_{q,p}$ for certain p and q.

We also emphasize that in both, the null and alternative hypotheses of (4), the space of the interclass structure is spanned by the same set of matrices. In view of the previous comment, this implies that one of the quadratic subspaces (representing the interclass or intraclass covariance structure) must be the same in both, null and alternative hypotheses. For example, one can test a hypothesis such as $\boldsymbol{\Omega}_{\mathrm{BCS\_CS}}$ vs $\boldsymbol{\Omega}_{\mathrm{BCS\_CT}}$ (the intraclass structure is the same) or $\boldsymbol{\Omega}_{\mathrm{BCS\_CT}}$ vs $\boldsymbol{\Omega}_{\mathrm{BCT\_CT}}$ (the interclass structure is the same), although the hypothesis $\boldsymbol{\Omega}_{\mathrm{BCS\_CS}}$ vs $\boldsymbol{\Omega}_{\mathrm{BCT\_CT}}$ (with $p, q \geq 4$) is outside the scope of this paper. For the same reason, hypothesis (4) does not cover the case when, for example, we examine the independence of measurements at various occasions versus a BCS structure (cf. Filipiak et al. 2023a), as this would be equivalent to testing $\boldsymbol{\Omega}_{\mathrm{BI}}$ vs. $\boldsymbol{\Omega}_{\mathrm{BCS}}$. Moreover, it is easy to see that such a structure is not hierarchical, as it is focused only on the interclass covariance structure.

Finally, note that the separable covariance structures of the form $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$, where $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are arbitrary unstructured symmetric positive definite matrices of order q and p, are not considered in this paper, as long as $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ remain unstructured. This follows from the fact that the tensor space of Kronecker products of two unstructured matrices is not a linear space (is bilinear) and the space of unstructured matrices is not a commutative quadratic subspace. Nevertheless, if one or both of the components of the Kronecker product were to have a structure that belongs to a commutative quadratic subspace, the obtained covariance matrix could then be tested via the methods presented here; for example $\boldsymbol{\Omega}_{\mathrm{CS}} \otimes \boldsymbol{\Omega}_{\mathrm{UN}} = \boldsymbol{\Omega}_{\mathrm{BCS}}$, $\boldsymbol{\Omega}_{\mathrm{CS}} \otimes \boldsymbol{\Omega}_{\mathrm{CS}} = \boldsymbol{\Omega}_{\mathrm{BCS\_CS}}$, where the subscript UN denotes “unstructured".

3 Test statistics

In this section the main results of the paper are presented. First, we determine the form of the RST, and in the second part we derive the likelihood ratio statistic and its exact distribution. Since for both cases MLEs are needed, the considerations related to the test are preceded by establishing the form of the MLE of a structured covariance matrix.

3.1 Maximum likelihood estimators

Let us denote by $\boldsymbol{P}_A = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$ the orthogonal projector onto the column space of $\boldsymbol{A}$, and let $\boldsymbol{Q}_A = \boldsymbol{I} - \boldsymbol{P}_A$. Observing that under model (1) the expectation can be presented as $\mathbb{E}(\operatorname{vec}\boldsymbol{X}) = (\boldsymbol{I}_{qp} \otimes \mathbf{1}_n)\boldsymbol{\mu}$ and the covariance matrix $\mathbb{D}(\operatorname{vec} X) = \boldsymbol{\Omega} \otimes \boldsymbol{I}_n := \boldsymbol{\Omega}_{\mathrm{UN}} \otimes \boldsymbol{I}_n$, where $\boldsymbol{\Omega}_{\mathrm{UN}}$ denotes an unstructured matrix, it can be seen that the space of expectation commutes with the space of covariance matrices. Indeed,

$$
\begin{aligned}
\boldsymbol{P}_{I_{qp}\otimes 1_n}(\boldsymbol{\Omega}_{\mathrm{UN}} \otimes \boldsymbol{I}_n) &= (I_{qp} \otimes \boldsymbol{P}_{1_n})(\boldsymbol{\Omega}_{\mathrm{UN}} \otimes \boldsymbol{I}_n) \\
&= \qquad \boldsymbol{\Omega}_{\mathrm{UN}} \otimes \boldsymbol{P}_{1_n} \\
&= (\boldsymbol{\Omega}_{\mathrm{UN}} \otimes \boldsymbol{I}_n)(I_{qp} \otimes \boldsymbol{P}_{1_n}) = (\boldsymbol{\Omega}_{\mathrm{UN}} \otimes \boldsymbol{I}_n)\boldsymbol{P}_{I_{qp}\otimes 1_n}.
\end{aligned}
\tag{7}
$$

It is known that if (7) holds, then the likelihood equation for expectation does not depend on the covariance matrix and thus its solution is an ordinary least squares estimator. Therefore, the MLE of $\boldsymbol{\mu}$ is simply the sample mean, $\widehat{\boldsymbol{\mu}} = \frac{1}{n}\boldsymbol{X}'\mathbf{1}_n$, and the MLE of $\boldsymbol{\Omega}_{\mathrm{UN}}$ has the following explicit form:

$$
\widehat{\boldsymbol{\Omega}}_{\mathrm{UN}} = \tfrac{1}{n}\boldsymbol{X}'\boldsymbol{Q}_{1_n}\boldsymbol{X} \;\; =: \tfrac{1}{n}\boldsymbol{S}, \tag{8}
$$

where $\boldsymbol{S} \sim \mathcal{W}_{qp}(\boldsymbol{\Omega}_{\mathrm{UN}}, n-1)$; cf. Anderson (2003); Filipiak et al. (2020). Since the commutativity of the space of expectation and covariance matrix, as presented in (7), holds regardless of the possible structure of $\boldsymbol{\Omega}$, the MLE of $\boldsymbol{\mu}$ is always equal to the sample mean. However, imposing the structure on the covariance matrix requires determination of its MLE.

Following Filipiak et al. (2020), since both $\mathcal{V} \boxtimes \mathcal{U}$ and $\mathcal{V} \boxtimes (\mathcal{W}\backslash\mathcal{U})$ are quadratic subspaces, the MLEs of $\boldsymbol{\Omega}_0$ and $\boldsymbol{\Omega}_1$ can be represented as the projections of $\widehat{\boldsymbol{\Omega}}_{\mathrm{UN}}$ onto the space of corresponding structures. Thus, we may formulate the following proposition, with a detailed proof in Appendix A. To present the MLE of a covariance structure belonging to a quadratic subspace, we use the block trace operator, denoted by $\operatorname{BTr}_p \boldsymbol{A}$, defined as the sum of all diagonal $p \times p$ blocks of the $qp \times qp$ partitioned matrix $\boldsymbol{A}$; cf. Filipiak et al. (2018). Moreover, for $i = 1, \ldots, \nu$ and $j = 1, \ldots, \omega$ we use the notation $v_i = \operatorname{tr}(\boldsymbol{V}_i)$ and $u_j = \operatorname{tr}(\boldsymbol{U}_j)$, where $\boldsymbol{V}_i$ and $\boldsymbol{U}_j$ are the basis matrices of $\mathcal{V}$ and $\mathcal{U}$, respectively.

Proposition 1 *Under model (1) the MLE of the covariance matrix for the null and alternative hypotheses in (4) is given, respectively, by*

$$
\begin{aligned}
\widehat{\boldsymbol{\Omega}}_0 &= \sum_{i=1}^{\nu}\sum_{j=1}^{\omega}\widehat{\delta}_{ij}\left(\boldsymbol{V}_i \otimes \boldsymbol{U}_j\right) \quad \textit{with} \quad \widehat{\delta}_{ij} = \tfrac{1}{n\cdot v_i\cdot u_j}\operatorname{tr}\left[\left(\boldsymbol{V}_i \otimes \boldsymbol{U}_j\right)\boldsymbol{S}\right], \\
\widehat{\boldsymbol{\Omega}}_1 &= \sum_{i=1}^{\nu}\boldsymbol{V}_i \otimes \widehat{\boldsymbol{\Delta}}_i \qquad\qquad\quad \textit{with} \quad \widehat{\boldsymbol{\Delta}}_i = \tfrac{1}{n\cdot v_i}\operatorname{BTr}_p\left[(\boldsymbol{V}_i \otimes \boldsymbol{I}_p)\boldsymbol{S}\right],
\end{aligned}
$$

where $\boldsymbol{S}$ is defined in (8).

Note that an alternative version of the proof might be obtained by starting with a transformation of the original model (1) by $(\boldsymbol{H} \otimes \boldsymbol{I}_p)$, where $\boldsymbol{H} = (\boldsymbol{H}_1, \ldots, \boldsymbol{H}_\nu)$ consists of matrices forming a basis of $\mathcal{V}$, that is,

$$\boldsymbol{Y} = (\boldsymbol{X} - \mathbf{1}_n \boldsymbol{\mu}')(\boldsymbol{H} \otimes \boldsymbol{I}_p) \sim N_{n,qp}\left(\mathbf{0}_{n\times qp}, \boldsymbol{I}_n, (\boldsymbol{H}' \otimes \boldsymbol{I}_p)\boldsymbol{\Omega}(\boldsymbol{H} \otimes \boldsymbol{I}_p)\right),$$

which is identical to ν independent MANOVA (multivariate analysis of variance) models with zero mean.

3.2 Rao score test

Following Rao (2005), the Rao score test statistic is defined as

$$\mathrm{RS} = \boldsymbol{s}'(\widehat{\boldsymbol{\theta}}_0)\boldsymbol{F}^{-1}(\widehat{\boldsymbol{\theta}}_0)\boldsymbol{s}(\widehat{\boldsymbol{\theta}}_0),$$

where the score vector $\boldsymbol{s}(\boldsymbol{\theta})$ consists of the first derivatives of the log-likelihood function with respect to the vector of parameters under the alternative hypothesis, $\boldsymbol{F}(\boldsymbol{\theta})$ is the Fisher information matrix, and $\widehat{\boldsymbol{\theta}}_0$ is the MLE of the vector of parameters under the null hypothesis. For a symmetric matrix $\boldsymbol{A}$ of order p, we define the vec-half operator as the $p(p+1)/2 \times 1$ vector that is obtained from vec $\boldsymbol{A}$ by eliminating all upper triangle elements of $\boldsymbol{A}$. In other words, if $\boldsymbol{D}_p$ is a $p^2 \times p(p+1)/2$ duplication matrix (cf. Magnus and Neudecker, 1986), then $\boldsymbol{D}_p$ vech $\boldsymbol{A}$ = vec $\boldsymbol{A}$, and the vector of parameters can be written under the alternative hypothesis as $\boldsymbol{\theta} = \left(\boldsymbol{\mu}', \mathrm{vech}'\,\boldsymbol{\Delta}_1, \ldots, \mathrm{vech}'\,\boldsymbol{\Delta}_\nu\right)'$, while under the null hypothesis as $\boldsymbol{\theta}_0 = \left(\boldsymbol{\mu}', \delta_{11}, \delta_{12} \ldots, \delta_{1\omega}, \ldots, \delta_{\nu 2}, \ldots, \delta_{\nu\omega}\right)'$. Observe that since hypothesis (4) does not contain any restrictions on $\boldsymbol{\mu}$, the first entry of the score vector (first derivative with respect to $\boldsymbol{\mu}$) will reduce to $\mathbf{0}$ when $\boldsymbol{\mu}$ is replaced by its MLE. Moreover, for the same reason, the off-diagonal entries of the first row and column of the Fisher information matrix also reduce to 0. Thus, throughout this section we consider $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_0$ without the first component, as well as a Fisher information matrix that is of order $\nu\omega$ instead of $\nu\omega + qp$. Note also that all matrix derivatives in this paper are computed according to the rules given in Magnus and Neudecker (1988).

We give the following theorem presenting the form of the RST statistic.

Theorem 1 *Under the hypothesis (4) the Rao score test statistic has the form*

$$\mathrm{RS} = \frac{n}{2}\,\mathrm{tr}\left\{\left[\boldsymbol{I}_{qp} - \widehat{\boldsymbol{\Omega}}_1\widehat{\boldsymbol{\Omega}}_0^{-1}\right]^2\right\},$$

where $\widehat{\boldsymbol{\Omega}}_0$ and $\widehat{\boldsymbol{\Omega}}_1$, the MLEs of $\boldsymbol{\Omega}_0$ and $\boldsymbol{\Omega}_1$, are as given in Proposition 1.

Proof The log-likelihood function under H_1 can be presented as

$$\ln L(\boldsymbol{\mu}, \boldsymbol{\Omega}_1) = -\frac{nqp}{2}\ln(2\pi) - \frac{n}{2}\sum_{i=1}^{\nu}(v_i \ln|\boldsymbol{\Delta}_i|) - \frac{1}{2}\sum_{i=1}^{\nu}\mathrm{tr}\left[\boldsymbol{Z}'\boldsymbol{Z}(\boldsymbol{V}_i \otimes \boldsymbol{\Delta}_i^{-1})\right],$$

with $\boldsymbol{Z} = \boldsymbol{Z}(\boldsymbol{\mu}) = \boldsymbol{X} - \mathbf{1}_n\boldsymbol{\mu}' \sim N(\mathbf{0}, \boldsymbol{I}_n, \boldsymbol{\Omega}_1)$. Differentiation of the above with respect to $\boldsymbol{\Delta}_i$, $i = 1, \ldots, \nu$, with the use of Magnus and Neudecker (1986), Magnus and Neudecker (1988), and Corollary 2.10 of Filipiak et al. (2018), gives

$$\begin{aligned}\frac{\partial \ln L}{\partial \boldsymbol{\Delta}_i} &= \Big[-\frac{n v_i}{2} \operatorname{vec}' \boldsymbol{\Delta}_i^{-1} \\ &\quad + \tfrac{1}{2} \operatorname{vec}'(\boldsymbol{Z}'\boldsymbol{Z})(\boldsymbol{I}_q \otimes \boldsymbol{K}_{p,q} \otimes \boldsymbol{I}_p)(\operatorname{vec} \boldsymbol{V}_i \otimes \boldsymbol{I}_{p^2})(\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1})\Big]\boldsymbol{D}_p \\ &= \tfrac{n}{2}\left(-v_i \operatorname{vec}' \boldsymbol{\Delta}_i^{-1} + \operatorname{vec}'\left\{\mathrm{BTr}_p[(\boldsymbol{V}_i \otimes \boldsymbol{I}_p)\tfrac{1}{n}\boldsymbol{Z}'\boldsymbol{Z}]\right\}(\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1})\right)\boldsymbol{D}_p,\end{aligned}$$

where $\boldsymbol{K}_{p,q}$ is a commutation matrix of order pq such that for a $p \times q$ matrix $\boldsymbol{A}$ we have $\boldsymbol{K}_{p,q} \operatorname{vec} \boldsymbol{A} = \operatorname{vec} \boldsymbol{A}'$; cf. Kollo and von Rosen (2005). Putting $\widehat{\boldsymbol{\theta}}_0$ as given by Proposition 1 and $\widehat{\boldsymbol{\mu}}$ as given in (8) into the above derivative, and using the notation

$$\widehat{\boldsymbol{\Gamma}}_i = \sum_{j=1}^{\omega} \widehat{\delta}_{ij}\boldsymbol{U}_j, \tag{9}$$

we obtain

$$\boldsymbol{s}(\widehat{\boldsymbol{\theta}}_0) = -\frac{n}{2}(\boldsymbol{I}_\nu \otimes \boldsymbol{D}_p')(\boldsymbol{s}_1', \ldots, \boldsymbol{s}_\nu')',$$

where $\boldsymbol{s}_i = v_i \operatorname{vec} \widehat{\boldsymbol{\Gamma}}_i^{-1} - \frac{1}{n}\left(\widehat{\boldsymbol{\Gamma}}_i^{-1} \otimes \widehat{\boldsymbol{\Gamma}}_i^{-1}\right) \operatorname{vec}\{\mathrm{BTr}_p[(\boldsymbol{V}_i \otimes \boldsymbol{I}_p)\boldsymbol{S}]\}$, $i = 1, \ldots, \nu$, and $\boldsymbol{S}$ is given in (8). Since the block trace of the matrix given in this formula is proportional to the MLE of $\boldsymbol{\Delta}_i$ (see Proposition 1), we may express each $\boldsymbol{s}_i$ as

$$\boldsymbol{s}_i = v_i \left[\left(\widehat{\boldsymbol{\Gamma}}_i^{-1} \otimes \widehat{\boldsymbol{\Gamma}}_i^{-1}\right) \operatorname{vec}\left(\widehat{\boldsymbol{\Gamma}}_i - \widehat{\boldsymbol{\Delta}}_i\right)\right].$$

To compute the Fisher information matrix, $\boldsymbol{F}(\boldsymbol{\theta}) = (\boldsymbol{F}_{ik}(\boldsymbol{\theta}))$, $i, k = 1, \ldots, \nu$, it is necessary to determine second-order partial derivatives and to calculate their expected values. Using formula (1.4.23) from Kollo and von Rosen (2005), for each $i = 1, \ldots, \nu$, we have

$$\begin{aligned}\frac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_i^2} &= \boldsymbol{D}_p'\Big[\tfrac{n v_i}{2}(\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1}) \\ &\quad -\tfrac{1}{2}(\operatorname{vec}'\{\mathrm{BTr}_p[(\boldsymbol{V}_i \otimes \boldsymbol{I}_p)\boldsymbol{Z}'\boldsymbol{Z}]\} \otimes \boldsymbol{I}_{p^2})(\boldsymbol{I}_p \otimes \boldsymbol{K}_{p,p} \otimes \boldsymbol{I}_p) \\ &\qquad \cdot(\boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \operatorname{vec} \boldsymbol{\Delta}_i^{-1} + \operatorname{vec} \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1} \otimes \boldsymbol{\Delta}_i^{-1})\Big]\boldsymbol{D}_p,\end{aligned}$$

and $\dfrac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_i \partial \boldsymbol{\Delta}_k} = \mathbf{0}$ for $i \neq k$ and $i, k = 1, 2, \ldots, \nu$.

Following Kollo and von Rosen (2005, Theorem 2.2.9(i)), under the alternative hypothesis, $\mathrm{E}(\boldsymbol{Z}'\boldsymbol{Z}) = \operatorname{tr} \boldsymbol{I}_n \cdot \boldsymbol{\Omega}_1$. Hence, using orthogonality and idempotency of $\boldsymbol{V}_i$ for $i = 1, 2, \ldots, \nu$, we obtain

$$
\begin{aligned}
\boldsymbol{F}_{ii}(\boldsymbol{\theta}) &= -\mathrm{E}\left(\frac{\partial^2 \ln L}{\partial \boldsymbol{\Delta}_i^2}\right) \\
&= -\tfrac{n}{2}\boldsymbol{D}_p'\Big[v_i(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1}) \\
&\qquad -\{\mathrm{vec}'[\mathrm{BTr}_p(\boldsymbol{V}_i\otimes\boldsymbol{\Delta}_i)]\otimes\boldsymbol{I}_{p^2}\}(\boldsymbol{I}_p\otimes\boldsymbol{K}_{p,p}\otimes\boldsymbol{I}_p) \\
&\qquad \cdot(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1}\otimes\mathrm{vec}\,\boldsymbol{\Delta}_i^{-1}+\mathrm{vec}\,\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\Big]\boldsymbol{D}_p.
\end{aligned}
$$

From Filipiak et al. (Filipiak et al. (2018), Lemma 2.11, Lemma 2.13) it can be shown that $\mathrm{BTr}_p(\boldsymbol{V}_i\otimes\boldsymbol{\Delta}_i)=v_i\boldsymbol{\Delta}_i$. Moreover, from Filipiak et al. (Filipiak et al. (2016), Lemma 1) it is known that for a symmetric matrix $\boldsymbol{A}$

$$
\begin{aligned}
&(\mathrm{vec}'\,\boldsymbol{A}\otimes\boldsymbol{D}_p')(\boldsymbol{I}_p\otimes\boldsymbol{K}_{p,p}\otimes\boldsymbol{I}_p)(\mathrm{vec}\,\boldsymbol{A}^{-1}\otimes\boldsymbol{I}_{p^2}) \\
&\quad=(\mathrm{vec}'\,\boldsymbol{A}\otimes\boldsymbol{D}_p')(\boldsymbol{I}_p\otimes\boldsymbol{K}_{p,p}\otimes\boldsymbol{I}_p)(\boldsymbol{I}_{p^2}\otimes\mathrm{vec}\,\boldsymbol{A}^{-1})=\boldsymbol{D}_p'.
\end{aligned}
$$

Hence, for every $i=1,2,...,\nu$, we obtain

$$
\begin{aligned}
\boldsymbol{F}_{ii}(\boldsymbol{\theta}) &= -\tfrac{nv_i}{2}\Big[\boldsymbol{D}_p'(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\boldsymbol{D}_p \\
&\quad -(\mathrm{vec}'\,\boldsymbol{\Delta}_i\otimes\boldsymbol{D}_p')(\boldsymbol{I}_p\otimes\boldsymbol{K}_{p,p}\otimes\boldsymbol{I}_p)(\boldsymbol{I}_{p^2}\otimes\mathrm{vec}'\,\boldsymbol{\Delta}_i^{-1})(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\boldsymbol{D}_p \\
&\quad -(\mathrm{vec}'\,\boldsymbol{\Delta}_i\otimes\boldsymbol{D}_p')(\boldsymbol{I}_p\otimes\boldsymbol{K}_{p,p}\otimes\boldsymbol{I}_p)(\mathrm{vec}'\,\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{I}_{p^2})(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\boldsymbol{D}_p\Big] \\
&= -\tfrac{nv_i}{2}\Big[\boldsymbol{D}_p'(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\boldsymbol{D}_p-2\boldsymbol{D}_p'(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\boldsymbol{D}_p\Big] \\
&= \tfrac{nv_i}{2}\boldsymbol{D}_p'(\boldsymbol{\Delta}_i^{-1}\otimes\boldsymbol{\Delta}_i^{-1})\boldsymbol{D}_p.
\end{aligned}
$$

Putting $\widehat{\boldsymbol{\theta}}_0$ as given by Proposition 1 into the above formula, we finally obtain

$$
\boldsymbol{F}_{ii}(\widehat{\boldsymbol{\theta}}_0)=\frac{nv_i}{2}\boldsymbol{D}_p'(\widehat{\boldsymbol{\Gamma}}_i^{-1}\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1})\boldsymbol{D}_p,\quad i=1,2,...,\nu,
$$

with $\widehat{\boldsymbol{\Gamma}}_i$ as defined in (9).

From Filipiak et al. (Filipiak et al. (2016), Proposition 1 (iv)) it is known, that for any nonsingular matrix $\boldsymbol{A}$ of order m

$$
\left[\boldsymbol{D}_p'(\boldsymbol{A}^{-1}\otimes\boldsymbol{A}^{-1})\boldsymbol{D}_p\right]^{-1}=\boldsymbol{D}_p^{+}(\boldsymbol{A}\otimes\boldsymbol{A})\boldsymbol{D}_p^{+'},
$$

where $\boldsymbol{D}_p^{+}$ is the Moore-Penrose inverse of $\boldsymbol{D}_p$. Therefore $\boldsymbol{F}_{ii}^{-1}(\widehat{\boldsymbol{\theta}}_0)=\boldsymbol{D}_p^{+}(\widehat{\boldsymbol{\Gamma}}_i\otimes\widehat{\boldsymbol{\Gamma}}_i)\boldsymbol{D}_p^{+'}$, $i=1,2,...,v$, and finally

$$
\boldsymbol{F}^{-1}(\widehat{\boldsymbol{\theta}}_0)=\frac{2}{n}(\boldsymbol{I}_\nu\otimes\boldsymbol{D}_p^{+})\,\mathrm{BDiag}\left(\widehat{\boldsymbol{\Gamma}}_i\otimes\widehat{\boldsymbol{\Gamma}}_i\right)(\boldsymbol{I}_\nu\otimes\boldsymbol{D}_p^{+'}).
$$

Taking the product of the score vector and Fisher information matrix we obtain

$$\begin{aligned}\mathrm{RS} &= \tfrac{n}{2}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_\nu)(\boldsymbol{I}_\nu\otimes\boldsymbol{D}_p\boldsymbol{D}_p^+)\,\mathrm{BDiag}\left[\tfrac{1}{v_i}\left(\widehat{\boldsymbol{\Gamma}}_i\otimes\widehat{\boldsymbol{\Gamma}}_i\right)\right](\boldsymbol{I}_\nu\otimes(\boldsymbol{D}_p\boldsymbol{D}_p^+)')(\boldsymbol{s}_1',\ldots,\boldsymbol{s}_\nu')'\\ &= \tfrac{n}{2}\sum_{i=1}^{\nu}\Big[v_i\,\mathrm{vec}'\left(\widehat{\boldsymbol{\Gamma}}_i-\widehat{\boldsymbol{\Delta}}_i\right)\left(\widehat{\boldsymbol{\Gamma}}_i^{-1}\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\boldsymbol{D}_p\boldsymbol{D}_p^+\left(\widehat{\boldsymbol{\Gamma}}_i\otimes\widehat{\boldsymbol{\Gamma}}_i\right)(\boldsymbol{D}_p\boldsymbol{D}_p^+)'\\ &\qquad\cdot\left(\widehat{\boldsymbol{\Gamma}}_i^{-1}\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\mathrm{vec}\left(\widehat{\boldsymbol{\Gamma}}_i-\widehat{\boldsymbol{\Delta}}_i\right)\Big].\end{aligned}$$

Since $\boldsymbol{D}_p\boldsymbol{D}_p^+ = \boldsymbol{N}_p$, where $\boldsymbol{N}_p = \frac{1}{2}(\boldsymbol{I}_{p^2}+\boldsymbol{K}_{p,p})$ (cf. Magnus and Neudecker, 1986, formulas (54) and (33)) and because for any symmetric matrix $\boldsymbol{A}$ the formula $\boldsymbol{N}_p\,\mathrm{vec}\,\boldsymbol{A} = \mathrm{vec}\,\boldsymbol{A}$ holds (cf. Magnus and Neudecker, 1986, formula (36)), from the symmetry of each $\widehat{\boldsymbol{\Gamma}}_i$ we obtain

$$\begin{aligned}\mathrm{RS} &= \tfrac{n}{2}\sum_{i=1}^{\nu} v_i\,\mathrm{vec}'(\widehat{\boldsymbol{\Gamma}}_i-\widehat{\boldsymbol{\Delta}}_i)(\widehat{\boldsymbol{\Gamma}}_i^{-1}\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1})\,\mathrm{vec}(\widehat{\boldsymbol{\Gamma}}_i-\widehat{\boldsymbol{\Delta}}_i)\\ &= \tfrac{n}{2}\sum_{i=1}^{\nu} v_i\,\mathrm{tr}\left[(\widehat{\boldsymbol{\Gamma}}_i-\widehat{\boldsymbol{\Delta}}_i)\widehat{\boldsymbol{\Gamma}}_i^{-1}(\widehat{\boldsymbol{\Gamma}}_i-\widehat{\boldsymbol{\Delta}}_i)\widehat{\boldsymbol{\Gamma}}_i^{-1}\right]\\ &= \tfrac{n}{2}\sum_{i=1}^{\nu} v_i\,\mathrm{tr}\left(\boldsymbol{I}_p-2\widehat{\boldsymbol{\Delta}}_i\widehat{\boldsymbol{\Gamma}}_i^{-1}+\widehat{\boldsymbol{\Delta}}_i\widehat{\boldsymbol{\Gamma}}_i^{-1}\widehat{\boldsymbol{\Delta}}_i\widehat{\boldsymbol{\Gamma}}_i^{-1}\right).\end{aligned}$$

From the idempotency of $\boldsymbol{V}_i$, $i = 1, 2, \ldots, \nu$, and since $v_i\,\mathrm{tr}\,\boldsymbol{I}_p = \mathrm{tr}\left(\boldsymbol{V}_i\otimes\boldsymbol{I}_p\right)$ we may rewrite the above as

$$\begin{aligned}\mathrm{RS} = \tfrac{n}{2}\,\mathrm{tr}\Big[\boldsymbol{I}_{qp} &- 2\sum_{i=1}^{\nu}\left(\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Delta}}_i\right)\left(\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\\ &-\sum_{i=1}^{\nu}\left(\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Delta}}_i\right)\left(\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\left(\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Delta}}_i\right)\left(\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\Big],\end{aligned}$$

and finally, again using the orthogonality of $\boldsymbol{V}_i$, $i = 1, \ldots, \nu$,

$$\begin{aligned}\mathrm{RS} &= \frac{n}{2}\,\mathrm{tr}\Bigg[\boldsymbol{I}_{qp} - 2\left(\sum_{i=1}^{\nu}\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Delta}}_i\right)\left(\sum_{i=1}^{\nu}\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\\ &\qquad+\left(\sum_{i=1}^{\nu}\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Delta}}_i\right)\left(\sum_{i=1}^{\nu}\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\left(\sum_{i=1}^{\nu}\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Delta}}_i\right)\left(\sum_{i=1}^{\nu}\boldsymbol{V}_i\otimes\widehat{\boldsymbol{\Gamma}}_i^{-1}\right)\Bigg]\\ &= \frac{n}{2}\,\mathrm{tr}\left\{\left[\boldsymbol{I}_{qp}-\widehat{\boldsymbol{\Omega}}_1\widehat{\boldsymbol{\Omega}}_0^{-1}\right]^2\right\}.\end{aligned}$$

□

3.3 Likelihood ratio test

The MLE of $\boldsymbol{\Delta}_i$ given in Proposition 1 is represented in terms of $\mathrm{BTr}_p\left[(\boldsymbol{V}_i\otimes\boldsymbol{I}_p)\boldsymbol{S}\right]$, which can be equivalently written as $\mathrm{BTr}_p\left[(\boldsymbol{H}_i'\otimes\boldsymbol{I}_p)\boldsymbol{S}(\boldsymbol{H}_i\otimes\boldsymbol{I}_p)\right]$ with $\boldsymbol{V}_i = \boldsymbol{H}_i\boldsymbol{H}_i'$; cf. Filipiak et al. (2018). Thus, one can see that the block trace operator merely sums v_i blocks of order p of a $v_i p\times v_i p$ partitioned matrix. Such an operation is equivalent to $\sum_{j=1}^{v_i}(\boldsymbol{h}_{ij}'\otimes\boldsymbol{I}_p)\boldsymbol{S}(\boldsymbol{h}_{ij}\otimes\boldsymbol{I}_p)$, with $\boldsymbol{h}_{ij}$ being the jth column of $\boldsymbol{H}_i$, or $\sum_{j=1}^{v_i} \boldsymbol{Y}_{ij}' \boldsymbol{Q}_{1_n} \boldsymbol{Y}_{ij}$, with $\boldsymbol{Y}_{ij} = \boldsymbol{X}(\boldsymbol{h}_{ij} \otimes \boldsymbol{I}_p)$, denoted by Liang et al. (2022) as $\boldsymbol{S}_i$. The representation of the MLEs of the covariance structure under the null and alternative hypotheses in terms of $\boldsymbol{S}_i$ will be convenient for determining the likelihood ratio statistic. Therefore, let

$$\boldsymbol{S}_i = \mathrm{BTr}_p\left[(\boldsymbol{H}_i' \otimes \boldsymbol{I}_p)\boldsymbol{S}(\boldsymbol{H}_i \otimes \boldsymbol{I}_p)\right]. \tag{10}$$

Using the fact that tr $\boldsymbol{A} = \mathrm{tr}(\mathrm{BTr}\,\boldsymbol{A})$ and Filipiak et al. (Filipiak et al. (2018), Lemma 2.11), the MLEs of δ_{ij} and $\boldsymbol{\Delta}_i$ in Proposition 1, $i = 1, \ldots, \nu$, $j = 1, \ldots, \omega$, can be represented as

$$\widehat{\delta}_{ij} = \frac{1}{n \cdot v_i \cdot u_j} \mathrm{tr}\left(\boldsymbol{U}_j \boldsymbol{S}_i\right) \quad \text{and} \quad \widehat{\boldsymbol{\Delta}}_i = \frac{1}{n \cdot v_i} \boldsymbol{S}_i. \tag{11}$$

Now we can formulate the likelihood ratio criterion.

Theorem 2 *The likelihood ratio criterion* Λ *for testing (4) is given by*

$$\Lambda = \left\{ \prod_{i=1}^{\nu} \left[|\boldsymbol{S}_i|^{v_i} \prod_{j=1}^{\omega} \left(\frac{u_j}{t_{ij}} \right)^{v_i u_j} \right] \right\}^{n/2},$$

where $t_{ij} = \mathrm{tr}(\boldsymbol{S}_i \boldsymbol{U}_j)$ *and* $\boldsymbol{S}_i$ *is as defined in (10).*

Proof The likelihood ratio is given by

$$\Lambda = \frac{\max_{H_0} L(\boldsymbol{\mu}, \boldsymbol{\Omega}_0)}{\max_{H_1} L(\boldsymbol{\mu}, \boldsymbol{\Omega}_1)},$$

with the maximum attained at the MLEs of $\boldsymbol{\Omega}_0$ and $\boldsymbol{\Omega}_1$ given in Proposition 1. Using the representation (11), we can write

$$\max_{H_0} L(\boldsymbol{\mu}, \boldsymbol{\Omega}_0) = L(\widehat{\boldsymbol{\mu}}, \widehat{\delta}_{ij}) = (2\pi)^{-\frac{qpn}{2}} \prod_{i=1}^{\nu} \prod_{j=1}^{\omega} \left(\frac{\mathrm{tr}(\boldsymbol{S}_i \boldsymbol{U}_j)}{n v_i u_j} \right)^{-\frac{n v_i u_j}{2}} e^{-\frac{qpn}{2}}$$

and

$$\max_{H_1} L(\boldsymbol{\mu}, \boldsymbol{\Omega}_1) = L(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Delta}}_i) = (2\pi)^{-\frac{qpn}{2}} \prod_{i=1}^{\nu} \left(\frac{|\boldsymbol{S}_i|}{(n v_i)^p} \right)^{-\frac{n v_i}{2}} e^{-\frac{qpn}{2}},$$

and the expression follows. □

It is known that for increasing sample size the likelihood ratio test statistic, LRT $= -2 \ln \Lambda$, tends to the chi-square distribution with $(p(p+1)/2 - \omega)\nu$ degrees of freedom; cf. Rao (2005). Nevertheless, it is also worthwhile to derive the exact distribution of the LRT. For this purpose we first present auxiliary lemmas which will be used in the subsequent proof. The notation "$\overset{d}{\sim}$" is used to indicate that a random variable "has the same distribution as" another, and $\beta(a, b)$ denotes the beta distribution with parameters a and b.

Lemma 1 (Muirhead, 1982, Theorem 3.2.15, p. 100) *If $\boldsymbol{A} \sim \mathcal{W}_m(\boldsymbol{\Sigma}, n)$, where $n \geq m$ is an integer, then $|\boldsymbol{A}| / |\boldsymbol{\Sigma}| \overset{d}{\sim} \prod_{i=1}^{m} X_i$, where $X_i \sim \chi^2_{n-i+1}$, $i = 1, \ldots, m$, are independent random variables.*

Lemma 2 (Olkin and Press, 1969, Lemma 2) *Let $W_0, W_1, \ldots, W_m$ be independently distributed random variables, $W_j \sim \chi^2_{2a_j}$, $j = 0, 1, \ldots, m$. If*

$$L = m^m \frac{\prod_{j=1}^{m} W_j}{(W_0 + \sum_{j=1}^{m} W_j)^m},$$

then $L \overset{d}{\sim} \prod_{j=1}^{m} X_j$, where $X_1, \ldots, X_m$ are independently distributed,

$$X_j \sim \beta\left(a_j, b_j\right), \qquad b_j = \frac{a + j - 1}{m} - a_j, \quad a = \sum_{j=0}^{m} a_j.$$

Lemma 3 *If $\boldsymbol{A} \sim \mathcal{W}_{qp}(\boldsymbol{\Sigma}, n)$, where $\boldsymbol{\Sigma}$ can be represented as (3), then*

$$\mathrm{BTr}_p[(\boldsymbol{V}_i \otimes \boldsymbol{I}_p)\boldsymbol{A}] \sim \mathcal{W}_p(\boldsymbol{\Delta}_i, n \cdot \mathrm{tr}\, \boldsymbol{V}_i).$$

Proof See Appendix B. □

We can now formulate the following theorem.

Theorem 3 *If the null hypothesis of (4) holds and $n > p$, then*

$$\mathrm{LRT} \overset{\mathrm{d}}{\sim} \begin{cases} -n \sum_{i=1}^{\nu} \sum_{j=2}^{\omega} \sum_{k=1}^{u_j} v_i \ln B_{ijk} & \text{if} \quad u_1 = 1 \\ -n \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \sum_{k=1}^{u_j} v_i \ln B_{ijk} & \text{if} \quad u_1 \geq 2, \end{cases}$$

where

$$B_{ijk} \sim \beta\left(\frac{(n-1)v_i}{2} - \frac{s_{jk} - 1}{2}, \frac{k-1}{u_j} + \frac{s_{jk} - 1}{2}\right)$$

and $s_{jk} = k + \sum_{c=0}^{j-1} u_c$, $u_0 = 0$.

Proof Under H_0 of (4), let us represent each matrix $\boldsymbol{U}_j$ generating the basis of the quadratic subspace $\mathcal{U}$ as $\boldsymbol{G}_j\boldsymbol{G}_j'$, $j = 1, \ldots, \omega$, where $\boldsymbol{G}_j'\boldsymbol{G}_j = \boldsymbol{I}_{u_j}$ and $\boldsymbol{G}_j'\boldsymbol{G}_{j'} = \boldsymbol{0}_{u_j \times u_{j'}}$ and let $\boldsymbol{G} = (\boldsymbol{G}_1, \ldots, \boldsymbol{G}_\omega)$. From the orthogonality of $\boldsymbol{G}$ we have $\left|\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}\right| =$

$|\boldsymbol{S}_i|$, $i = 1, \ldots, \nu$. Moreover, in Theorem 2 each $t_{ij} = \operatorname{tr}(\boldsymbol{S}_i \boldsymbol{U}_j)$ is equal to the sum of respective diagonal entries of matrices $\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}$. Thus, $\Lambda^{2/n}$ can be written as

$$\Lambda^{2/n} = \prod_{i=1}^{\nu} \left\{ \frac{\left|\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}\right| \prod_{j=1}^{\omega} u_j^{u_j}}{\prod_{j=1}^{\omega} \left[\sum_{k=1}^{u_j} \left(\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}\right)_{s_{jk},s_{jk}}\right]^{u_j}} \right\}^{v_i}, \tag{12}$$

where $s_{jk} = k + \sum_{c=0}^{j-1} u_c$ and $u_0 = 0$. Due to Lemma 3, the matrices

$$\boldsymbol{S}_i \sim \mathcal{W}_p(\boldsymbol{\Delta}_i, (n-1)v_i),$$

$i = 1, \ldots, \nu$, are mutually independent, and hence all the factors in (12) are independent. Furthermore, under the null hypothesis, if we denote by $\boldsymbol{D}_i$ the block-diagonal matrix with blocks $\delta_{ij}\boldsymbol{I}_{u_j}$, $j = 1, \ldots, \omega$, on the diagonal, then for every $i = 1, \ldots, \nu$ it holds that

$$\boldsymbol{D}_i^{-1/2}\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}\boldsymbol{D}_i^{-1/2} \sim \mathcal{W}_p(\boldsymbol{I}_p, (n-1)v_i).$$

Decomposing the above to $\boldsymbol{W}_i\boldsymbol{W}_i'$, where $\boldsymbol{W}_i$ is a lower triangular matrix with positive diagonal elements, we have that all elements of $\boldsymbol{W}_i$ are independent, the nonzero off-diagonal entries being normally distributed, that is, $W_{i,kk'} \sim N(0,1)$, $p \geq k > k' \geq 1$, and $W_{i,kk}^2 \sim \chi^2_{(n-1)v_i-k+1}$; cf. Kollo and von Rosen (2005, Theorem 2.4.4(i)). Since

$$|\boldsymbol{S}_i| = |\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}| = \left|\boldsymbol{D}_i^{1/2}\boldsymbol{W}_i\boldsymbol{W}_i'\boldsymbol{D}_i^{1/2}\right| = |\boldsymbol{D}_i| \cdot \prod_{k=1}^{p} W_{i,kk}^2 = \left(\prod_{j=1}^{\omega} \delta_{ij}^{u_j}\right) \prod_{k=1}^{p} W_{i,kk}^2,$$

we can write

$$\prod_{k=1}^{p} W_{i,kk}^2 = \prod_{j=1}^{\omega}\prod_{k=1}^{u_j} Z_{is_{jk}}^{(1)}, \tag{13}$$

where $Z_{is_{jk}}^{(1)} \sim \chi^2_{(n-1)v_i-s_{jk}+1}$. Note now that the factors in the denominator of (12) can be represented as

$$\begin{aligned}
\sum_{k=1}^{u_j} \left(\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}\right)_{s_{jk},s_{jk}} &= \delta_{ij} \sum_{k=1}^{u_j} \left(\boldsymbol{W}_i\boldsymbol{W}_i'\right)_{s_{jk},s_{jk}} \\
&= \begin{cases}
\delta_{ij} W_{i,11}^2 & \text{for } u_1 = 1 \\
\delta_{ij} \left(\sum_{k=1}^{u_1} W_{i,kk}^2 + \sum_{k=2}^{u_1}\sum_{k'=1}^{k-1} W_{i,kk'}^2\right) & \text{for } j = 1 \text{ and } u_1 \geq 2 \\
\delta_{ij} \left(\sum_{k=1}^{u_j} W_{i,s_{jk}s_{jk}}^2 + \sum_{k=1}^{u_j}\sum_{k'=1}^{s_{jk}-1} W_{i,s_{jk}k'}^2\right), & \text{for } j = 2, \ldots, \omega.
\end{cases}
\end{aligned}$$

Since each squared diagonal component of $\boldsymbol{W}_i$ is distributed as $Z^{(1)}_{is_{jk}}$, each squared off-diagonal component is distributed as $Z \sim \chi^2_1$, and all of these off-diagonal entries are independent, we obtain

$$\sum_{k=1}^{u_j} \left(\boldsymbol{G}'\boldsymbol{S}_i\boldsymbol{G}\right)_{s_{jk},s_{jk}} \overset{d}{\sim} \begin{cases} \delta_{ij} Z^{(1)}_{i1} & \text{for } u_1 = 1 \\ \delta_{ij}\left(\sum_{k=1}^{u_j} Z^{(1)}_{is_{jk}} + Z^{(2)}_{ij}\right) & \text{for } j = 1, \ldots, \omega \ \text{ and } \ u_1 \geq 2, \end{cases} \tag{14}$$

where $Z^{(2)}_{ij} \sim \chi^2_{d_j}$ with $d_j = (2s_{j0} + u_j - 1)u_j/2$. Substituting (13) and (14) into (12), we obtain the following:

$$\Lambda^{2/n} \overset{d}{\sim} \begin{cases} \prod_{i=1}^{\nu} \prod_{j=2}^{\omega} \left[\frac{u_j^{u_j} \cdot \prod_{k=1}^{u_j} Z^{(1)}_{is_{jk}}}{\left(\sum_{k=1}^{u_j} Z^{(1)}_{is_{jk}} + Z^{(2)}_{ij}\right)^{u_j}} \right]^{v_i} & \text{for } u_1 = 1 \\ \prod_{i=1}^{\nu} \prod_{j=1}^{\omega} \left[\frac{u_j^{u_j} \cdot \prod_{k=1}^{u_j} Z^{(1)}_{is_{jk}}}{\left(\sum_{k=1}^{u_j} Z^{(1)}_{is_{jk}} + Z^{(2)}_{ij}\right)^{u_j}} \right]^{v_i} & \text{for } u_1 \geq 2. \end{cases}$$

Now, using Lemma 2, we can write

$$\Lambda^{2/n} \overset{d}{\sim} \begin{cases} \prod_{i=1}^{\nu} \prod_{j=2}^{\omega} \prod_{k=1}^{u_j} B_{ijk}^{v_i} \text{ if } u_1 = 1 \\ \prod_{i=1}^{\nu} \prod_{j=1}^{\omega} \prod_{k=1}^{u_j} B_{ijk}^{v_i} \text{ if } u_1 \geq 2, \end{cases}$$

where

$$B_{ijk} \sim \beta\left(a^{(1)}_{is_{jk}}, \frac{a_{ij} + k - 1}{u_j} - a^{(1)}_{is_{jk}}\right)$$

with

$$a^{(1)}_{is_{jk}} = \frac{(n-1)v_i - s_{jk} + 1}{2} \quad \text{and} \quad a_{ij} = \sum_{k=1}^{u_j} a^{(1)}_{is_{jk}} + \frac{d_j}{2}.$$

Using the formulas for d_j and s_{jk}, the expression for a_{ij} can be reduced as follows:

$$\begin{aligned} a_{ij} &= \sum_{k=1}^{u_j} \frac{(n-1)v_i - s_{jk} + 1}{2} + \frac{(2s_{j0} + u_j - 1)u_j}{4} \\ &= u_j \frac{(n-1)v_i + 1}{2} - \frac{1}{2}\sum_{k=1}^{u_j} s_{jk} + \frac{\left(2\sum_{c=0}^{j-1} u_c + u_j - 1\right)u_j}{4} \end{aligned}$$

$$= u_j \frac{(n-1)v_i + 1}{2} - \frac{1}{2}\sum_{k=1}^{u_j}\left(\sum_{c=0}^{j-1} u_c + k\right) + \frac{\left(2\sum_{c=0}^{j} u_c - u_j - 1\right)u_j}{4}$$

$$= u_j \frac{(n-1)v_i + 1}{2} - \frac{u_j \cdot \sum_{c=0}^{j-1} u_c}{2} - \frac{u_j(u_j+1)}{4} + \frac{\left(2\sum_{c=0}^{j-1} u_c + u_j - 1\right)u_j}{4}$$

$$= \frac{u_j}{2}\left[(n-1)v_i + 1 - \sum_{c=0}^{j-1} u_c - \frac{u_j}{2} - \frac{1}{2} + \sum_{c=0}^{j} u_c + \frac{u_j}{2} - \frac{1}{2}\right]$$

$$= \frac{(n-1)u_j v_i}{2}$$

and the assertion follows. □

4 Comparison of tests – simulation studies

It is known (Rao 2005) that for increasing sample size the null distributions of both RST and LRT statistics tend to the χ^2 distribution with the number of degrees of freedom equal to the difference between the numbers of parameters under the alternative and null hypotheses. To assess the performance of the proposed tests, we conduct simulations to determine the speed of convergence to the limiting chi-square distribution and the empirical power, computed as the ratio between the number of rejected null hypotheses and the number of simulation runs. We consider the hypothesis

$$H_0 : \ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_{\mathrm{BCT\_CT}} \quad \text{vs} \quad H_1 : \ \boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_{\mathrm{BCT}}, \tag{15}$$

as one of the most general examples of hypothesis (4) related to known structures. Note also that for this hypothesis the number of degrees of freedom of the limiting distribution is $(p(p+1)/2 - \omega)\nu$.

To analyze the empirical null distribution of the RST and to obtain the quantiles necessary for determining the power of RST, for particular sets of parameters (p, q, n), 10,000 observation matrices from $N_{n,qp}(\mathbf{0}, \boldsymbol{I}_n, \boldsymbol{\Omega}_{\mathrm{BCT\_CT}})$ were generated. To obtain the exact null distribution of the LRT (and also its quantiles) we used the R package `CharFunToolR` developed in Gajdoš (2018). In both tests, we always set $n > p$.

The following subsections show that the RST outperforms the LRT with respect to the speed of convergence of the null distribution to the χ^2 distribution, while there is no significant difference in the behavior of the power of the tests: both of them are competitive.

4.1 Convergence to the limiting distribution

The empirical distribution of the RST and the exact distribution of the LRT, together with the limiting χ^2 distribution, for various combinations of p and q equal to 3 and 5 and to 3 and 10 (except $(p, q) = (3, 3)$) with $n \in \{p+1, 25, 50\}$ are presented in

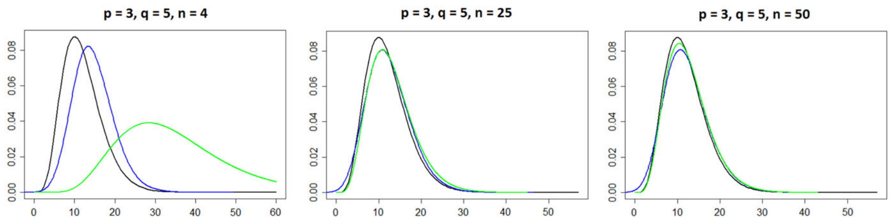


Fig. 1 Empirical null distribution of RST (blue) and exact distribution of LRT (green) together with the χ^2_{12} distribution (black) for $p = 3, q = 5$ (color figure online)

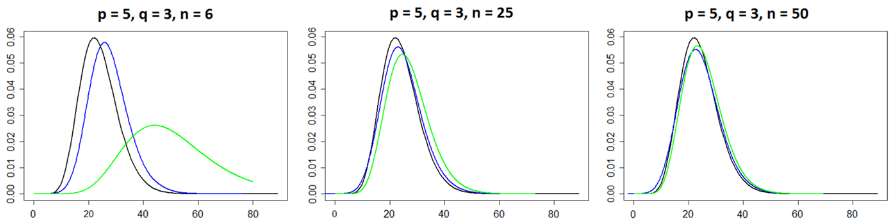


Fig. 2 Empirical null distribution of RST (blue) and exact distribution of LRT (green) together with the χ^2_{24} distribution (black) for $p = 5, q = 3$ (color figure online)

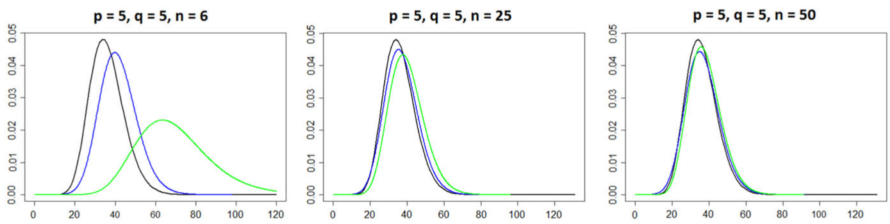


Fig. 3 Empirical null distribution of RST (blue) and exact distribution of LRT (green) together with the χ^2_{36} distribution (black) for $p = 5, q = 5$ (color figure online)

Figs. 1, 2, 3, 4, 5, 6. It is observed that with increasing n both null distributions tend to the limiting distribution; however, the convergence of the RST is faster than that of the LRT in all cases under consideration. Moreover, for relatively small p and q, the empirical null distribution of the RST is close to the limiting distribution even for a small sample size. In Figs. 1, 2, 3, with $n = 50$, the exact distribution center for the LRT seems to be closer to the chi-square distribution than the corresponding center for the RST. However, the tails, which are crucial for concluding a testing procedure, are indistinguishable for both null distributions. Note, however, that increasing p and q (especially p) slows down the convergence of both null distributions.

4.2 Power analysis

In this section we use a Monte Carlo simulation study to test the behavior of a power function of the RST and LRT with respect to the discrepancy between the null and alternative hypotheses, given respectively in (5) and (3), and with respect to the sample size. As a measure of discrepancy between two multivariate normal distributions with

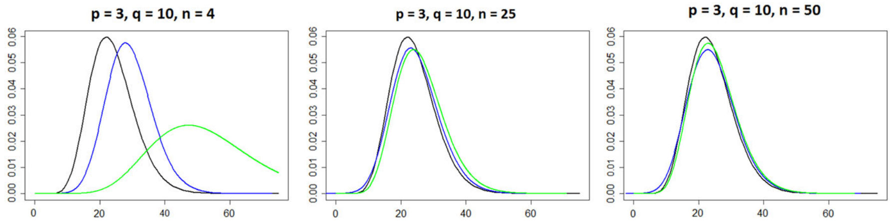


Fig. 4 Empirical null distribution of RST (blue) and exact distribution of LRT (green) together with the χ^2_{24} distribution (black) for $p = 3, q = 10$ (color figure online)

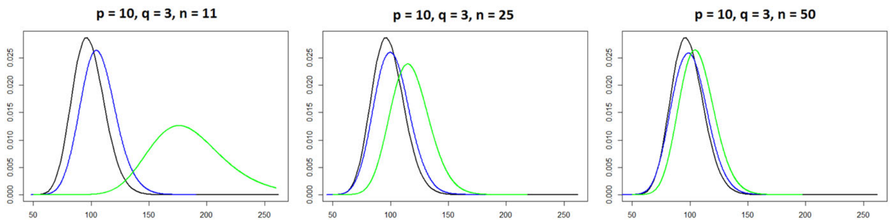


Fig. 5 Empirical null distribution of RST (blue) and exact distribution of LRT (green) together with the χ^2_{98} distribution (black) for $p = 10, q = 3$ (color figure online)

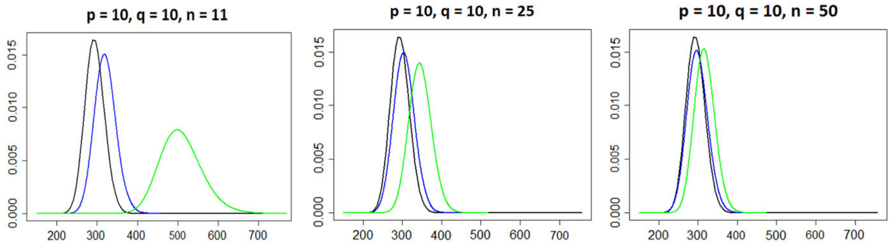


Fig. 6 Empirical null distribution of RST (blue) and exact distribution of LRT (green) together with the χ^2_{294} distribution (black) for $p = 10, q = 10$ (color figure online)

different covariance matrices, we use Kullback–Leibler divergence (cf. Stein 1956, James and Stein 1961, Dey and Srinivasan 1985, Lin et al. 2014), which for $\delta = \{\delta_{ij} : i = 1, \ldots, \nu,\ j = 1, \ldots, \omega\}$ and $\mathcal{D} = \{\boldsymbol{\Delta}_i : i = 1, \ldots, \nu\}$ has the form

$$\zeta(\delta, \mathcal{D}) = \operatorname{tr}\left[\left(\sum_{i=1}^{\nu} \boldsymbol{V}_i \otimes \boldsymbol{\Delta}_i^{-1}\right)\left(\sum_{i=1}^{\nu}\sum_{j=1}^{\omega} \delta_{ij}\left(\boldsymbol{V}_i \otimes \boldsymbol{U}_j\right)\right)\right] - \ln\left|\left(\sum_{i=1}^{\nu} \boldsymbol{V}_i \otimes \boldsymbol{\Delta}_i^{-1}\right)\left(\sum_{i=1}^{\nu}\sum_{j=1}^{\omega} \delta_{ij}\left(\boldsymbol{V}_i \otimes \boldsymbol{U}_j\right)\right)\right| - pq.$$

Thus, for a given alternative (the set $\mathcal{D}$), to determine the structure under the null hypothesis (the set δ) with the lowest discrepancy we should determine

$$\xi(\widetilde{\delta}) = \min_{\delta_{ij}} \zeta(\delta \mid \mathcal{D}).$$

Recalling that the matrices $\boldsymbol{V}_i$, $i = 1, \ldots, \nu$, are idempotent and orthogonal, and using (17), the problem simplifies to

$$\xi(\widetilde{\delta}) = \min_{\delta_{ij}} \left[\sum_{i=1}^{\nu} \sum_{j=1}^{\omega} v_i \delta_{ij} \operatorname{tr}\left(\boldsymbol{\Delta}_i^{-1} \boldsymbol{U}_j\right) + \ln \left| \sum_{i=1}^{\nu} \boldsymbol{V}_i \otimes \boldsymbol{\Delta}_i \right| - \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} v_i u_j \ln(\delta_{ij}) - pq \right].$$

Differentiating the above with respect to δ_{ij}, we obtain

$$\widetilde{\delta} = \left\{ \widetilde{\delta}_{ij} = \frac{u_j}{\operatorname{tr}(\boldsymbol{\Delta}_i^{-1} \boldsymbol{U}_j)} : i = 1, \ldots, \nu, \; j = 1, \ldots, \omega \right\}.$$

Since the Kullback-Leibler divergence function is not upper bounded, for power analysis we use a transformation of $\xi(\widetilde{\delta})$ in the form $\eta(\widetilde{\delta}) = 1 - 1/(1 - \xi(\widetilde{\delta}))$, to restrict the possible discrepancy to the interval [0,1).

For simulation purposes, we henceforth assume $\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_{\text{BCT}}$ and $\boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_{\text{BCT\_CT}}$.

In the first step we check the behavior of the power function of the RST and LRT with respect to the discrepancy $\eta(\widetilde{\delta})$. For this purpose we set the parameters of the experiment as $(p, q) \in \{(3, 5), (5, 3), (5, 5)\}$ and $n = 25$, and we generate 100 p.d. matrices Ω_{BCT}, for which the discrepancies $\eta(\widetilde{\delta})$ are computed. Then, for each $\boldsymbol{\Omega}_{\text{BCT}}$ we generate 10,000 observation matrices from $N_{n,qp}(\boldsymbol{\mu}, \boldsymbol{I}_n, \boldsymbol{\Omega}_{\text{BCT}})$, and for every observation matrix we test the hypothesis (15) using the quantiles of, respectively, the empirical null distribution of the RST and the exact null distribution of the LRT. In this way we obtain 100 values of the power of each test, computed as the ratio between the number of rejected null hypotheses and the number of simulation runs (10,000). The results are presented in Fig. 7. In all of these graphs the power shows an upward trend with increasing discrepancy, with slightly higher deviations between powers for the RST than for the LRT. A similar phenomenon is observed by Filipiak et al. (2023b), where various discrepancy measures are studied in the context of testing separability.

Finally, we compare the power of the RST and LRT with respect to the sample size. For this purpose we choose $p = 3$, $q = 5$, and $n \in \{4, 15, 25, 35, 50, 75, 100\}$, and we choose two matrices $\boldsymbol{\Omega}_{\text{BCT}}$ such that $\eta(\widetilde{\delta}) = 0.4$ and two further matrices for which $\eta(\widetilde{\delta}) = 0.6$. Each subgraph in Fig. 8 represents the empirical power of the RST (blue line) and LRT (green line) for 10,000 data matrices generated from $N_{n,qp}(\boldsymbol{\mu}, \boldsymbol{I}_n, \boldsymbol{\Omega}_{\text{BCT}})$, where the first column contains the results of simulation studies for two matrices $\boldsymbol{\Omega}_{\text{BCT}}$ with discrepancy equal to $\eta(\widetilde{\delta}) = 0.4$, and the second contains results for matrices $\boldsymbol{\Omega}_{\text{BCT}}$ with discrepancy $\eta(\widetilde{\delta}) = 0.6$ (the exact forms of these matrices are available from the authors upon request). It can be seen that both of the tests are competitively powerful: in some cases the power of the RST exceeds the power of the LRT, while in others the order is reversed. It is also observed that, as expected, the power of both tests increases with sample size, and the power of each test depends on the discrepancy.

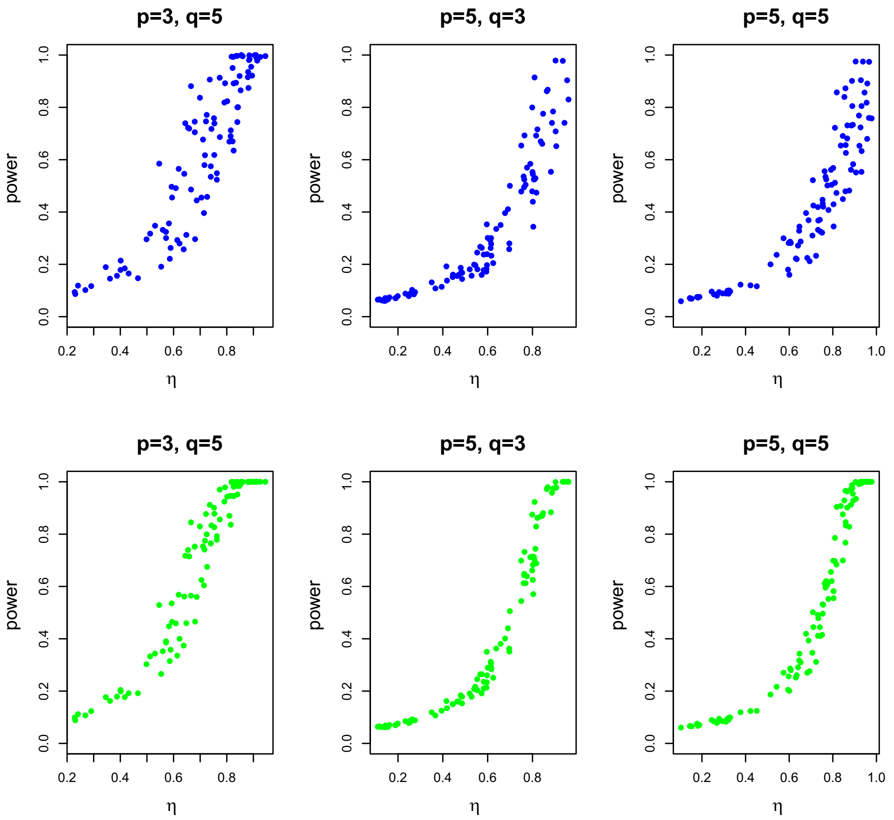


Fig. 7 Empirical power of RST (first row) and LRT (second row) with respect to discrepancy, for various combinations of p, q, and $n = 25$

5 Applications of the results—examples

In this section we illustrate the above results with examples based on simulated and real-life data sets. For this purpose we consider three special cases of hypothesis (4), namely

$$\begin{array}{llll} H_0^{(a)} : \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_{\mathrm{BCT\_CT}} & \text{vs } H_1^{(a)} : \boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_{\mathrm{BCT}}, \\ H_0^{(b)} : \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_{\mathrm{BCS\_CS}} & \text{vs } H_1^{(b)} : \boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_{\mathrm{BCS}}, \\ H_0^{(c)} : \boldsymbol{\Omega}_0 = \boldsymbol{I}_q \otimes \boldsymbol{\Omega}_{\mathrm{CT}} & \text{vs } H_1^{(c)} : \boldsymbol{\Omega}_1 = \boldsymbol{I}_q \otimes \boldsymbol{\Omega}_{\mathrm{UN}}. \end{array} \qquad (16)$$

To clarify how the values of RST and LRT are computed, in Appendix C we present the basis of the structures considered in (16). In Appendix C we also give the basis of each structure with respect to the parameters q and p; see Table 3 for a simulated example and Table 4 for a real-life example. Finally, also in Appendix C, due to Theorem 3, we show the parameters of the beta distribution used to construct the exact

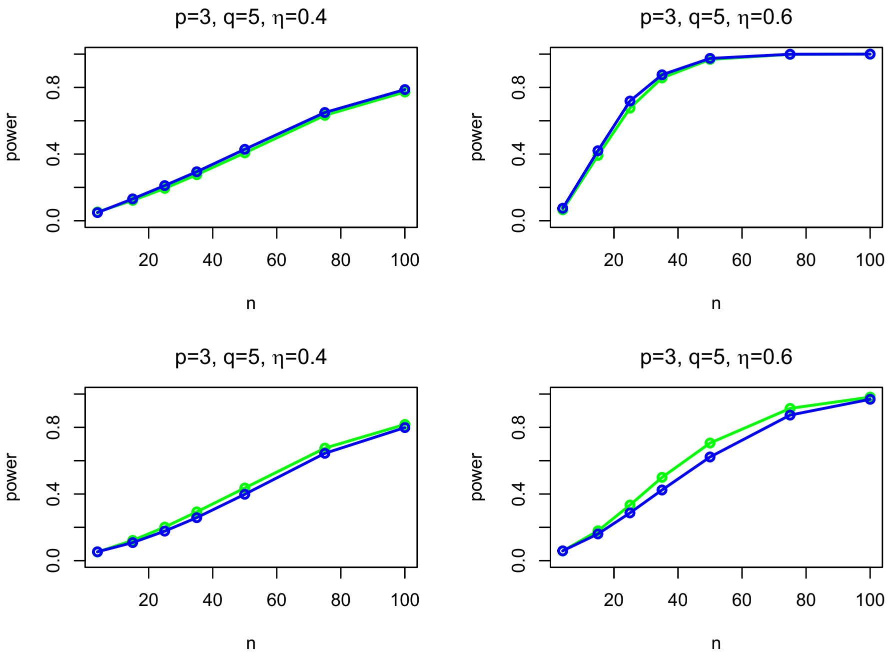


Fig. 8 Empirical power of RST (blue) and exact power of LRT (green) for $p = 3, q = 5$ and two alternatives with $\eta(\widetilde{\delta}) = 0.4$ (left column) and $\eta(\widetilde{\delta}) = 0.6$ (right column)

distribution of the LRT; see Table 5 for the simulated example and Table 6 for the real-life example.

5.1 Simulated data example

We generate the data set from $N_{n,qp}(\mathbf{0}, \boldsymbol{I}_n, \boldsymbol{\Omega}_{\text{BCS\_CS}})$ with $n = 25$, $q = 5$, $p = 7$ and $\boldsymbol{\Omega}_{\text{BCS\_CS}} = \boldsymbol{P}_{1_q} \otimes \boldsymbol{\Delta}_1 + \boldsymbol{Q}_{1_q} \otimes \boldsymbol{\Delta}_2$, where $\boldsymbol{\Delta}_1 = 2\boldsymbol{P}_{1_p} + 3\boldsymbol{Q}_{1_p}$, $\boldsymbol{\Delta}_2 = 3\boldsymbol{P}_{1_p} + 2\boldsymbol{Q}_{1_p}$, and the projectors $\boldsymbol{P}_1$, $\boldsymbol{Q}_1$ of relevant order are as defined in Sect. 3; the data set is available from the authors upon request.

For every hypothesis in (16), we computed the values of the RST and LRT statistics, and also determined the empirical p-value for the RST using an empirical null distribution, and the exact p-value for the LRT using the `CharFunToolR` package in R (cf. Gajdoš 2018). The results are presented in Table 1, together with the p-values of the limiting χ^2 distribution with the following degrees of freedom: $(p(p+1)/2-\omega)\nu = 72$ under $H_0^{(a)}$, $p(p+1) - 4 = 52$ under $H_0^{(b)}$, and $p(p+1)/2 - \omega = 24$ under $H_0^{(c)}$. Moreover, in Fig. 9 we present the empirical null distribution of the RST (blue line), the exact distribution of the LRT (green line), and the limiting χ^2 distribution with the respective number of degrees of freedom (black line) for each hypothesis.

It can be seen that for $n = 25$ the distribution of both the RST and LRT is close to the limiting distribution. Thus, assuming $\alpha = 0.1$, the decision taken based on the empirical/exact distribution and the limiting distribution remains the same: reject

Table 1 Test statistics and respective p-values for hypotheses (16) for a simulated data set

	(a)		(b)		(c)	
	RST	LRT	RST	LRT	RST	LRT
Statistic	101.154	107.681	65.517	68.386	41.062	38.460
Empirical / exact p-val	0.022	0.037	0.129	0.189	0.028	0.054
χ^2 p-val	0.013	0.004	0.099	0.063	0.016	0.031

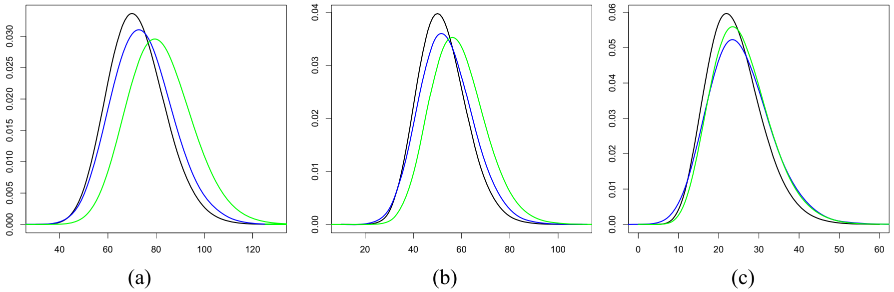


Fig. 9 Empirical null distribution of RST statistic (blue) and exact distribution of LRT statistic (green) together with the respective χ^2 distribution (black) for $n = 25$, $q = 5$, $p = 7$, and null hypotheses in (16)

$H_0^{(a)}$, fail to reject $H_0^{(b)}$, and reject $H_0^{(c)}$. For $\alpha = 0.05$ we observe that for hypothesis (c) the p-value taken from the exact distribution of LRT slightly exceeds the nominal level, while for the limiting distribution it is smaller than 0.05. This seems surprising, especially given that the null distribution of the LRT is closer to the chi-square distribution than in the case of hypothesis (a), which can be observed either directly from Fig. 9 or from the difference between the exact and limiting p-values for both of the hypotheses. However, when we look into the values of the LRT statistics, we observe that the area under the curve is larger for hypothesis (c) than for hypothesis (a). Moreover, when we compare the values of the LRT and RST for hypotheses (c) and (a), we see that in the case of (c) the value of the RST is higher than that of the LRT, while for hypothesis (a) the order is reversed.

Finally note that, as expected from the assumptions of the experiment, both tests fail to reject $H_0^{(b)}$.

5.2 Real data example

In this section we consider the real data example presented in Liang et al. (2015). The data comes from an experiment in which $n = 11$ plants were selected, in each plant three flowers were randomly chosen ($q = 3$), and then the lengths of all four petals from each flower were measured ($p = 4$). As observed in Liang et al. (2015), since the arrangement of petals in the flower is circular, it is reasonable to assume that the correlation between the lengths of any two petals within a single flower depends on the number of petals between them. In other words, we assume equal correlations between observations from neighboring petals and equal correlations between obser-

Table 2 Test statistics and respective p-values for hypotheses (16) for real data

	(a)		(b)		(c)	
	RST	LRT	RST	LRT	RST	LRT
Statistic	20.82	24.89	23.98	27.56	12.36	12.01
Empirical / exact p-val	0.133	0.159	0.107	0.158	0.111	0.177
χ^2 p-val	0.106	**0.036**	0.090	**0.036**	0.089	0.100

vations from opposite petals. Hence, our aim is to test the circular Toeplitz structure of intraclass dependence between petals on a single flower. Moreover, since the order of flowers on a single plant is not important, a compound symmetry structure can be applied to describe the inter-flower correlation (interclass dependence). Observe, however, that since only three flowers are selected, the compound symmetry structure is equivalent to the circular Toeplitz structure. Thus, in this case hypothesis (16)(a) may be of interest.

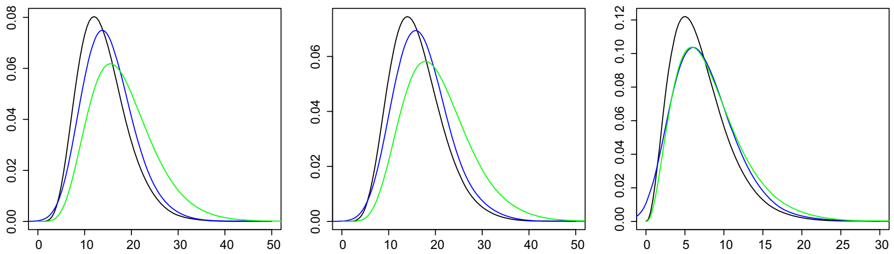
If the researcher is additionally interested in verifying whether the correlation between the length of petals depends on the distance between petals, it is reasonable to study hypothesis (16)(b).

Another possibility is to test whether the measurements made on different flowers are uncorrelated, while the lengths of petals on a particular flower depend on the distance between the petals. In this case, hypothesis (16)(c) is of interest.

For the data under consideration, a hypothesis concerning the particular structure of the mean and circular Toeplitz or compound symmetry structure of covariance between petals, assuming a compound symmetry structure of covariance between flowers, was tested in Liang et al. (2022) using the LRT. Similarly to Liang et al. (2022), in this paper we also test hypotheses (a) and (b) from (16); however, since we do not consider any special structure of expectation, our results differ from those of Liang et al. (2022). Finally, to illustrate the LRT and RST formulated in a general form, we additionally test the hypothesis of block-sphericity with circular Toeplitz blocks given in (c) in (16).

The computed values of the RST and LRT statistics, together with corresponding p-values obtained using the exact and limiting chi-square distributions, are presented in Table 2. The empirical p-value of the RST is determined from the empirical null distribution of RST, while the p-value of the LRT is computed from the exact distribution of the LRT determined using the `CharFunToolR` package (cf. Gajdoš, 2018). Note that the number of degrees of freedom of the limiting chi-square distribution is $(p(p+1)/2 - \omega)\nu = 14$ under $H_0^{(a)}$, $p(p+1) - 4 = 16$ under $H_0^{(b)}$, and $p(p+1)/2 - \omega = 7$ under $H_0^{(c)}$.

Assuming a significance level of $\alpha = 0.05$, it is observed that for hypotheses (a) and (b) the decisions based on the exact and limiting distributions of the LRT differ (thus, the respective p-values of chi-square distributions are given in bold). This is because the exact distribution of the LRT is quite distant from the limiting distribution for a small sample size ($n = 11$); cf. Fig. 10a, b. Such a problem does not arise either for hypothesis (c) or for the inference based on the RST.



(a) (b) (c)

Fig. 10 Empirical null distribution of RST statistic (blue) and exact distribution of LRT statistic (green) together with the respective χ^2 distribution (black) for $n = 11$, $q = 3$, $p = 4$, and null hypotheses in (16) (color figure online)

Note that, reasoning from the empirical or limiting distribution of the RST or the exact distribution of the LRT, all three null hypotheses are not rejected. Thus, one may ask which structure to choose for further analysis. One approach would be to choose the structure with the smallest p-values. Another good choice would be the structure with the smallest number of parameters, because in this case more degrees of freedom are left for estimation of the expectation.

Summing up, both tests work well; however, for the LRT the inference should be based on the exact rather than the limiting distribution, and this requires the use of specialized software and packages. Thus, we would recommend to practitioners the use of the RST, as the limiting chi-square distribution is a good approximation of the distribution even for a relatively small sample size, and is well known to researchers.

6 Final remarks

In this paper we have shown, via simulations, that for testing (4) the RST performs better than the LRT in the sense of the speed of convergence to the limiting chi-square distribution when $n \to \infty$ and p, q are fixed. Since the exact distribution of the RST remains unknown, we are not aware of any theoretical results supporting the conclusions drawn from the simulation studies. Moreover, one can see that the condition $n > p$ imposed for both tests, where p is fixed, is required. Obviously, because the relation (6) makes it possible to exchange the role of interclass and intraclass structures, in such a case the requirement $n > p$ will be replaced by $n > q$. Nevertheless, it would be beneficial to develop tests for the high-dimensional case where both p (respectively q) and n tend to infinity, with $p/n \to c \in (0, 1]$ (respectively $q/n \to c \in (0, 1]$). However, the study in high-dimensional setup appears much more challenging, and hence will be the topic of future research.

Acknowledgements The authors thank to the Referees for very constructive comments. Special thanks go to the International Banach Center of the Polish Academy of Sciences for providing the opportunity and support for this paper.
This research is partially supported by the Rector of Poznań University of Technology, project no. 0213/SIGR/2154 (Mateusz John).

Funding Open access funding provided by Linnaeus University.

Appendix A: Proof of Proposition 1

Proof Assume the covariance structure under H_0. Because $\mathcal{V} \otimes \mathcal{U}$ is a commutative quadratic subspace, from Filipiak et al. (2020, Theorem 4.1), the MLE of $\boldsymbol{\Omega}_0$ can be obtained by projecting $\widehat{\boldsymbol{\Omega}}_{\text{UN}}$ as given in (8) onto the orthonormal basis of $\mathcal{V} \otimes \mathcal{U}$. Thus,

$$\widehat{\boldsymbol{\Omega}}_0 = \sum_{i=1}^{\nu} \sum_{j=1}^{\omega} \left\langle \tfrac{1}{n}\boldsymbol{S}, \left(\tfrac{1}{\sqrt{v_i}}\boldsymbol{V}_i \otimes \tfrac{1}{\sqrt{u_j}}\boldsymbol{U}_j\right)\right\rangle \left(\tfrac{1}{\sqrt{v_i}}\boldsymbol{V}_i \otimes \tfrac{1}{\sqrt{u_j}}\boldsymbol{U}_j\right),$$

where $\langle \bullet, \bullet \rangle$ denotes the standard inner product. Hence, after simple calculations we obtain the MLE of δ_{ij}, $i = 1, \ldots, \nu$, $j = 1, \ldots, \omega$, given in the Proposition.

Assume the covariance structure under H_1. Note that $\boldsymbol{\Omega}_1^{-1} = \sum_{i=1}^{\nu} \boldsymbol{V}_i \otimes \boldsymbol{\Delta}_i^{-1}$. Moreover,

$$\begin{aligned} |\boldsymbol{\Omega}_1| &= \left|(\boldsymbol{H}' \otimes \boldsymbol{I}_p)\boldsymbol{\Omega}_1(\boldsymbol{H} \otimes \boldsymbol{I}_p)\right| \\ &= \left|\sum_{i=1}^{\nu} \boldsymbol{H}_i'\boldsymbol{V}_i\boldsymbol{H}_i \otimes \boldsymbol{\Delta}_i\right| = \left|\sum_{i=1}^{\nu} \boldsymbol{I}_{v_i} \otimes \boldsymbol{\Delta}_i\right| = \prod_{i=1}^{\nu} |\boldsymbol{\Delta}_i|^{v_i}, \end{aligned} \tag{17}$$

where $\boldsymbol{H} = (\boldsymbol{H}_1, \ldots, \boldsymbol{H}_\nu)$. Therefore, differentiating the log-likelihood function

$$\ln L(\boldsymbol{\Omega}) = -\frac{nqp}{2}\ln(2\pi) - \frac{n}{2}\sum_{i=1}^{\nu}(v_i \cdot \ln|\boldsymbol{\Delta}_i|) - \frac{1}{2}\sum_{i=1}^{\nu}\operatorname{tr}\left[\boldsymbol{S}(\boldsymbol{V}_i \otimes \boldsymbol{\Delta}_i^{-1})\right]$$

with respect to $\boldsymbol{\Delta}_i$, we obtain ν normal equations of the form

$$n \cdot v_i \cdot \operatorname{vec}\boldsymbol{\Delta}_i = (\operatorname{vec}'\boldsymbol{V}_i \otimes \boldsymbol{I}_{p^2})(\boldsymbol{I}_q \otimes \boldsymbol{K}_{q,p} \otimes \boldsymbol{I}_p)\operatorname{vec}\boldsymbol{S},$$

where $\boldsymbol{K}_{q,p}$ is the commutation matrix; cf. Kollo and von Rosen (2005). From Filipiak et al. (2018, Corollary 2.10) we obtain the MLEs of $\boldsymbol{\Delta}_i$ presented in the Proposition. □

Appendix B: Proof of Lemma 3

Proof Let $\boldsymbol{A} \sim \mathcal{W}_{qp}(\boldsymbol{\Sigma}, n)$, where $\boldsymbol{\Sigma}$ can be represented as (3). Then, writing $\boldsymbol{H} = (\boldsymbol{H}_1, \ldots, \boldsymbol{H}_\nu)$, we have

$$(\boldsymbol{H}' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{H} \otimes \boldsymbol{I}_p) \sim \mathcal{W}_{qp}(\mathrm{BDiag}(\boldsymbol{I}_{v_i} \otimes \boldsymbol{\Delta}_i), n)$$

(cf. Kollo and von Rosen, 2005, Th. 2.4.2), where $\mathrm{BDiag}(\cdot)$ is a block diagonal matrix. This implies that the diagonal blocks of the above $(\boldsymbol{H}_i' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{H}_i \otimes \boldsymbol{I}_p)$ and $(\boldsymbol{H}_j' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{H}_j \otimes \boldsymbol{I}_p)$, $i, j = 1, \ldots, \nu$, $i \neq j$, are independent, and

$$(\boldsymbol{H}_i' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{H}_i \otimes \boldsymbol{I}_p) \sim \mathcal{W}_{v_i p}(\boldsymbol{I}_{v_i} \otimes \boldsymbol{\Delta}_i, n), \qquad i = 1, \ldots, \nu;$$

cf. Anderson (2003, Theorem 7.3.5, p. 261). Moreover, since each $\boldsymbol{H}_i$ consists of orthogonal columns, $\boldsymbol{H}_i = (\boldsymbol{h}_{i1}, \ldots, \boldsymbol{h}_{iv_i})$, the diagonal blocks of $(\boldsymbol{H}_i' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{H}_i \otimes \boldsymbol{I}_p)$, $i = 1, \ldots, \nu$, are independent, and

$$(\boldsymbol{h}_{ik}' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{h}_{ik} \otimes \boldsymbol{I}_p) \sim \mathcal{W}_p(\boldsymbol{\Delta}_i, n), \qquad i = 1, \ldots, \nu, \quad k = 1, \ldots, v_i.$$

Note now that, since

$$\mathrm{BTr}_p\left[(\boldsymbol{H}_i' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{H}_i \otimes \boldsymbol{I}_p)\right] = \sum_{k=1}^{v_i}(\boldsymbol{h}_{ik}' \otimes \boldsymbol{I}_p)\boldsymbol{A}(\boldsymbol{h}_{ik} \otimes \boldsymbol{I}_p)$$

from Anderson (2003, Theorem 7.3.2, p. 260) , we obtain the assertion. □

Appendix C: Basis of covariance structures and parameters of beta distribution for simulated and real data examples

To clarify how the values of the RST and LRT are computed, we give here the basis of the structures considered in (16). For simplicity we define the basis that follows from the spectral decomposition, consisting of τ matrices $\boldsymbol{W}_k$ of size $t \times t$, which for the interclass structure become a basis of ν matrices $\boldsymbol{V}_i$ of size $q \times q$, and for the intraclass structure a basis of ω matrices $\boldsymbol{U}_j$ of size $p \times p$. Denoting by $\boldsymbol{L}$ the orthogonal matrix whose columns are the eigenvectors of the relevant covariance structure, we can construct $\boldsymbol{W}_k$ as a product $\boldsymbol{L}_k\boldsymbol{L}_k'$, where:

- for matrices proportional to the identity: $\tau = 1$ and $\boldsymbol{L}_1 = \boldsymbol{I}_t$;
- for CS matrices: $\tau = 2$ and $\boldsymbol{L} = (\boldsymbol{L}_1 : \boldsymbol{L}_2)$, with $\boldsymbol{L}_1 = \frac{1}{\sqrt{t}}\mathbf{1}_t$ and $\boldsymbol{L}_2$ consisting of $t-1$ eigenvectors being orthogonal to $\mathbf{1}_t$ (cf. Olkin and Press, 1969);
- for CT matrices: $\tau = \lfloor t/2 \rfloor + 1$ and $\boldsymbol{L} = (\frac{1}{\sqrt{t}}\mathbf{1}_t, \boldsymbol{\ell}_2, \ldots, \boldsymbol{\ell}_t)$, with the ath element of $\boldsymbol{\ell}_b$, $a \in \{1, \ldots, t\}$, $b \in \{2, \ldots, t\}$, given by

$$\ell_b^a = \tfrac{1}{\sqrt{t}}\left[\cos\left(\tfrac{2\pi}{t}(b-1)(a-1)\right) + \sin\left(\tfrac{2\pi}{t}(b-1)(a-1)\right)\right]$$

Table 3 The basis of the covariance structure related to the hypotheses considered in (16) for $q = 5$ and $p = 7$

Hypothesis	Interclass			intraclass		
	ν	v_i	$\boldsymbol{V}_i$	ω	u_j	$\boldsymbol{U}_j$
(a)	3	$v_1 = 1$	$\boldsymbol{V}_1 = \boldsymbol{P}_{1_5}$	4	$u_1 = 1$	$\boldsymbol{U}_1 = \boldsymbol{P}_{1_7}$
		$v_2 = 2$	$\boldsymbol{V}_2 = \boldsymbol{\ell}_2\boldsymbol{\ell}_2' + \boldsymbol{\ell}_5\boldsymbol{\ell}_5'$		$u_2 = 2$	$\boldsymbol{U}_2 = \boldsymbol{\ell}_2\boldsymbol{\ell}_2' + \boldsymbol{\ell}_7\boldsymbol{\ell}_7'$
		$v_3 = 2$	$\boldsymbol{V}_3 = \boldsymbol{\ell}_3\boldsymbol{\ell}_3' + \boldsymbol{\ell}_4\boldsymbol{\ell}_4'$		$u_3 = 2$	$\boldsymbol{U}_3 = \boldsymbol{\ell}_3\boldsymbol{\ell}_3' + \boldsymbol{\ell}_6\boldsymbol{\ell}_6'$
					$u_4 = 2$	$\boldsymbol{U}_4 = \boldsymbol{\ell}_4\boldsymbol{\ell}_4' + \boldsymbol{\ell}_5\boldsymbol{\ell}_5'$
(b)	2	$v_1 = 1$	$\boldsymbol{V}_1 = \boldsymbol{P}_{1_5}$	2	$u_1 = 1$	$\boldsymbol{U}_1 = \boldsymbol{P}_{1_7}$
		$v_2 = 4$	$\boldsymbol{V}_2 = \boldsymbol{Q}_{1_5}$		$u_2 = 6$	$\boldsymbol{U}_2 = \boldsymbol{Q}_{1_7}$
(c)	1	$v_1 = 5$	$\boldsymbol{V}_1 = \boldsymbol{I}_5$	4	$u_1 = 1$	$\boldsymbol{U}_1 = \boldsymbol{P}_{1_7}$
					$u_2 = 2$	$\boldsymbol{U}_2 = \boldsymbol{\ell}_2\boldsymbol{\ell}_2' + \boldsymbol{\ell}_7\boldsymbol{\ell}_7'$
					$u_3 = 2$	$\boldsymbol{U}_3 = \boldsymbol{\ell}_3\boldsymbol{\ell}_3' + \boldsymbol{\ell}_6\boldsymbol{\ell}_6'$
					$u_4 = 2$	$\boldsymbol{U}_4 = \boldsymbol{\ell}_4\boldsymbol{\ell}_4' + \boldsymbol{\ell}_5\boldsymbol{\ell}_5'$

Table 4 Basis of the covariance structure related to the hypotheses considered in (16) for $q = 3$ and $p = 4$

Hypothesis	Interclass			Intraclass		
	ν	v_i	$\boldsymbol{V}_i$	ω	u_j	$\boldsymbol{U}_j$
(a)	2	$v_1 = 1$	$\boldsymbol{V}_1 = \boldsymbol{P}_{1_3}$	3	$u_1 = 1$	$\boldsymbol{U}_1 = \boldsymbol{P}_{1_4}$
		$v_2 = 2$	$\boldsymbol{V}_2 = \boldsymbol{Q}_{1_3}$		$u_2 = 2$	$\boldsymbol{U}_2 = \boldsymbol{\ell}_2\boldsymbol{\ell}_2' + \boldsymbol{\ell}_4\boldsymbol{\ell}_4'$
					$u_3 = 1$	$\boldsymbol{U}_3 = \boldsymbol{\ell}_3\boldsymbol{\ell}_3'$
(b)	2	$v_1 = 1$	$\boldsymbol{V}_1 = \boldsymbol{P}_{1_3}$	2	$u_1 = 1$	$\boldsymbol{U}_1 = \boldsymbol{P}_{1_4}$
		$v_2 = 2$	$\boldsymbol{V}_2 = \boldsymbol{Q}_{1_3}$		$u_2 = 3$	$\boldsymbol{U}_2 = \boldsymbol{Q}_{1_4}$
(c)	1	$v_1 = 3$	$\boldsymbol{V}_1 = \boldsymbol{I}_3$	3	$u_1 = 1$	$\boldsymbol{U}_1 = \boldsymbol{P}_{1_4}$
					$u_2 = 2$	$\boldsymbol{U}_2 = \boldsymbol{\ell}_2\boldsymbol{\ell}_2' + \boldsymbol{\ell}_4\boldsymbol{\ell}_4'$
					$u_3 = 1$	$\boldsymbol{U}_3 = \boldsymbol{\ell}_3\boldsymbol{\ell}_3'$

(cf. Basilevsky, 1983, Olkin and Press, 1969), where $\boldsymbol{L}_1 = \frac{1}{\sqrt{t}}\boldsymbol{1}_t$, $\boldsymbol{L}_k = (\boldsymbol{\ell}_k : \boldsymbol{\ell}_{t-k+2})$, $k = 2, \ldots, \tau - 1$, and $\boldsymbol{L}_\tau = \begin{cases} (\boldsymbol{\ell}_\tau : \boldsymbol{\ell}_{\tau+1}) & \text{if } t \text{ is odd} \\ \boldsymbol{\ell}_\tau & \text{if } t \text{ is even.} \end{cases}$

It is also worthwhile to note that for the CS structure, $\boldsymbol{W}_1$ is simply $\boldsymbol{P}_{1_t}$, while $\boldsymbol{W}_2 = \boldsymbol{Q}_{1_t}$, defined in Sect. 3.

In Tables 3 and 4 we give the basis of structures considered for each hypothesis, respectively for $q = 5$, $p = 7$ and $q = 3$, $p = 4$. Recall that in each case the same interclass structure is assumed in the null and alternative hypotheses, while the intraclass dependency has a specified structure in the null hypotheses and remains unstructured in each alternative. Therefore, the basis matrices $\boldsymbol{V}_i$ for the interclass structure in each alternative hypothesis is the same as in the respective null hypothesis.

In Tables 5 and 6 we give the parameters of the beta distributions used for construction of the exact distribution of the LRT, presented in Theorem 3, respectively

Table 5 Parameters of beta distributions related to the hypotheses considered in (16) for $q = 5$ and $p = 7$

Hypothesis (a) , $\nu = 3, \omega = 4, v_1 = 1, v_2 = v_3 = 2, u_1 = 1, u_2 = u_3 = u_4 = 2$				
j	k	$s_{jk} - 1$	beta distribution parameters	
2	1	1	$[(n-1)v_i - 1]/2$	1/2
	2	2	$[(n-1)v_i - 2]/2$	3/2
3	1	3	$[(n-1)v_i - 3]/2$	3/2
	2	4	$[(n-1)v_i - 4]/2$	5/2
4	1	5	$[(n-1)v_i - 5]/2$	5/2
	2	6	$[(n-1)v_i - 6]/2$	7/2
Hypothesis (b) , $\nu = 2, \omega = 2, v_1 = 1, v_2 = 4, u_1 = 1, u_2 = 6$				
j	k	$s_{jk} - 1$	beta distribution parameters	
2	1	1	$[(n-1)v_i - 1]/2$	1/2
	2	2	$[(n-1)v_i - 2]/2$	3/2
	3	3	$[(n-1)v_i - 3]/2$	5/2
	4	4	$[(n-1)v_i - 4]/2$	7/2
	5	5	$[(n-1)v_i - 5]/2$	9/2
	6	6	$[(n-1)v_i - 6]/2$	11/2
Hypothesis (c) , $\nu = 1, \omega = 4, v_1 = 5, u_1 = 1, u_2 = u_3 = u_4 = 2$				
j	k	$s_{jk} - 1$	beta distribution parameters	
2	1	1	$[(n-1)q - 1]/2$	1/2
	2	2	$[(n-1)q - 2]/2$	3/2
3	1	3	$[(n-1)q - 3]/2$	3/2
	2	4	$[(n-1)q - 4]/2$	5/2
4	1	5	$[(n-1)q - 5]/2$	5/2
	2	6	$[(n-1)q - 6]/2$	7/2

for $q = 5$, $p = 7$ and $q = 3$, $p = 4$. Note that in the examples considered $u_1 = 1$ for each hypothesis, and thus each set of parameters begins with $j = 2$.

Table 6 Parameters of beta distributions related to the hypotheses considered in (16) for $q = 3$ and $p = 4$

Hypothesis (a), $\nu = 2, \omega = 3, v_1 = 1, v_2 = 2, u_1 = 1, u_2 = 2, u_3 = 1$				
j	k	$s_{jk} - 1$	beta distribution parameters	
2	1	1	$[(n-1)v_i - 1]/2$	1/2
	2	2	$[(n-1)v_i - 2]/2$	3/2
3	1	3	$[(n-1)v_i - 3]/2$	3/2
Hypothesis (b), $\nu = 2, \omega = 2, v_1 = 1, v_2 = 2, u_1 = 1, u_2 = 3$				
j	k	$s_{jk} - 1$	beta distribution parameters	
2	1	1	$[(n-1)v_i - 1]/2$	1/2
	2	2	$[(n-1)v_i - 2]/2$	3/2
	3	3	$[(n-1)v_i - 3]/2$	5/2
Hypothesis (c), $\nu = 1, \omega = 3, v_1 = 3, u_1 = 1, u_2 = 2, u_3 = 1$				
j	k	$s_{jk} - 1$	beta distribution parameters	
2	1	1	$[(n-1)q - 1]/2$	1/2
	2	2	$[(n-1)q - 2]/2$	3/2
3	1	3	$[(n-1)q - 3]/2$	3/2

References

Anderson TW (2003) Introduction to multivariate statistics, 3rd edn. John Wiley & Sons, Hoboken, New Jersey

Arnold SF (1973) Application of the theory of products of problems to certain patterned covariance matrices. Ann Stat 1(4):682–699

Arnold SF (1979) Linear models with exchangeably distributed errors. J Am Stat Assoc 74(365):194–199

Bai Z, Jiang D, Yao JF et al (2009) Corrections to lrt on large dimensional covariance matrix by rmt. Ann Stat 37(6B):3822–3840

Basilevsky A (1983) Applied matrix algebra in the statistical sciences. North-Holland

Chi EM, Reinsel GC (1989) Models for longitudinal data with random effects and ar (1) errors. J Am Stat Assoc 84(406):452–459

Dey DK, Srinivasan C (1985) Estimation of a covariance matrix under stein's loss. Ann Stat 13(4):1581–1591

Eaton ML (1983) Multivariate statistics: a vector space approach. Wiley, New York

Filipiak K, Klein D, Roy A (2016) Score test for a separable covariance structure with the first component as compound symmetric correlation matrix. J Multivar Anal 150:105–124

Filipiak K, Klein D, Roy A (2017) A comparison of likelihood ratio tests and rao's score test for three separable covariance matrix structures. Biom J 59(1):192–215

Filipiak K, Klein D, Vojtková E (2018) The properties of partial trace and block trace operators of partitioned matrices. Electron J Linear Algebra 33:3–15

Filipiak K, John M, Markiewicz A (2020) Comments on maximum likelihood estimation and projections under multivariate statistical models. In: Recent developments in multivariate and random matrix analysis. Springer, p 51–66

Filipiak K, John M, Klein D (2023) Testing independence under a block compound symmetry covariance structure. Stat Pap 64(2):677–704

Filipiak K, Klein D, Mokrzycka M (2023b) Discrepancy between structured matrices in the context of power analysis, under revision

Gajdoš A (2018) CharfuntoolR: the characteristic functions toolbox (R). Available from: https://github.com/gajdosandrej/CharFunToolR

James W, Stein C (1961) Estimation with quadratic loss. In: Proceedings of the fourth berkeley symposium on mathematical statistics and probability: the statistical laboratory, University of California, June 20-July 30, 1960, University of California Press, pp 361–379

Jiang D (2016) Tests for large-dimensional covariance structure based on rao's score test. J Multivar Anal 152:28–39
Jiang D, Jiang T, Yang F (2012) Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. J Stat Plan Inference 142(8):2241–2256
John S (1971) Some optimal multivariate tests. Biometrika 58(1):123–127
Kato N, Yamada T, Fujikoshi Y (2010) High-dimensional asymptotic expansion of lr statistic for testing intraclass correlation structure and its error bound. J Multivar Anal 101(1):101–112
Klein D, Pielaszkiewicz J, Filipiak K (2022) Approximate normality in testing an exchangeable covariance structure under large-and high-dimensional settings. J Multivar Anal 192:105049
Kollo T, von Rosen D (2005) Advanced multivariate statistics with matrices. Springer-Verlag, Dordrecht
Lai J, Wang X, Zhao K et al (2023) Block-diagonal test for high-dimensional covariance matrices. TEST 32(1):447–466
Liang Y, von Rosen D, von Rosen T (2015) On estimation in hierarchical models with block circular covariance structures. Ann Inst Stat Math 67(4):773–791
Liang Y, Coelho CA, von Rosen T (2022) Hypothesis testing in multivariate normal models with block circular covariance structures. Biom J 64:557–576
Lin L, Higham NJ, Pan J (2014) Covariance structure regularization via entropy loss function. Comput Stat Data Anal 72:315–327
Louden C, Roy A (2010) Classification rules under autoregressive and general circulant covariance. Commun Stat-Theory Methods 39:3294–3315
Magnus J, Neudecker H (1986) Symmetry, 0–1 matrices and jacobians. Economet Theor 2:157–190
Magnus J, Neudecker H (1988) Matrix differential calculus with applications in statistics and econometrics. Wiley, New York
Muirhead RJ (1982) Aspects of multivariate statistical theory. Wiley, New York
Nahtman T, von Rosen D (2008) Shift permutation invariance in linear random factor models. Math Methods Statist 17:173–185
Olkin I (1973) Inference for a normal population when the parameters exhibit some structure. Tech. Rep. No.77, Department of Statistics, Stanford University
Olkin I, Press S (1969) Testing and estimation for a circular stationary model. Ann Math Stat 40:1358–1373
Rao C (2005) Score test: historical review and recent developments. Adv Ranking Sel, Mult Comp Reliab pp 3–20
Roy A, Filipiak K, Klein D (2018) Testing a block exchangeable covariance matrix. Statistics 52(2):393–408
Seely J (1971) Quadratic subspaces and completeness. Ann Math Stat 42(2):710–721
Srivastava MS, Singull M (2012) Profile analysis with random-effects covariance structure. J Jpn Stat Soc 42:145–164
Stein C (1956) Some problems in multivariate analysis, part i. Department of Statistics, Stanford University, Tech. rep
Steinmetz H, Isidor R, Baeuerle N (2012) Testing the circular structure of human values: a meta-analytical structural equation modelling approach. Surv Res Methods 6(1):61–75
Tang P, Nn Xiao, Js Xie (2022) A test on high-dimensional intraclass correlation structure. Chin Q J Math 37(1):10
Wilks SS (1946) Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. Ann Math Stat 17:257–281
Xie J, Sun G (2019) A test for block circular symmetric covariance structure with divergent dimension. ESAIM Prob Stat 23:672–696
Yamada T (2012) High-dimensional edgeworth expansion of lr statistic for testing circular symmetric covariance structure and its error bound. Commun Stat-Theory Methods 41(11):1887–1910
Yi L, Xie J (2018) A high-dimensional likelihood ratio test for circular symmetric covariance structure. Commun Stat-Theory Methods 47(6):1392–1402
Zmyślony R, Žežula I, Kozioł A (2018) Application of jordan algebra for testing hypotheses about structure of mean vector in model with block compound symmetric covariance structure. Electron J Linear Algebra 33:41–52