REVSTAT – Statistical Journal
Volume 16, Number 3, July 2018, 365–403

LIKELIHOOD RATIO TEST FOR THE HYPER-BLOCK MATRIX SPHERICITY COVARIANCE STRUCTURE — CHARACTERIZATION OF THE EXACT DISTRIBUTION AND DEVELOPMENT OF NEAR-EXACT DISTRIBUTIONS FOR THE TEST STATISTIC

Authors: Bárbara R. Correia
– CMA/UNL – Centro de Matemática e Aplicações,
Universidade Nova de Lisboa, Portugal
regadas.correia@gmail.com

Carlos A. Coelho
– Departamento de Matemática,
CMA/UNL – Centro de Matemática e Aplicações,
Universidade Nova de Lisboa, Portugal
cmac@fct.unl.pt

Filipe J. Marques
– Departamento de Matemática,
CMA/UNL – Centro de Matemática e Aplicações,
Universidade Nova de Lisboa, Portugal
fjm@fct.unl.pt

Received: March 2016 Revised: July 2016 Accepted: September 2016

Abstract:

- In this paper the authors introduce the hyper-block matrix sphericity test which is a generalization of both the block-matrix and the block-scalar sphericity tests and as such also of the common sphericity test. This test is a tool of crucial importance to verify elaborate assumptions on covariance matrix structures, namely on meta-analysis and error covariance structures in mixed models and models for longitudinal data. The authors show how by adequately decomposing the null hypothesis of the hyper-block matrix sphericity test it is possible to easily obtain the expression for the likelihood ratio test statistic as well as the expression for its moments. From the factorization of the exact characteristic function of the logarithm of the statistic, induced by the decomposition of the null hypothesis, and by adequately replacing some of the factors with an asymptotic result, it is possible to obtain near-exact distributions that lie very close to the exact distribution. The performance of these near-exact distributions is assessed through the use of a measure of proximity between distributions, based on the corresponding characteristic functions.

Key-Words:

- *equality of matrices test; Generalized Integer Gamma distribution; Generalized Near-Integer Gamma distribution; independence test; near-exact distributions; mixtures of distributions.*

AMS Subject Classification:

- 62H15, 62H10, 62E10, 62E15, 62E20.

1. INTRODUCTION

Likelihood ratio tests (l.r.t.'s) have a large scope of application in different fields of research such as for example engineering, economics, medicine and ecology [27, 20, 5, 21]. However, in most cases, the exact distribution of the l.r.t. statistics has a very complicated expression which makes difficult the practical use of the testing procedure. On the other hand the commonly used asymptotic approximations [2, 26] display lack of precision mainly in extreme situations such as for high number of variables and/or small sample sizes [14, 9] and situations where the parameters of interest and/or nuisance parameters are on the boundary of the parameter space [11]. This is a well known and recognized problem in standard likelihood ratio testing procedures which becomes even more serious when one wants to perform tests for more elaborate covariance structures. These elaborate structures have recently become very important in different statistical techniques for the validation of assumptions required in different models such as in hierarchical or mixed linear univariate and multivariate models.

In this paper the authors introduce the hyper-block matrix (HBM) sphericity test. This test is a useful generalization of the block-matrix and of the block-scalar sphericity tests and is of crucial importance to validate elaborate assumptions on covariance matrix structures, for example on meta-analysis and error covariance structures in mixed models and models for longitudinal data.

We will say that a covariance matrix Σ has a HBM spherical structure if we can write

$$(1.1)\qquad \Sigma = \begin{bmatrix} I_{k_1}\otimes\Delta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & I_{k_m}\otimes\Delta_m \end{bmatrix}, \qquad (\Delta_\ell \text{ unspecified}, \ \ell = 1, ..., m)\,.$$

where $\otimes$ denotes the Kronecker product, and for $\ell = 1, \ldots, m$, I_{k_ℓ} denotes the identity matrix of order k_ℓ and Δ_ℓ is a positive-definite matrix.

The HBM spherical structure may arise in many situations and has as particular cases many interesting and important structures which may be of interest not only as covariance structures in multivariate analysis as well as covariance structures for the error in linear mixed and repeated measures models.

Let us consider a situation in which the same p^* random variables (r.v.'s), $X_1, \ldots, X_{p^*}$, are measured in m "locals", in the ℓ-th of which ($\ell = 1, \ldots, m$) we take k_ℓ measurements, that is, a sample of size k_ℓ, and let us suppose we organize such a meta-sample in a matrix X of dimensions $p^*{\times}n$, with $n = \sum_{\ell=1}^{m} k_\ell$, as in Figure 1.

$$\underset{(p^*\times n)}{X} = \left[\begin{array}{l} \overbrace{X_{11}\ldots X_{1k_1}}^{k_1} \;\ldots\; \overbrace{X_{1,k_1+\cdots+k_{\ell-1}+1}\ldots X_{1,k_1+\cdots+k_\ell}}^{k_\ell} \;\ldots \\ X_{21}\ldots X_{2k_1} \;\ldots\; X_{2,k_1+\cdots+k_{\ell-1}+1}\ldots X_{2,k_1+\cdots+k_\ell} \;\ldots \\ \vdots \qquad\quad \vdots \quad \vdots \qquad\qquad\qquad\qquad \vdots \\ X_{p^*1}\ldots X_{p^*k_1} \;\ldots\; X_{p^*,k_1+\cdots+k_{\ell-1}+1}\ldots X_{p^*,k_1+\cdots+k_\ell} \;\ldots \end{array}\right.$$

$$\left.\begin{array}{l} \ldots\; \overbrace{X_{1,k_1+\cdots+k_{m-1}+1} \;\ldots\; X_{1,k_1+\cdots+k_m}}^{k_m} \\ \ldots\; X_{2,k_1+\cdots+k_{m-1}+1} \;\ldots\; X_{2,k_1+\cdots+k_m} \\ \qquad\quad \vdots \qquad\qquad\qquad\quad \vdots \\ \ldots\; X_{p^*,k_1+\cdots+k_{m-1}+1} \;\ldots\; X_{p^*,k_1+\cdots+k_m} \end{array}\right]$$

Figure 1: Data matrix illustrating a situation of a meta-sample from m "locals", with sample size k_ℓ for the ℓ-th local and $p^*_\ell = p^*$ $(\ell = 1, \ldots, m)$.

Here "locals" is a general designation for example for different locals, factories, companies, hospitals, etc., and if we consider the p^* r.v.'s $X_1, \ldots, X_{p^*}$ organized in the random vector $\underline{X}^* = [X_1, \ldots, X_{p^*}]'$, with $Cov(\underline{X}^*) = \Delta$, then we have

$$Cov(X) = Cov(vec(X)) = \begin{bmatrix} I_{k_1} \otimes \Delta & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & I_{k_m} \otimes \Delta \end{bmatrix}.$$

But, the HBM sphericity setup allows for more general situations as for example those in which we may want to study or model possible differences in strength break in a set of p^* components manufactured by m different companies, or the measurements of p^* variables thought to be possible important indicators of some disorder or disease, measured across m hospitals, or measurements of p^* pollutants in m different locals, or measurements of p^* atmospheric variables and indicators in m different cities, by taking a sample of size k_ℓ in the ℓ-th "local", but that, furthermore, not in every "local", city, hospital or company, it was possible to obtain measurements of all p^* variables, although we still want to consider as many of these in each "local" as possible. Then we may want to consider a meta-sample as the one illustrated in Figure 2, for $p^* = 5$.

In this case, since in different "locals" we may have different subsets of the p^* variables being analyzed, we may end-up with a covariance setup for the matrix X as the one in (1.1), with different covariance matrices for each "local".

Once we assume the HBM spherical structure for the covariance structure in our model, we may then be interested in testing if that is indeed a plausible model for our covariances. The issues are thus: (i) how can we carry out a test for such an elaborate structure, and (ii) in case we find a way of doing so, how will we then be able to obtain p-values and/or quantiles for our test statistic, since this may have a quite elaborate exact distribution.

$$
\begin{array}{rl}
 & \begin{array}{cccc} \overbrace{\hphantom{X_{11}\dots X_{1k_1}}}^{k_1} & \hphantom{\dots} & \overbrace{\hphantom{X_{1,k_1+\cdots+k_{\ell-1}+1}\cdots X_{1,k_1+\cdots+k_\ell}}}^{k_\ell} & \hphantom{\dots} \end{array} \\
\underset{(p^*\times n)}{X} = & \left[\begin{array}{cccc}
X_{11}\dots X_{1k_1} & \dots & X_{1,k_1+\cdots+k_{\ell-1}+1}\cdots X_{1,k_1+\cdots+k_\ell} & \dots \\
X_{21}\dots X_{2k_1} & \dots & \text{———} \quad \text{—} \quad \text{———} & \dots \\
X_{31}\dots X_{3k_1} & \dots & X_{3,k_1+\cdots+k_{\ell-1}+1}\cdots X_{3,k_1+\cdots+k_\ell} & \dots \\
\text{—}\ \text{—}\ \ \text{—} & \dots & X_{4,k_1+\cdots+k_{\ell-1}+1}\cdots X_{4,k_1+\cdots+k_\ell} & \dots \\
X_{51}\dots X_{5k_1} & \dots & \text{———} \quad \text{—} \quad \text{———} & \dots
\end{array}\right.
\end{array}
$$

$$
\begin{array}{rl}
 & \begin{array}{cc} \hphantom{\dots} & \overbrace{\hphantom{X_{1,k_1+\cdots+k_{m-1}+1} \dots X_{1,k_1+\cdots+k_m}}}^{k_m} \end{array} \\
 & \left.\begin{array}{cccc}
\dots & X_{1,k_1+\cdots+k_{m-1}+1} & \dots & X_{1,k_1+\cdots+k_m} \\
\dots & X_{2,k_1+\cdots+k_{m-1}+1} & \dots & X_{2,k_1+\cdots+k_m} \\
\dots & \text{———} & \text{—} & \text{———} \\
\dots & X_{4,k_1+\cdots+k_{m-1}+1} & \dots & X_{4,k_1+\cdots+k_m} \\
\dots & X_{5,k_1+\cdots+k_{m-1}+1} & \dots & X_{5,k_1+\cdots+k_m}
\end{array}\right]
\end{array}
$$

Figure 2: Data matrix illustrating a situation of a meta-sample from m "locals", with sample size k_ℓ for the ℓ-th local ($\ell = 1, \dots, m$), $p_1^* = 4$, $p_\ell^* = 3$ and $p_m^* = 4$, with different covariance matrices Δ_ℓ ($\ell = 1, \dots, m$).

These are indeed the issues we propose to address in this paper, namely showing how one can quite easily build the l.r.t. statistic for the test of the HBM Spherical structure and how we can then obtain the expression for the moments of the statistic and even for the characteristic function (c.f.) of its logarithm, from which factorization we will then be able to obtain very sharp approximations for the exact distribution of the statistic.

The HBM sphericity test is thus a test where the null hypothesis is written as

$$
(1.2) \qquad H_0 : \Sigma = \begin{bmatrix} I_{k_1} \otimes \Delta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & I_{k_m} \otimes \Delta_m \end{bmatrix}, \qquad (\Delta_\ell \text{ unspecified}, \ \ell = 1, ..., m)
$$

where Σ is the covariance matrix of the random vector $\underline{X}$ and the matrices Δ_ℓ are $p_\ell^* \times p_\ell^*$, $(\ell = 1, ..., m)$, with $p_\ell = k_\ell \times p_\ell^*$ and $p = \sum_{\ell=1}^m p_\ell$.

This test is a generalization of the standard sphericity test and it has as particular cases a number of interesting and important tests:

(i) the block-matrix sphericity (BM-Sph) test, for $m = 1$ [4, 3, 15],

(ii) the block-scalar sphericity (BS-Sph) test, for $p^*_\ell = 1, (\ell = 1, ..., m)$ [19, 18, 13],

(iii) the block independence (BI) test, for $k_\ell = 1, (\ell = 1, ..., m)$ [24, 25], [1, Chap. 9], [17, Sec. 11.2], [6, 7],

(iv) the common independence (Ind) test, for $p^*_\ell = k_\ell = 1, (\ell = 1, ..., m)$ [22, Sec. 7.4.3], [9, Secs. 1,2], and

(v) the sphericity (Sph) test, for $m = p^*_1 = 1$ [16], [1, Sec. 10.7], [17, Sec. 8.3], [9].

These particular cases as well as their relations may be analyzed in Figure 3.

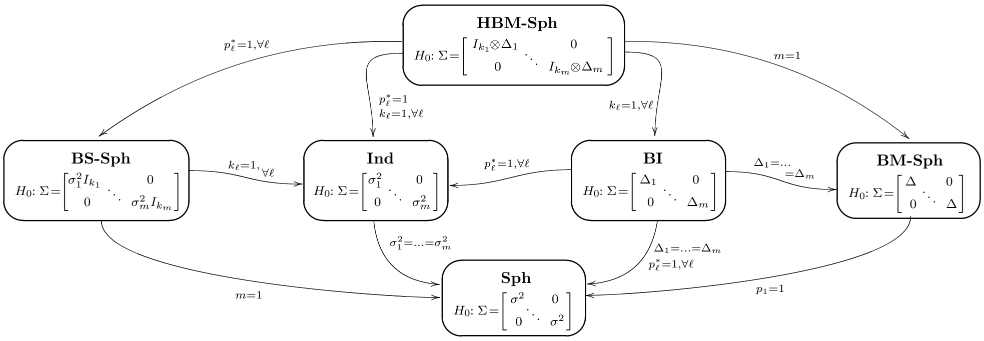

Figure 3: Particular cases of the HBM (Hyper-Block Matrix) test and their inter-relations:

HBM-Sph: Hyper-block matrix sphericity test;
BM-Sph: Block-matrix sphericity test;
BS-Sph: Block-scalar sphericity test;
Sph: Sphericity test; **Ind**: Independence test; **BI**: Block independence test.

The exact distribution of the HBM sphericity test statistic is almost intractable in practical terms, thus our propose is to develop near-exact distributions for the test statistic and its logarithm, based on an adequate factorization of the c.f. of the logarithm of the test statistic.

In Section 2 we will show how we may decompose the overall null hypothesis in (1.2) into a set of three conditionally independent hypotheses and then how from this decomposition (see [8]) we may derive expressions for the l.r.t. statistic and its h-th moment. In Section 3 we will show how we may easily obtain the expression for the c.f. of the logarithm of the test statistic and how we may use the decomposition of the null hypothesis in Section 2, to induce an adequate factorization of this c.f. in two factors, one that is the c.f. a Generalized Integer Gamma (GIG) distribution [6], and the other the c.f. of a sum of independent r.v.'s whose exponentials have Beta distributions. Then, in Section 4 we will use this factorization to build very sharp near-exact distributions both for the test statistic and its logarithm.

Near-exact distributions are asymptotic distributions built using a different approach. Usually working from an adequate factorization of the c.f. of the logarithm of the l.r.t. statistic, we leave unchanged the set of factors that correspond to a manageable distribution and approximate asymptotically the remaining set of factors, in such a way that the resulting c.f., which we will call a near-exact c.f., corresponds to a known manageable distribution, from which p-values and quantiles may be easily computed. These near-exact distributions lie much closer to the exact distribution than any common asymptotic distribution and, when correctly built for statistics used in Multivariate Analysis, will show a marked asymptotic behavior not only for increasing sample sizes but also for increasing number of variables involved.

In Section 5 we will use a measure of proximity between the exact distribution and the near-exact distributions, based on the corresponding characteristic functions in order to assess the quality and the asymptotic properties of the near-exact distributions developed.

Section 6 is dedicated to power studies, where the very good behavior of the test is revealed, through studies based on 1 000 000 pseudo-random samples and carried out on several scenarios of violation of the null hypothesis of HBM sphericity.

2. THE TEST STATISTICS AND ITS MOMENTS

In general terms, a null hypothesis H_0 may be decomposed into a sequence of three conditionally independent null hypotheses, if H_0 admits the decomposition

$$H_0 \equiv H_{03|1,2} \circ H_{02|1} \circ H_{01}$$

where '$\circ$' is to be read as 'after', as long as

$$\Omega_{H_0} \equiv \Omega_{H_{03|1,2}} \subset \Omega_{H_{13|1,2}} \equiv \Omega_{H_{02|1}} \subset \Omega_{H_{12|1}} \equiv \Omega_{H_{01}} \subset \Omega_{H_{11}} \equiv \Omega_{H_1}$$

where Ω_{H_0} is the parameter space under H_0 and Ω_{H_1} the union of the parameter spaces under H_0 and H_1, and where H_{1*} represents the alternative hypothesis to H_{0*} (where '$*$' is used as a wildcard).

The null hypothesis

$$H_{01}:\ \Sigma = bdiag\,(\Sigma_{\ell\ell}; \ell = 1, ..., m) \tag{2.1}$$

corresponds to the test of independence of m groups of variables, the ℓ-th group having $p_\ell = p^*_\ell \times k_\ell$ variables ($\ell = 1, ..., m$).

If we consider that the random vector $\underline{X}$ has a p-variate Normal distribution with expected value vector $\underline{\mu}$ and covariance matrix Σ, that is, if we consider the vector $\underline{X} \sim N_p\left(\underline{\mu}, \Sigma\right)$ and suppose that we have a sample of size $N\,(>p)$ from $\underline{X}$, then the l.r.t. statistic used to test H_{01} and its h-th moment are respectively given by (see secs. 9.2 and 9.3.2 in [1])

$$\Lambda_1 = \frac{|A|^{\frac{N}{2}}}{\prod_{\ell=1}^{m} |A_{\ell\ell}|^{\frac{N}{2}}} \tag{2.2}$$

and

$$E\left[(\Lambda_1)^h\right] = \prod_{\ell=1}^{m-1}\prod_{k=1}^{p_\ell} \frac{\Gamma\left(\frac{N-q_\ell-k}{2}+\frac{N}{2}h\right)}{\Gamma\left(\frac{N-q_\ell-k}{2}\right)} \frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2}+\frac{N}{2}h\right)}\,, \qquad \left(h > \frac{p-p_m}{N}-1\right) \tag{2.3}$$

where the matrix A is the maximum likelihood estimator (m.l.e.) of Σ, $A_{\ell\ell}$ its ℓ-th diagonal block of order p_ℓ $(\ell=1,...,m)$ and $q_\ell = p_{\ell+1}+...+p_m$.

The null hypothesis

$$H_{02|1} = \bigwedge_{\ell=1}^{m} H_{02|1}^{\ell} \tag{2.4}$$

where for $\ell = 1,...,m$

$$\begin{aligned} H_{02|1}^{\ell} :\ & \Sigma_{\ell\ell} = bdiag\left(\Sigma_{vv}^{\ell}, v=1,...,k_\ell\right) \\ & \text{assuming } \Sigma = bdiag\left(\Sigma_{\ell\ell}, \ell=1,...,m\right) \\ & \text{that is, assuming } H_{01} \end{aligned} \tag{2.5}$$

is the null hypothesis of a test of independence of k_ℓ groups of variables, with p_ℓ^* variables each. The l.r.t. statistic to test $H_{02|1}^{\ell}$ in (2.5) and its h-th moment are respectively given by (see Secs. 9.2 and 9.3.2 in [1])

$$\Lambda_{2|1}^{\ell} = \frac{|A_{\ell\ell}|^{\frac{N}{2}}}{\prod_{v=1}^{k_\ell} \left|A_{\ell\ell}^{v}\right|^{\frac{N}{2}}}$$

and

$$E\left[\left(\Lambda_{2|1}^{\ell}\right)^h\right] = \prod_{v=1}^{k_\ell-1}\prod_{k=1}^{p_\ell^*} \frac{\Gamma\left(\frac{N-q_v^\ell-k}{2}+\frac{N}{2}h\right)}{\Gamma\left(\frac{N-q_v^\ell-k}{2}\right)} \frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2}+\frac{N}{2}h\right)}\,, \qquad \left(h > \frac{p_\ell}{N}-1\right)$$

where the matrix $A_{\ell\ell}$ is the maximum likelihood estimator of $\Sigma_{\ell\ell}$, $A_{\ell\ell}^{v}$ its v-th $(v=1,...,k_\ell)$ diagonal block of order p_ℓ^* and $q_v^\ell = (k_\ell - v)\,p_\ell^*$, $(v=1,...,k_\ell)$.

The l.r.t. statistic to test the null hypothesis in (2.4) is thus

$$\Lambda_{2|1} = \prod_{\ell=1}^{m}\Lambda_{2|1}^{\ell} = \prod_{\ell=1}^{m} \frac{|A_{\ell\ell}|^{\frac{N}{2}}}{\prod_{v=1}^{k_\ell}\left|A_{\ell\ell}^{v}\right|^{\frac{N}{2}}} \tag{2.6}$$

and, given the fact that the statistics $\Lambda_{2|1}^{\ell}$, $\ell = 1, \ldots, m$, form a set of m independent statistics, since under H_{01} in (2.1) the m matrices $A_{\ell\ell}$ are independent and each statistic $\Lambda_{2|1}^{\ell}$ is built only from $A_{\ell\ell}$, the h-th moment of $\Lambda_{2|1}$ is

$$
\begin{aligned}
E\left[\left(\Lambda_{2|1}\right)^h\right] &= \prod_{\ell=1}^{m} E\left[\left(\Lambda_{2|1}^{\ell}\right)^h\right] \\
&= \prod_{\ell=1}^{m}\prod_{v=1}^{k_\ell-1}\prod_{k=1}^{p_\ell^*} \frac{\Gamma\left(\frac{N-q_v^\ell-k}{2}+\frac{N}{2}h\right)}{\Gamma\left(\frac{N-q_v^\ell-k}{2}\right)} \frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2}+\frac{N}{2}h\right)}\,, \\
&\qquad \left(h > \max\left\{\tfrac{p_\ell}{N}-1, \ell=1,...,m\right\}\right).
\end{aligned} \tag{2.7}
$$

Finally, the null hypothesis $H_{03|1,2}$ may be written as

$$
H_{03|1,2} = \bigwedge_{\ell=1}^{m} H_{03|1,2}^{\ell} \tag{2.8}
$$

where, for $\ell = 1, ..., m$,

$$
\begin{aligned}
H_{03|1,2}^{\ell}:\ & \Sigma_{11}^{\ell} = \cdots = \Sigma_{k_\ell k_\ell}^{\ell} = \Delta_\ell \ \ (\Delta_\ell \text{ unspecified}) \\
& \text{assuming } \Sigma = bdiag\left(\Sigma_{\ell\ell} = \left(bdiag\left(\Sigma_{vv}^{\ell}, v=1,...,k_\ell\right)\right)\right) \\
& \text{that is, assuming } H_{02|1} \text{ and } H_{01}
\end{aligned} \tag{2.9}
$$

is the null hypothesis corresponding to the test of equality of k_ℓ covariance matrices each with dimensions $p_\ell^* \times p_\ell^*$.

Since under $H_{02|1}$, for each $\ell = 1, \ldots, m$, the k_ℓ matrices $A_{\ell\ell}^{v}$ $(v = 1, \ldots, k_\ell)$ are independent, The l.r.t. statistic to test each null hypothesis $H_{03|1,2}^{\ell}$ in (2.9) and its h-th moment are respectively, (see Secs. 10.2 and 10.4.2 in [1])

$$
\Lambda_{3|1,2}^{\ell} = \frac{\prod_{v=1}^{k_\ell} |A_{\ell\ell}^{v}|^{\frac{N}{2}}}{\left|A_\ell^*\right|^{k_\ell\frac{N}{2}}}\, k_\ell^{\frac{Np_\ell}{2}}
$$

and

$$
E\left[\left(\Lambda_{3|1,2}^{\ell}\right)^h\right] = \prod_{k=1}^{p_\ell^*}\prod_{v=1}^{k_\ell} \frac{\Gamma\left(\frac{N-1}{2}-\frac{k-1}{2k_\ell}+\frac{v-1}{k_\ell}\right)}{\Gamma\left(\frac{N-1}{2}-\frac{k-1}{2k_\ell}+\frac{v-1}{k_\ell}+\frac{N}{2}h\right)} \frac{\Gamma\left(\frac{N-k}{2}+\frac{N}{2}h\right)}{\Gamma\left(\frac{N-k}{2}\right)}
$$

where the matrix $A_{\ell\ell}$ is the maximum likelihood estimator of $\Sigma_{\ell\ell}$, $A_{\ell\ell}^{v}$ its v-th diagonal block of order p_ℓ^* $(v = 1, ..., k_\ell)$ and $A_\ell^* = A_{\ell\ell}^{1} + \cdots + A_{\ell\ell}^{k_\ell}$.

The l.r.t. statistic to test (2.8) is thus

$$
\Lambda_{3|1,2} = \prod_{\ell=1}^{m} \Lambda_{3|1,2}^{\ell} = \prod_{\ell=1}^{m} \frac{\prod_{v=1}^{k_\ell} |A_{\ell\ell}^{v}|^{\frac{N}{2}}}{\left|A_\ell^*\right|^{k_\ell\frac{N}{2}}}\, k_\ell^{\frac{Np_\ell}{2}} \tag{2.10}
$$

and, since under H_{01} in (2.1), the m statistics $\Lambda^{\ell}_{3|1,2}$ are independent, given that each statistic $\Lambda^{\ell}_{3|1,2}$ is built only from $A_{\ell\ell}$ and under H_{01} the m matrices $A_{\ell\ell}$ $(\ell = 1, \ldots, m)$ are independent, the h-th moment of $\Lambda_{3|1,2}$ is given by

$$\begin{aligned} E\left[\left(\Lambda_{3|1,2}\right)^h\right] &= \prod_{\ell=1}^{m} E\left[\left(\Lambda^{\ell}_{3|1,2}\right)^h\right] \\ &= \prod_{\ell=1}^{m}\prod_{k=1}^{p^*_\ell}\prod_{v=1}^{k_\ell} \frac{\Gamma\left(\frac{N-1}{2}-\frac{k-1}{2k_\ell}+\frac{v-1}{k_\ell}\right)\Gamma\left(\frac{N-k}{2}+\frac{N}{2}h\right)}{\Gamma\left(\frac{N-1}{2}-\frac{k-1}{2k_\ell}+\frac{v-1}{k_\ell}+\frac{N}{2}h\right)\Gamma\left(\frac{N-k}{2}\right)}, \\ &\qquad\qquad \left(h > \max\left\{\frac{p^*_\ell}{N}-1, \ell = 1, ..., m\right\}\right). \end{aligned}$$

From Lemma 10.3.1 in [1], the l.r.t. statistic to test (1.2) is thus the product of the l.r.t. statistics used to test the null hypotheses in (2.1), (2.4) and (2.8), that is, the product of the l.r.t. statistics in (2.2), (2.6) and (2.10), thus with

$$\begin{aligned} \Lambda &= \Lambda_1 \times \Lambda_{2|1} \times \Lambda_{3|1,2} \\ &= \frac{|A|^{\frac{N}{2}}}{\prod_{\ell=1}^{m}|A_{\ell\ell}|^{\frac{N}{2}}} \times \left\{\prod_{\ell=1}^{m}\frac{|A_{\ell\ell}|^{\frac{N}{2}}}{\prod_{v=1}^{k_\ell}|A^v_{\ell\ell}|^{\frac{N}{2}}}\right\} \times \left\{\prod_{\ell=1}^{m}\frac{\prod_{v=1}^{k_\ell}|A^v_{\ell\ell}|^{\frac{N}{2}}}{|A^*_\ell|^{k_\ell\frac{N}{2}}}\,k_\ell^{\frac{Np_\ell}{2}}\right\} \qquad (2.11)\\ &= \left\{\prod_{\ell=1}^{m}k_\ell^{\frac{Np_\ell}{2}}\right\} \times \frac{|A|^{\frac{N}{2}}}{\prod_{\ell=1}^{m}|A^*_\ell|^{k_\ell\frac{N}{2}}} \end{aligned}$$

where the matrix A is the m.l.e. of Σ, $A_{\ell\ell}$ is the ℓ-th diagonal block of order p_ℓ of A $(\ell = 1, ..., m)$, with $p = \sum_{\ell=1}^{m} p_\ell$ and $A^*_\ell = A^1_{\ell\ell} + \cdots + A^{k_\ell}_{\ell\ell}$, where $A^v_{\ell\ell}$ is the v-th $(v = 1, ..., k_\ell)$ diagonal block of order p^*_ℓ of $A_{\ell\ell}$ and $p_\ell = k_\ell \times p^*_\ell$. We should note that the expression in (2.11) is identical to the one we obtain when we use the usual method of derivation of the l.r.t. statistic, through its definition (see Appendix A).

The hypotheses $H_{03|1,2}$, $H_{02|1}$ and H_{01} are independent in the sense that under the overall null hypothesis H_0 it is possible to prove that the l.r.t. statistics used to test these hypotheses are independent. Indeed, by Lemma 10.4.1 in [1] or Theorem 5 in [10], the l.r.t. statistic Λ_1 in (2.2) is independent of the m matrices $A_{\ell\ell}$, $(\ell = 1, ..., m)$, so that since $\Lambda_{2|1}$ and $\Lambda_{3|1,2}$ are built only from the m matrices $A_{\ell\ell}$ $(\ell = 1, ..., m)$, these l.r.t. statistics are independent of Λ_1. But then, since we may use the same two results to argue that each statistic $\Lambda^{\ell}_{2|1}$ is independent of the k_ℓ matrices $A^v_{\ell\ell}$ $(v = 1, \ldots, k_\ell)$, the statistic $\Lambda_{2|1}$ is independent of all $\sum_{\ell=1}^{m} k_\ell$ matrices $A^v_{\ell\ell}$ $(\ell = 1, \ldots, k_\ell; v = 1, \ldots, m)$ and as such independent of $\Lambda_{3|1,2}$ which is built only on these matrices.

As such, we may obtain the expression of the h-th moment of Λ from the expressions for the h-th moment of each of the statistics $\Lambda_{3|1,2}$, $\Lambda_{2|1}$ and Λ_1 writing it as

$$\begin{aligned}
E\left[(\Lambda)^h\right] &= E\left[(\Lambda_1)^h\right] \times E\left[\left(\Lambda_{2|1}\right)^h\right] \times E\left[\left(\Lambda_{3|1,2}\right)^h\right] \\
&= \prod_{\ell=1}^{m} \prod_{k=1}^{p_\ell^*} \prod_{v=1}^{k_\ell-1} \frac{\Gamma\left(\frac{N-1}{2} - \frac{k-1}{2k_\ell} + \frac{v-1}{k_\ell}\right)}{\Gamma\left(\frac{N-1}{2} - \frac{k-1}{2k_\ell} + \frac{v-1}{k_\ell} + \frac{N}{2}h\right)} \frac{\Gamma\left(\frac{N-q_v^\ell-k}{2} + \frac{N}{2}h\right)}{\Gamma\left(\frac{N-q_v^\ell-k}{2}\right)} \\
&\quad \times \prod_{\ell=1}^{m} \prod_{k=1}^{p_\ell^*} \frac{\Gamma\left(\frac{N-1}{2} - \frac{k-1}{2k_\ell} + \frac{k_\ell-1}{k_\ell}\right)}{\Gamma\left(\frac{N-1}{2} - \frac{k-1}{2k_\ell} + \frac{k_\ell-1}{k_\ell} + \frac{N}{2}h\right)} \frac{\Gamma\left(\frac{N-k}{2} + \frac{N}{2}h\right)}{\Gamma\left(\frac{N-k}{2}\right)} \\
&\quad \times \prod_{\ell=1}^{m-1} \prod_{k=1}^{p_\ell} \frac{\Gamma\left(\frac{N-q_\ell-k}{2} + \frac{N}{2}h\right)}{\Gamma\left(\frac{N-q_\ell-k}{2}\right)} \frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2} + \frac{N}{2}h\right)},
\end{aligned}$$

for $h > \max\{(p-p_m)/N-1, p_\ell/N-1 (\ell=1,...,m)\}$ and where $p = \sum_{\ell=1}^m p_\ell$, $q_\ell = p_{\ell+1} + ... + p_m$, $p_\ell = k_\ell \times p_\ell^*$ and $q_v^\ell = (k_\ell - v)\, p_\ell^*$ $(\ell = 1, ..., m; v = 1, ..., k_\ell)$.

Actually we may note that the two first null hypotheses may be condensed into a single null hypothesis to test the independence of $q^{**} = \sum_{\ell=1}^m k_\ell$ groups of variables, the ν-th group with p_ν^{**} variables for

$$p^{**} = \Big[\underbrace{p_1^*, \ldots, p_1^*}_{k_1}, \underbrace{p_2^*, \ldots, p_2^*}_{k_2}, \;\ldots\;, \underbrace{p_\ell^*, \ldots, p_\ell^*}_{k_\ell}, \;\ldots\;, \underbrace{p_m^*, \ldots, p_m^*}_{k_m}\Big], \tag{2.12}$$

with $\nu = 1, \ldots, \sum_{\ell=1}^m k_\ell$, that is, with

$$p_\nu^{**} = p_\ell^*\,, \qquad \text{for} \qquad 1 + \sum_{i=1}^{\ell-1} k_i \le \nu \le \sum_{i=1}^{\ell} k_i\,.$$

This null hypothesis may be written as

$$\begin{aligned}
H_{0,12} : \Sigma = bdiag\Big(&\underbrace{\Delta_1, \ldots, \Delta_{k_1}}_{\text{of order } p_1^*}, \;\ldots\;, \underbrace{\Delta_{k_1+\cdots+k_{\ell-1}+1}, \ldots, \Delta_{k_1+\cdots+k_\ell}}_{\text{of order } p_\ell^*}, \;\ldots \\
&\ldots\;, \underbrace{\Delta_{k_1+\cdots+k_{m-1}+1}, \ldots, \Delta_{k_1+\cdots+k_m}}_{\text{of order } p_m^*}\Big).
\end{aligned}$$

The l.r.t. statistic to test $H_{0,12}$ is given by

$$\Lambda_{1,2} = \frac{|A|^{\frac{N}{2}}}{\prod_{\nu=1}^{q^{**}} |A_{\nu\nu}^*|^{\frac{N}{2}}} \tag{2.13}$$

where $A^*_{\nu\nu}$ is the ν-th diagonal block of order p^{**}_ν ($\nu = 1, ..., q^{**}$), and the expression of its h-th moment is given by (see secs. 9.2 and 9.3.2 in [1])

$$E\left[(\Lambda_{1,2})^h\right] = \prod_{\nu=1}^{q^{**}-1} \prod_{k=1}^{p^{**}_\nu} \frac{\Gamma\left(\frac{N-q^*_\nu-k}{2} + \frac{N}{2}h\right)}{\Gamma\left(\frac{N-q^*_\nu-k}{2}\right)} \frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2} + \frac{N}{2}h\right)}$$

for

$$q^*_\nu = p^{**}_{\nu+1} + \cdots + p^{**}_{q^{**}} \quad \text{where} \quad q^{**} = \sum_{\ell=1}^{m} k_\ell \,.$$

Note that the l.r.t. statistic in (2.13) may be also given by the product of the l.r.t.'s in (2.2) and (2.6), used to test the null hypotheses in (2.1) and (2.4) respectively, and the expression of its h-th moment may also be given by the product of the expressions of the h-th moments in (2.3) and (2.7) of the l.r.t.'s in (2.2) and (2.6) respectively.

Finally, the expression of the h-th moment of Λ may be re-written as

$$\begin{aligned} E\left[(\Lambda)^h\right] &= E\left[(\Lambda_{1,2})^h\right] \times E\left[\left(\Lambda_{3|1,2}\right)^h\right] \\ &= \prod_{\nu=1}^{q^{**}-1} \prod_{k=1}^{p^{**}_\nu} \frac{\Gamma\left(\frac{N-q^*_\nu-k}{2} + \frac{N}{2}h\right)}{\Gamma\left(\frac{N-q^*_\nu-k}{2}\right)} \frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2} + \frac{N}{2}h\right)} \\ &\quad \times \prod_{\ell=1}^{m} \prod_{k=1}^{p^*_\ell} \prod_{v=1}^{k_\ell} \frac{\Gamma\left(\frac{N-1}{2} - \frac{k-1}{2k_\ell} + \frac{v-1}{k_\ell}\right)}{\Gamma\left(\frac{N-1}{2} - \frac{k-1}{2k_\ell} + \frac{v-1}{k_\ell} + \frac{N}{2}h\right)} \frac{\Gamma\left(\frac{N-k}{2} + \frac{N}{2}h\right)}{\Gamma\left(\frac{N-k}{2}\right)} \,. \end{aligned} \tag{2.14}$$

The factorization of the c.f. of $W = -\log \Lambda$ developed in the next section will have as a starting base this last expression.

3. THE CHARACTERISTIC FUNCTION OF $W = -\log(\Lambda)$

Since in (2.14) the Gamma functions remain valid for any strictly complex h, if we take $W_{1,2} = -\log \Lambda_{1,2}$ and $W_3 = -\log \Lambda_{3|1,2}$, we may write the c.f. of $W = -\log \Lambda$ as

$$
\begin{aligned}
\Phi_W(t) &= E\left(e^{-it\log\Lambda}\right) = E\left[\Lambda^{-it}\right] = E\left[\Lambda_{1,2}^{-it}\right]\, E\left[\Lambda_{3|1,2}^{-it}\right] \\
&= \Phi_{W_{1,2}}(t)\,\Phi_{W_3}(t) \\
&= \underbrace{\prod_{\nu=1}^{q^{**}-1}\prod_{k=1}^{p_\nu^{**}} \frac{\Gamma\left(\frac{N-q_\nu^*-k}{2}-\frac{N}{2}it\right)}{\Gamma\left(\frac{N-q_\nu^*-k}{2}\right)}\,\frac{\Gamma\left(\frac{N-k}{2}\right)}{\Gamma\left(\frac{N-k}{2}-\frac{N}{2}it\right)}}_{\Phi_{W_{1,2}}(t)} \\
&\quad\times \underbrace{\prod_{\ell=1}^{m}\prod_{k=1}^{p_\ell^*}\prod_{v=1}^{k_\ell} \frac{\Gamma\left(\frac{N-1}{2}-\frac{k-1}{2k_\ell}+\frac{v-1}{k_\ell}\right)}{\Gamma\left(\frac{N-1}{2}-\frac{k-1}{2k_\ell}+\frac{v-1}{k_\ell}-\frac{N}{2}it\right)}\,\frac{\Gamma\left(\frac{N-k}{2}-\frac{N}{2}it\right)}{\Gamma\left(\frac{N-k}{2}\right)}}_{\Phi_{W_3}(t)}\,,
\end{aligned}
\tag{3.1}
$$

where $t \in \mathbb{R}$, $i = \sqrt{-1}$, $q_\nu^* = p_{\nu+1}^{**} + \cdots + p_{q^{**}}^{**}$, $q^{**} = \sum_{\ell=1}^m k_\ell$, and p_ν^{**} are defined in (2.12). From this expression we may state that

$$
\Lambda \stackrel{d}{\equiv} \left\{\prod_{\nu=1}^{q^{**}-1}\prod_{k=1}^{p_\nu^{**}} (Y_{\nu k})^{N/2}\right\} \times \left\{\prod_{\ell=1}^{m}\prod_{k=1}^{p_\ell^*}\prod_{v=1}^{k_\ell} (Y_{\ell k v})^{N/2}\right\}
$$

where

$$
Y_{\nu k} \sim Beta\left(\frac{N-q_\nu^*-k}{2}, \frac{q_\nu^*}{2}\right) \quad \text{and} \quad Y_{\ell k v} \sim Beta\left(\frac{N-k}{2}, \frac{v-1}{k_\ell} + \frac{k-1}{2}\,\frac{k_\ell-1}{k_\ell}\right)
$$

are independent r.v.'s.

In order to be able to build sharp near-exact distributions for W and Λ, we need to further factorize $\Phi_{W_{1,2}}(t)$ and $\Phi_{W_3}(t)$, by writing each one of these c.f.'s as the product of two factors, one that is the c.f. of a Generalized Integer Gamma (GIG) distribution and the other the c.f. of a sum of independent r.v.'s whose exponentials have Beta distributions.

3.1. The factorization of the characteristic function of $W_{1,2} = -\log\Lambda_{1,2}$

The results in [7, 14] may be used to show that $\Phi_{W_{1,2}}(t)$ in (3.1) may be written as

$$
\Phi_{W_{1,2}}(t) = \Phi_{W_{1,2,a}}(t) \times \Phi_{W_{1,2,b}}(t)\,,
$$

where

$$
\Phi_{W_{1,2,a}}(t) = \prod_{k=3}^{p} \left(\frac{N-k}{N}\right)^{r_{1,k}} \left(\frac{N-k}{N} - it\right)^{-r_{1,k}}
\tag{3.2}
$$

is the c.f. of the sum of $p-2$ independent integer Gamma r.v.'s, that is a Generalized Integer Gamma distribution of depth $p-2$, with integer shape parameters

$r_{1,k}$ given by

$$r_{1,k} = \begin{cases} h_{1,k-2} + (-1)^k\, k^* & k = 3,4 \\ r_{1,k-2} + h_{1,k-2}, & k = 5, ..., p \end{cases} \tag{3.3}$$

with $k^* = \left\lfloor \frac{m^*}{2} \right\rfloor$, where m^* is the number of sets of variables with an odd number of variables, among the q^{**} groups of variables, the ν-th of which with p_ν^{**} variables, and

$$h_{1;k} = (\#\text{ of } p_\nu^{**}\,(\nu = 1, .., q^{**}) \geq k) - 1, \qquad k = 1, ..., p-2$$

and

$$\Phi_{W_{1,2,b}}(t) = \left(\frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N-2}{2}\right)} \frac{\Gamma\left(\frac{N-2}{2} - \frac{N}{2}it\right)}{\Gamma\left(\frac{N-1}{2} - \frac{N}{2}it\right)} \right)^{k^*} \tag{3.4}$$

is the c.f. of the sum of k^* independent r.v.'s with Logbeta distributions multiplied by $\frac{N}{2}$. We should note that when $k^* = 0$, $\Phi_{W_{1,2,b}}(t) = 1$.

3.2. The factorization of the characteristic function of $W_3 = -\log \Lambda_{3|1,2}$

Based on the results in [14] we may re-write the c.f. of W_3, $\Phi_{W_3}(t)$ in (3.1), as

$$\Phi_{W_3}(t) = \Phi_{W_{3,a}}(t) \times \Phi_{W_{3,b}}(t)$$

where

$$\Phi_{W_{3,a}}(t) = \prod_{\ell=1}^{m} \prod_{k=2}^{p_\ell^*} \left(\frac{N-k}{N}\right)^{r_{3,k-1}^\ell} \left(\frac{N-k}{N} - it\right)^{-r_{3,k-1}^\ell} \tag{3.5}$$

is the c.f. of the sum of $\sum\limits_{\ell=1}^{m} p_\ell^* - m$ independent Gamma r.v.'s, that is, a Generalized Integer Gamma distribution of depth $\sum\limits_{\ell=1}^{m} p_\ell^* - m$ with integer shape parameters $r_{3;k}^\ell$ given by (B.1) in Appendix B, and

$$\begin{aligned} \Phi_{W_{3,2}}(t) = \prod_{\ell=1}^{m} \Bigg\{ & \prod_{k=1}^{\lfloor p_\ell^*/2 \rfloor} \prod_{v=1}^{k_\ell} \frac{\Gamma\left(a_k^\ell + b_{kv}^\ell\right)}{\Gamma\left(a_k^\ell + b_{kv}^{\ell*}\right)} \frac{\Gamma\left(a_k^\ell + b_{vk}^{\ell*} - Nit\right)}{\Gamma\left(a_k^\ell + b_{vk}^\ell - Nit\right)} \\ & \times \left(\prod_{v=1}^{k_\ell} \frac{\Gamma\left(a_{p_\ell^*}^\ell + b_{p_\ell^* v}^\ell\right)}{\Gamma\left(a_{p_\ell^*}^\ell + b_{p_\ell^* v}^{\ell*}\right)} \frac{\Gamma\left(a_{p_\ell^*}^\ell + b_{p_\ell^* v}^{\ell*} - \frac{N}{2}it\right)}{\Gamma\left(a_{p_\ell^*}^\ell + b_{p_\ell^* v}^\ell - \frac{N}{2}it\right)} \right)^{\left(p_\ell^* \bmod 2\right)} \Bigg\} \end{aligned} \tag{3.6}$$

with

$$a_k^\ell = N - 2k,\ b_{kv}^\ell = 2k - 1 + \tfrac{v-2k}{k_\ell},\ b_{vk}^{\ell*} = \left\lfloor b_{vk}^\ell \right\rfloor \tag{3.7}$$

and

$$a^\ell_{p^*_\ell}=\tfrac{N-p^*_\ell}{2}\,,\ b^\ell_{p^*_\ell v}=\tfrac{p_\ell-k_\ell-p^*_\ell+2v-1}{2k_\ell}\,,\ b^{\ell *}_{p^*_\ell v}=\left\lfloor b^\ell_{p^*_\ell v}\right\rfloor\,, \tag{3.8}$$

is the c.f. of the sum of $\sum_{\ell=1}^{m}\left\lfloor \frac{p^*_\ell}{2}\right\rfloor k_\ell$ independent Logbeta r.v.'s multiplied by N and $\sum_{\ell=1}^{m} k_\ell\,(p^*_\ell \bmod 2)$ independent Logbeta r.v.'s multiplied by $\frac{N}{2}$.

As such, the c.f. of W may be written as

$$\Phi_W\,(t)=\Phi_1\,(t)\times\Phi_2\,(t) \tag{3.9}$$

where

$$\Phi_1\,(t)=\Phi_{W_{1,2,a}}\,(t)\times\Phi_{W_{3,a}}\,(t)\,, \tag{3.10}$$

with $\Phi_{W_{1,2,a}}\,(t)$ and $\Phi_{W_{3,a}}\,(t)$ given by (3.2) and (3.5), respectively, and

$$\Phi_2\,(t)=\Phi_{W_{1,2,b}}\,(t)\times\Phi_{W_{3,b}}\,(t) \tag{3.11}$$

with $\Phi_{W_{1,2,b}}\,(t)$ and $\Phi_{W_{3,b}}\,(t)$ in (3.4) and (3.6), respectively.

The c.f. $\Phi_1(t)$ in (3.10) can be seen as the c.f. of a GIG distribution of depth $p-1$, and it may be written as

$$\Phi_1\,(t)=\prod_{k=2}^{p}\left(\frac{N-k}{N}\right)^{r^+_k}\left(\frac{N-k}{N}-it\right)^{-r^+_k} \tag{3.12}$$

where

$$r^+_k=r^*_{1,k}+\sum_{\ell=1}^{m} r^{\ell *}_{3,k}\,, \tag{3.13}$$

with

$$r^*_{1,k}=\begin{cases}0 & k=2\\ r_{1,k} & k=3,\dots,p\end{cases} \qquad \text{and} \qquad r^{\ell *}_{3,k}=\begin{cases}r^\ell_{3,k} & k=2,\dots,p^*_\ell\\ 0 & k=p^*_\ell,\dots,p\end{cases} \tag{3.14}$$

with $r_{1,k}$ given by (3.3) and $r^\ell_{3,k}$ given by (B.1) in Appendix B, while $\Phi_2(t)$ is the c.f. of a sum of $k^*+\sum_{\ell=1}^{m}\left\lfloor \frac{p^*_\ell}{2}\right\rfloor k_\ell$ independent Logbeta r.v.'s multiplied by N and $\sum_{\ell=1}^{m} k_\ell\,(p^*_\ell \bmod 2)$ independent Logbeta r.v.'s multiplied by $\frac{N}{2}$.

From this alternative expression for the c.f. of $W=-\log\Lambda$ given by (3.9), we may see that the exact distribution of Λ in (2.11) may be written as

$$\Lambda \stackrel{d}{\equiv} \left\{\prod_{k=2}^{p} e^{Z_k}\right\}\left\{\prod_{k=1}^{k^*}(Y_{1,k})^{\frac{N}{2}}\right\}\left\{\prod_{\ell=1}^{m}\prod_{k=1}^{\lfloor p^*_\ell/2\rfloor}\prod_{v=1}^{k_\ell}\left(Y^\ell_{3,kv}\right)^{N}\right\}$$
$$\times\left\{\prod_{\ell=1}^{m}\left\{\prod_{v=1}^{k\ell}(Y^\ell_{3,v})^{\frac{N}{2}}\right\}^{p^*_\ell \bmod 2}\right\}$$

where $\stackrel{d}{\equiv}$ means "equivalent in distribution" and all the r.v.'s involved are independent, with

$$
\begin{aligned}
& Z_k \sim \Gamma\left(r_k^+, \frac{N-k}{N}\right), && k=2,\ldots,p \\
& Y_{1,k} \sim Beta\left(\frac{N-2}{2}, \frac{1}{2}\right), && k=1,\ldots,k^* \\
& Y^\ell_{3,kv} \sim Beta\Big(a^\ell_k + b^{\ell*}_{kv}, b^\ell_{kv} - b^{\ell*}_{kv}\Big), && k=1,\ldots,\left\lfloor \tfrac{p^*_\ell}{2} \right\rfloor;\ v=1,\ldots,k_\ell; \\
& && \ell = 1,\ldots,m \\
& Y^\ell_{3,v} \sim Beta\left(a^\ell_{p^*_\ell} + b^{\ell*}_{p^*_\ell v}, b^\ell_{p^*_\ell v} - b^{\ell*}_{p^*_\ell v}\right), &&
\end{aligned}
\tag{3.15}
$$

with r_k^+ given by (3.13) and (3.14), a^ℓ_k, b^ℓ_{kv} and $b^{\ell*}_{kv}$ given by (3.7) and $a^\ell_{p^*_\ell}$, $b^\ell_{p^*_\ell v}$ and $b^{\ell*}_{p^*_\ell v}$ given by (3.8).

This representation will enable us to develop well-fitting near-exact distributions, which bear an extreme closeness to the exact distribution of Λ.

4. NEAR-EXACT DISTRIBUTIONS FOR W AND Λ

To build the near-exact distributions of $W = -\log\Lambda$ and Λ we will leave $\Phi_1(t)$ in (3.9) and (3.12) unchanged and we will replace $\Phi_2(t)$ in (3.9) and (3.11) by a sharp asymptotic approximation in such a way that the resulting c.f. corresponds to a known manageable distribution.

From the results in Section 5 of [23], which show that we may asymptotically replace a $Logbeta(a,b)$ distribution by an infinite mixture of $\Gamma(b+j,a)$ distributions, with $j = 0,1,\ldots,$ using a somewhat heuristic approach, we will replace $\Phi_2(t)$ by

$$
\Phi_2^*(t) = \sum_{j=0}^{m^+} \pi_j \theta^{r+j} (\theta - it)^{-(r+j)}\,,
\tag{4.1}
$$

which is the c.f. of a finite mixture of $\Gamma(r+j,\theta)$ distributions, where

$$
\begin{aligned}
r \;&=\; \tfrac{k^*}{2} + \sum_{\ell=1}^{m} \sum_{k=1}^{\lfloor p^*_\ell/2 \rfloor} \sum_{v=1}^{k_\ell} \tfrac{v-2k}{k_\ell} - \left\lfloor \tfrac{v-2k}{k_\ell} \right\rfloor \\
&\qquad\qquad + \sum_{\ell=1}^{m} \left(\sum_{v=1}^{k_\ell} \tfrac{2v-p^*_\ell-1}{2k_\ell} - \left\lfloor \tfrac{2v-p^*_\ell-1}{2k_\ell} \right\rfloor \right)^{p^*_\ell \bmod 2} \\
&=\; \tfrac{k^*}{2} + \sum_{\ell=1}^{m} \left\lfloor \tfrac{p^*_\ell+1}{2} \right\rfloor \tfrac{k_\ell - 1}{2}
\end{aligned}
\tag{4.2}
$$

is the sum of all the second parameters of the Beta r.v.'s in (3.15) and θ is obtained, together with s_1, s_2 and π^*, as the numerical solution of the system of equations

$$\frac{d^h}{dt^h}\Phi_2(t)\Big|_{t=0} = \frac{d^h}{dt^h}\Big(\pi^*\theta^{s_1}(\theta-it)^{-s_1}+(1-\pi^*)\theta^{s_2}(\theta-it)^{-s_2}\Big)\Big|_{t=0}\,, \qquad h=1,\dots,4\,, \tag{4.3}$$

that is, as the rate parameter of a mixture of two Gamma distributions with a common rate, which matches the first 4 derivatives of $\Phi_2(t)$, at $t=0$, so that $\Phi_1(t)\times\Big(\pi^*\theta^{s_1}(\theta-it)^{-s_1}+(1-\pi^*)\theta^{s_2}(\theta-it)^{-s_2}\Big)$ corresponds to a distribution that matches the first 4 exact moments of W. Then the weights π_j, $j=0,\dots,m^+-1$ are determined in such a way that

$$\frac{d^h}{dt^h}\Phi_2(t)\Big|_{t=0} = \frac{d^h}{dt^h}\Phi_2^*(t)\Big|_{t=0}\,, \qquad h=1,\dots,m^+\,, \tag{4.4}$$

with $\pi_{m^+}=1-\sum_{j=0}^{m^+-1}\pi_j$.

We will thus take as near-exact c.f. of W the c.f.

$$\begin{aligned}\Phi_W^*(t) &= \Phi_1(t)\times\Phi_2^*(t)\\ &= \left\{\prod_{k=2}^{p}\left(\frac{N-k}{N}\right)^{r_k^+}\left(\frac{N-k}{N}-it\right)^{-r_k^+}\right\}\\ &\qquad\qquad\times\left\{\sum_{j=0}^{m^+}\pi_j\,\theta^{r+j}\,(\theta-it)^{-(r+j)}\right\}\\ &= \sum_{j=0}^{m^+}\pi_j\left\{\theta^{r+j}\,(\theta-it)^{-(r+j)}\prod_{k=2}^{p}\left(\frac{N-k}{N}\right)^{r_k^+}\left(\frac{N-k}{N}-it\right)^{-r_k^+}\right\}\end{aligned} \tag{4.5}$$

with r and r_k^+ respectively given by (4.2) and (3.13), which is the c.f. of a mixture of m^++1 GNIG distributions of depth p that matches the first m^+ exact moments of W. This c.f. yields near-exact distributions for W with p.d.f.

$$f_W(w)=\sum_{j=0}^{m^+}\pi_j f^{GNIG}\Big(w\,\Big|\,r_2^+,r_3^+,\dots,r_p^+,r;\tfrac{N-2}{N},\tfrac{N-3}{N},\dots,\tfrac{N-p}{N},\theta;p\Big), \tag{4.6}$$

and c.d.f.

$$F_W(w)=\sum_{j=0}^{m^+}\pi_j F^{GNIG}\Big(w\,\Big|\,r_2^+,r_3^+,\dots,r_p^+,r;\tfrac{N-2}{N},\tfrac{N-3}{N},\dots,\tfrac{N-p}{N},\theta;p\Big), \tag{4.7}$$

for $w>0$, and near-exact distributions for Λ with p.d.f.

$$f_\Lambda(z)=\sum_{j=0}^{m^+}\pi_j f^{GNIG}\Big(-\log z\,\Big|\,r_2^+,r_3^+,\dots,r_p^+,r;\tfrac{N-2}{N},\tfrac{N-3}{N},\dots,\tfrac{N-p}{N},\theta;p\Big)\tfrac{1}{z}, \tag{4.8}$$

and c.d.f.

$$(4.9)\quad F_\Lambda(z)=\sum_{j=0}^{m^+}\pi_j\left(1-F^{GNIG}\left(-\log z\,\middle|\,r_2^+,r_3^+,...,r_p^+,r;\tfrac{N-2}{N},\tfrac{N-3}{N},...,\tfrac{N-p}{N},\theta;p\right)\right),$$

for $0 < z < 1$.

The modules for the GNIG c.d.f. and p.d.f. are available in [12] and on the web-page https://sites.google.com/site/nearexactdistributions/. Using these modules, the computation of the p.d.f.'s and c.d.f.'s of the near-exact distributions becomes easy and very manageable, once the system of equations in (4.4) is linear and as such very simple to solve, as it is also the case with the system of equations in (4.3). The authors make available a set of Mathematica® modules to implement the computation of p.d.f, c.d.f., p-values and quantiles for the near-exact distributions developed in the paper, as well as a module to compute the value of the l.r.t. statistic from a sample, on the web-page https://sites.google.com/site/nearexactdistributions/hyper-block-matrix-sphericity. In Appendix C the authors present a short manual for the use of these modules, along with some examples.

5. NUMERICAL STUDIES

In order to assess the performance of the near-exact distributions obtained in the previous section we will use the measure

$$(5.1)\qquad \Delta^* = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{\Phi_W(t)-\Phi_W^*(t)}{t}\right| dt\,,$$

with

$$\max_w |F_W(w)-F_W^*(w)| = \max_z |F_\Lambda(z)-F_\Lambda^*(z)| \,\le\, \Delta^*\,,$$

where $\Phi_W(t)$ and $\Phi_W^*(t)$ represent respectively the exact and near-exact c.f.'s of W; $F_W(\cdot)$, $F_W^*(\cdot)$, the exact and near-exact c.d.f.'s of W and $F_\Lambda(\cdot)$ and $F_\Lambda^*(\cdot)$ those of Λ. Values for this measure Δ^*, which is therefore an upper bound on the difference between the exact and near-exact c.d.f.'s of both W and Λ for the near-exact distributions developed in the previous section, for different values of p_ℓ^* and k_ℓ, may be analyzed in Tables 1 and 2, with smaller values of the measure indicating even better agreements between the near-exact and the exact distributions.

From Tables 1 and 2 we may see the clear asymptotic behavior of the near-exact distributions not only for increasing sample sizes but also for increasing values of p_ℓ^* and k_ℓ, as well as the very good performance of the near-exact distributions for very small sample sizes, which barely exceed the number of variables in use.

Table 1: Values of the measure Δ^* in (5.1) for different values of m^+ (the number of exact moments matched by the near-exact distributions) and for $m = 4$ and increasing values of k_ℓ, p_ℓ^* and N, for base values of $k_\ell = \{3, 2, 3, 4\}$, $p_\ell^* = \{3, 5, 6, 4\}$ and sample sizes $N = p + 2, 200, 500$, with $p = \sum_{\ell=1}^{m} k_\ell \times p_\ell^*$.

N	m^+					
	0	1	2	4	6	10
$k_\ell + 0,\ p_\ell^* + 0,\ p = 53$						
$p+2$	$7.06{\times}10^{-6}$	$4.36{\times}10^{-8}$	$1.33{\times}10^{-12}$	$2.31{\times}10^{-19}$	$1.30{\times}10^{-26}$	$9.24{\times}10^{-37}$
$p+200$	$7.40{\times}10^{-7}$	$1.04{\times}10^{-8}$	$3.21{\times}10^{-14}$	$4.24{\times}10^{-21}$	$9.44{\times}10^{-28}$	$8.24{\times}10^{-41}$
$p+500$	$1.60{\times}10^{-7}$	$2.35{\times}10^{-9}$	$1.52{\times}10^{-15}$	$1.39{\times}10^{-22}$	$5.96{\times}10^{-30}$	$3.50{\times}10^{-44}$
$k_\ell + 2,\ p_\ell^* + 0,\ p = 89$						
$p+2$	$2.24{\times}10^{-6}$	$7.56{\times}10^{-9}$	$1.67{\times}10^{-13}$	$1.49{\times}10^{-20}$	$1.51{\times}10^{-27}$	$2.76{\times}10^{-41}$
$p+200$	$5.28{\times}10^{-7}$	$4.22{\times}10^{-9}$	$4.99{\times}10^{-14}$	$2.71{\times}10^{-21}$	$1.49{\times}10^{-28}$	$7.42{\times}10^{-43}$
$p+500$	$1.35{\times}10^{-7}$	$1.15{\times}10^{-9}$	$6.45{\times}10^{-15}$	$9.96{\times}10^{-23}$	$1.50{\times}10^{-30}$	$5.65{\times}10^{-46}$
$k_\ell + 5,\ p_\ell^* + 0,\ p = 143$						
$p+2$	$8.66{\times}10^{-7}$	$1.70{\times}10^{-9}$	$2.33{\times}10^{-14}$	$4.31{\times}10^{-22}$	$8.30{\times}10^{-30}$	$4.34{\times}10^{-45}$
$p+200$	$3.75{\times}10^{-7}$	$1.77{\times}10^{-9}$	$2.13{\times}10^{-14}$	$4.19{\times}10^{-22}$	$8.01{\times}10^{-30}$	$3.79{\times}10^{-45}$
$p+500$	$1.16{\times}10^{-7}$	$5.99{\times}10^{-10}$	$3.99{\times}10^{-15}$	$2.71{\times}10^{-23}$	$1.77{\times}10^{-31}$	$9.58{\times}10^{-48}$
$k_\ell + 2,\ p_\ell^* + 2,\ p = 129$						
$p+2$	$1.14{\times}10^{-6}$	$2.50{\times}10^{-9}$	$1.58{\times}10^{-14}$	$3.24{\times}10^{-22}$	$7.86{\times}10^{-30}$	$6.94{\times}10^{-45}$
$p+200$	$4.34{\times}10^{-7}$	$2.30{\times}10^{-9}$	$8.28{\times}10^{-15}$	$1.72{\times}10^{-22}$	$3.85{\times}10^{-30}$	$2.66{\times}10^{-45}$
$p+500$	$1.28{\times}10^{-7}$	$7.39{\times}10^{-10}$	$1.19{\times}10^{-15}$	$8.62{\times}10^{-24}$	$6.44{\times}10^{-32}$	$4.83{\times}10^{-48}$
$k_\ell + 2,\ p_\ell^* + 5,\ p = 189$						
$p+2$	$6.62{\times}10^{-7}$	$9.46{\times}10^{-10}$	$2.22{\times}10^{-15}$	$1.41{\times}10^{-23}$	$1.27{\times}10^{-31}$	$1.88{\times}10^{-47}$
$p+200$	$3.84{\times}10^{-7}$	$1.32{\times}10^{-9}$	$2.92{\times}10^{-15}$	$1.69{\times}10^{-23}$	$2.33{\times}10^{-31}$	$7.10{\times}10^{-47}$
$p+500$	$1.35{\times}10^{-7}$	$5.18{\times}10^{-10}$	$2.47{\times}10^{-16}$	$1.05{\times}10^{-24}$	$6.62{\times}10^{-33}$	$3.62{\times}10^{-49}$
$k_\ell + 5,\ p_\ell^* + 2,\ p = 207$						
$p+2$	$3.75{\times}10^{-7}$	$4.84{\times}10^{-10}$	$1.88{\times}10^{-15}$	$9.07{\times}10^{-24}$	$4.51{\times}10^{-32}$	$1.36{\times}10^{-48}$
$p+200$	$2.38{\times}10^{-7}$	$7.36{\times}10^{-10}$	$2.92{\times}10^{-15}$	$2.24{\times}10^{-23}$	$1.66{\times}10^{-31}$	$1.01{\times}10^{-47}$
$p+500$	$8.99{\times}10^{-8}$	$3.03{\times}10^{-10}$	$7.05{\times}10^{-16}$	$2.33{\times}10^{-24}$	$7.17{\times}10^{-33}$	$7.46{\times}10^{-50}$
$k_\ell + 5,\ p_\ell^* + 5,\ p = 303$						
$p+2$	$1.75{\times}10^{-7}$	$1.47{\times}10^{-10}$	$1.53{\times}10^{-16}$	$1.85{\times}10^{-25}$	$2.43{\times}10^{-34}$	$5.01{\times}10^{-52}$
$p+200$	$1.52{\times}10^{-7}$	$3.02{\times}10^{-10}$	$3.43{\times}10^{-16}$	$9.02{\times}10^{-25}$	$2.43{\times}10^{-33}$	$1.90{\times}10^{-50}$
$p+500$	$6.79{\times}10^{-8}$	$1.54{\times}10^{-10}$	$1.04{\times}10^{-16}$	$1.50{\times}10^{-25}$	$2.12{\times}10^{-34}$	$4.44{\times}10^{-52}$
$p+1000$	$2.75{\times}10^{-8}$	$6.65{\times}10^{-11}$	$2.57{\times}10^{-17}$	$1.70{\times}10^{-26}$	$1.07{\times}10^{-35}$	$4.24{\times}10^{-54}$

This asymptotic behavior is more marked for the near-exact distributions that match more exact moments. This may be seen from the more accentuated decrease in the values of the measure Δ^* for these near-exact distributions, that is, e.g. increases either in sample size or in the number of variables make the values of the measure Δ^* to decrease more for the near-exact distributions that match 10 exact moments than for those that match 6 exact moments. It is interesting to note that even near-exact distributions that match a very small number of exact moments or even no exact moment, and that, as such, are much simpler in their structure, and faster to compute, exhibit these asymptotic behaviors, with the behavior of the near-exact distribution that matches no exact moment being absolutely remarkable. This latter one is a very simple near-exact distribution, for the computation of which we do not even need to solve the system of equations in (4.4). In this case we will have $m^+ = 0$, and from (4.5)-(4.9) it is easy to see that the near-exact distribution is just a GIG or a GNIG distribution, according to r in (4.2) being integer or not.

Table 2: Values of the measure Δ^* in (5.1) for different values of m^+ (the number of exact moments matched by the near-exact distributions) and for $m = 5$ and increasing values of k_ℓ, p_ℓ^* and N, for base values of $k_\ell = \{3, 2, 3, 4, 4\}$, $p_\ell^* = \{3, 5, 6, 4, 5\}$ and sample sizes $N = p + 2, 200, 500$, with $p = \sum_{\ell=1}^{m} k_\ell \times p_\ell^*$.

N	m^+					
	0	1	2	4	6	10
			$k_\ell + 0,\ p_\ell^* + 0,\ p = 73$			
$p+2$	$3.81{\times}10^{-6}$	$1.61{\times}10^{-8}$	$1.57{\times}10^{-13}$	$1.14{\times}10^{-21}$	$2.24{\times}10^{-27}$	$4.53{\times}10^{-40}$
$p+200$	$6.64{\times}10^{-7}$	$6.61{\times}10^{-9}$	$5.43{\times}10^{-15}$	$2.03{\times}10^{-21}$	$2.09{\times}10^{-28}$	$9.91{\times}10^{-43}$
$p+500$	$1.58{\times}10^{-7}$	$1.66{\times}10^{-9}$	$1.82{\times}10^{-15}$	$7.74{\times}10^{-23}$	$1.86{\times}10^{-30}$	$3.79{\times}10^{-46}$
			$k_\ell + 2,\ p_\ell^* + 0,\ p = 119$			
$p+2$	$1.39{\times}10^{-6}$	$3.35{\times}10^{-9}$	$4.07{\times}10^{-14}$	$1.39{\times}10^{-21}$	$5.16{\times}10^{-29}$	$1.00{\times}10^{-43}$
$p+200$	$4.80{\times}10^{-7}$	$2.78{\times}10^{-9}$	$2.26{\times}10^{-14}$	$6.90{\times}10^{-22}$	$2.09{\times}10^{-29}$	$2.40{\times}10^{-44}$
$p+500$	$1.37{\times}10^{-7}$	$8.61{\times}10^{-10}$	$3.48{\times}10^{-15}$	$3.46{\times}10^{-23}$	$3.29{\times}10^{-31}$	$3.67{\times}10^{-47}$
			$k_\ell + 5,\ p_\ell^* + 0,\ p = 188$			
$p+2$	$5.50{\times}10^{-7}$	$7.93{\times}10^{-10}$	$6.19{\times}10^{-15}$	$4.93{\times}10^{-23}$	$3.95{\times}10^{-31}$	$3.17{\times}10^{-47}$
$p+200$	$3.19{\times}10^{-7}$	$1.10{\times}10^{-9}$	$8.81{\times}10^{-15}$	$9.96{\times}10^{-23}$	$1.07{\times}10^{-30}$	$1.42{\times}10^{-46}$
$p+500$	$1.12{\times}10^{-7}$	$4.31{\times}10^{-10}$	$2.02{\times}10^{-15}$	$9.19{\times}10^{-24}$	$3.89{\times}10^{-32}$	$7.89{\times}10^{-49}$
			$k_\ell + 2,\ p_\ell^* + 2,\ p = 171$			
$p+2$	$6.57{\times}10^{-7}$	$1.05{\times}10^{-9}$	$3.25{\times}10^{-15}$	$2.49{\times}10^{-23}$	$2.25{\times}10^{-31}$	$2.55{\times}10^{-47}$
$p+200$	$3.44{\times}10^{-7}$	$1.33{\times}10^{-9}$	$2.73{\times}10^{-15}$	$2.87{\times}10^{-23}$	$3.31{\times}10^{-31}$	$5.50{\times}10^{-47}$
$p+500$	$1.16{\times}10^{-7}$	$4.92{\times}10^{-10}$	$4.57{\times}10^{-16}$	$1.95{\times}10^{-24}$	$8.74{\times}10^{-33}$	$2.10{\times}10^{-49}$
			$k_\ell + 2,\ p_\ell^* + 5,\ p = 249$			
$p+2$	$3.50{\times}10^{-7}$	$3.68{\times}10^{-10}$	$5.29{\times}10^{-16}$	$1.36{\times}10^{-24}$	$4.43{\times}10^{-33}$	$7.49{\times}10^{-50}$
$p+200$	$2.61{\times}10^{-7}$	$6.53{\times}10^{-10}$	$7.82{\times}10^{-16}$	$3.81{\times}10^{-24}$	$2.29{\times}10^{-32}$	$1.18{\times}10^{-48}$
$p+500$	$1.06{\times}10^{-7}$	$2.98{\times}10^{-10}$	$1.69{\times}10^{-16}$	$4.17{\times}10^{-25}$	$1.23{\times}10^{-33}$	$1.43{\times}10^{-50}$
			$k_\ell + 5,\ p_\ell^* + 2,\ p = 270$			
$p+2$	$2.38{\times}10^{-7}$	$2.28{\times}10^{-10}$	$5.34{\times}10^{-16}$	$1.13{\times}10^{-24}$	$2.43{\times}10^{-33}$	$1.28{\times}10^{-50}$
$p+200$	$1.90{\times}10^{-7}$	$4.31{\times}10^{-10}$	$1.78{\times}10^{-15}$	$4.88{\times}10^{-24}$	$1.93{\times}10^{-32}$	$3.20{\times}10^{-49}$
$p+500$	$8.00{\times}10^{-8}$	$2.06{\times}10^{-10}$	$3.53{\times}10^{-16}$	$7.31{\times}10^{-25}$	$1.40{\times}10^{-33}$	$5.37{\times}10^{-51}$
			$k_\ell + 5,\ p_\ell^* + 5,\ p = 393$			
$p+2$	$9.00{\times}10^{-8}$	$5.69{\times}10^{11}$	$2.98{\times}10^{-17}$	$1.40{\times}10^{-26}$	$7.15{\times}10^{-36}$	$2.13{\times}10^{-54}$
$p+200$	$9.26{\times}10^{-8}$	$1.35{\times}10^{-10}$	$8.75{\times}10^{-17}$	$1.03{\times}10^{-25}$	$1.25{\times}10^{-34}$	$1.93{\times}10^{-52}$
$p+500$	$4.72{\times}10^{-8}$	$7.98{\times}10^{-11}$	$3.32{\times}10^{-17}$	$2.48{\times}10^{-26}$	$1.82{\times}10^{-35}$	$1.00{\times}10^{-53}$
$p+1000$	$2.09{\times}10^{-8}$	$3.81{\times}10^{-11}$	$9.49{\times}10^{-18}$	$3.62{\times}10^{-27}$	$1.31{\times}10^{-36}$	$1.68{\times}10^{-55}$

6. POWER STUDIES

In order to try to assess the behavior of the test under the alternative hypothesis, some power studies, based on simulations, were carried out. These studies focused on two forms of violation of the null hypothesis: (i) the violation of the equality of the Δ_ℓ matrices inside each block of k_ℓ of these matrices and (ii) the violation of the block-independence inside each group of $p_\ell = p^*_\ell \times k_\ell$ variables (see (1.2)).

First of all we should bring to the attention of the reader the fact that we are working with a random vector

$$\underline{X} = \left[\,\underline{X}'_1, \ldots, \underline{X}'_m\right]',$$

where in turn, for $\ell = 1, \ldots, m$,

$$\underline{X}_\ell = \left[\underline{X}'_{\ell 1}, \ldots, \underline{X}'_{\ell k_\ell}\right]',$$

with

$$\underline{X}_{\ell j} \sim N_{p^*_\ell}\left(\underline{\mu}_{\ell j}, \Delta_\ell\right) \qquad j = 1, \ldots, k_\ell$$

for some positive-definite matrix Δ_ℓ and some real $p^*_\ell \times 1$ vector $\underline{\mu}_{\ell j}$, and with

$$Cov\left(\underline{X}_{\ell j}, \underline{X}_{\ell' j'}\right) = 0_{p^*_\ell \times p^*_{\ell'}}$$

for either $\ell = \ell'$ or $\ell \neq \ell'$, with $j \neq j'$ if $\ell = \ell'$.

To keep things not too much involved, mainly in terms of easiness of exposition and to restrain the number of possible scenarios, while at the same time being able to give a view of a quite wide variety of situations under the alternative hypothesis, we considered a case with $m = 2$ and $k_1 = 2$ and $k_2 = 3$ with $p^*_1 = 5$ and $p^*_2 = 2$, with

$$\Delta_1 = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 2 & 2/3 & 2/4 & 2/5 \\ 1/3 & 2/3 & 3 & 3/4 & 3/5 \\ 1/4 & 2/4 & 3/4 & 4 & 4/5 \\ 1/5 & 2/5 & 3/5 & 4/5 & 5 \end{bmatrix} \qquad \text{and} \qquad \Delta_2 = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 2 \end{bmatrix},$$

where the choice of Δ_1 and Δ_2 did not obey to any other particular criteria than that of being two positive-definite matrices.

In the next two subsections we will perform power studies for the cases of violation of the hypothesis of equality of the diagonal blocks within each block of k_ℓ matrices and the hypothesis of independence generating for each scenario 1 000 000 pseudo random samples of size 29.

6.1. Violation of the equality hypothesis

In order to implement the violation of the hypothesis of equality of the diagonal blocks inside each $I_{k_\ell} \otimes \Delta_\ell$ block ($\ell = 1, 2$), we considered 40 different scenarios, with Σ covariance matrices of the form

$$\begin{bmatrix} \delta_{11}\Delta_1 & 0 & 0 & 0 & 0 \\ 0 & \delta_{12}\Delta_1 & 0 & 0 & 0 \\ 0 & 0 & \delta_{21}\Delta_2 & 0 & 0 \\ 0 & 0 & 0 & \delta_{22}\Delta_2 & 0 \\ 0 & 0 & 0 & 0 & \delta_{23}\Delta_2 \end{bmatrix}$$

with δ_{11}, δ_{12}, δ_{21}, δ_{22} and δ_{23} assuming the values in Table 3. In this Table are also defined the values for δ_1^* and δ_2^*. These new parameters summarize in a single value, respectively, the variability of the combinations of the values of δ_{11}, δ_{12} and δ_{21}, δ_{22}, δ_{23}. The values of δ_1^* and δ_2^* in Table 3 are defined based on the rank of the values of $\sum_{i=1}^{k_\ell} \left(1/\delta_{\ell i} - 1/\,\overline{\delta}_\ell\right)^2$, ($\ell = 1, 2; k_\ell = 2, 3$), since this is for our purpose a more adequate measure of dispersion of the values of δ_{1j} and δ_{2j} ($j = 1, \ldots, k_\ell; \ell = 1, 2$) than the usual variance. We will see that with this choice

Table 3: Definition of the values for δ_1^* and δ_2^*.

δ_1^*	$\Longleftrightarrow$	δ_{11}	δ_{12}	δ_2^*	$\Longleftrightarrow$	δ_{21}	δ_{22}	δ_{23}
1		1	1	1		1	1	1
2		1	2	2		1	1	2
3		1/2	2	3		1/2	1	2
4		1/3	2	4		1/2	1/2	2
5		1/3	3	5		1/3	1	2
				6		1/3	1	3
				7		1/3	1/3	2
				8		1/3	1/3	3

for the definition of the values of δ_1^* and δ_2^*, the power of the test will be an increasing function of the values of both δ_1^* and δ_2^*. The values δ_{11}, δ_{12} and the values δ_{21}, δ_{22}, δ_{23} were indeed chosen in such a way that they would generate a wide range of values of δ_1^* and δ_2^* that could show how the power of the test behaves for this variety of values. We should remark that for $\delta_1^* = 1$ and $\delta_2^* = 1$ we are under the null hypothesis in (1.2), while for any other combination of values of δ_1^* and δ_2^* we will be under various forms of the alternative hypothesis due to the fact that for values of δ_1^* different from 1, the null hypothesis $H_{03|1,2}^1$ in (2.5) is violated, since in these cases we have $Var(\underline{X}_{11}) \neq Var(\underline{X}_{12})$, while for $\delta_2^* \neq 1$ it is the hypothesis $H_{03|1,2}^2$ in (2.5) that is violated, since for $\delta_2^* \neq 1$ we have at least two of $Var(\underline{X}_{21})$, $Var(\underline{X}_{22})$ or $Var(\underline{X}_{23})$ different. Increasing values of either δ_1^* or δ_2^* indicate a larger departure from H_0 in (1.2).

In Tables 4 and 5 we may analyze the power values for different values of δ_1^* and δ_2^*, respectively for $\alpha = 0.05$ and $\alpha = 0.01$. We may note how for $\delta_1^* = \delta_2^* = 1$, situation in which we are under the null hypothesis H_0 in (1.2), we obtain a value for power which coincides with the α value, showing the unbiasedness characteristic of the test. We may also see how power has a good rate of convergence towards 1 for increasing values of δ_1^* and δ_2^*.

Table 4: Power values, rounded to three decimal places, for $\alpha = 0.05$ and different values of δ_1^* and δ_2^*.

δ_1^*	δ_2^*							
	1	2	3	4	5	6	7	8
1	0.050	0.113	0.309	0.512	0.545	0.833	0.845	0.983
2	0.170	0.293	0.556	0.736	0.769	0.939	0.942	0.996
3	0.805	0.888	0.967	0.988	0.992	0.999	0.999	1.000
4	0.988	0.995	0.999	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5: Power values, rounded to three decimal places, for $\alpha = 0.01$ and different values of δ_1^* and δ_2^*.

δ_1^*	δ_2^*							
	1	2	3	4	5	6	7	8
1	0.010	0.030	0.118	0.259	0.278	0.605	0.634	0.927
2	0.050	0.109	0.290	0.482	0.513	0.802	0.816	0.975
3	0.559	0.696	0.871	0.941	0.954	0.992	0.992	1.000
4	0.940	0.970	0.993	0.998	0.999	1.000	1.000	1.000
5	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

In Figure 4 we present smoothed surface and line plots of the power values for the cases considered in Tables 4 and 5.

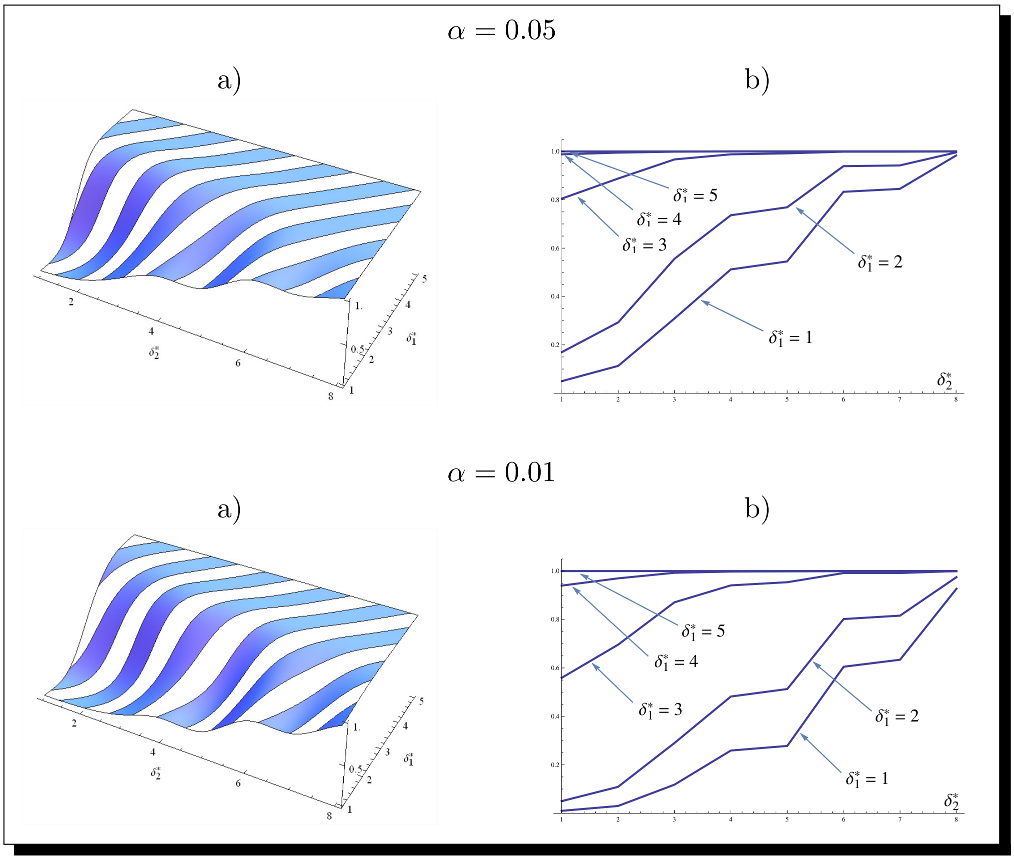

Figure 4: a) Smoothed surface plots and b) non-smoothed profile plots for different values of δ_1^* and running value of δ_2^*, for the values of power for the violation of the hypothesis of equality of diagonal blocks within each $I_{k_\ell} \otimes \Delta_\ell$ block ($\ell = 1, 2$).

6.2. Violation of the independence hypothesis

To implement the violation of the independence hypothesis, we consider 65 different scenarios with covariance matrices of the form

$$\begin{bmatrix} \Delta_1 & \gamma_1 C_1 & 0 & 0 & 0 \\ \gamma_1 C_1 & \Delta_1 & 0 & 0 & 0 \\ 0 & 0 & \Delta_2 & \gamma_{21} C_2 & \gamma_{22} C_2 \\ 0 & 0 & \gamma_{21} C_2 & \Delta_2 & \gamma_{23} C_2 \\ 0 & 0 & \gamma_{22} C_2 & \gamma_{23} C_2 & \Delta_2 \end{bmatrix}$$

where

$$C_1 = \frac{1}{10}\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 9 \end{bmatrix} \qquad \text{and} \qquad C_2 = \frac{1}{10}\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix},$$

and where γ_1 assumes the values $0.0, 1.0, 1.5, 1.75$ and 1.95 and γ_{21}, γ_{22} and γ_{23} assume the values in Table 6.

While for $\gamma_1 = 0$ we have the hypothesis $H^1_{02|1}$, of independence between $\underline{X}_{11}$ and $\underline{X}_{12}$, confirmed, for values of γ_1 different from zero we will be under the alternative hypothesis, since then the independence between these two sets of variables will be violated, with increasing values of γ_1 indicating an "increasing non-independence" of these two sets of variables, or equivalently, decreasing values of the determinant of the matrix

$$\Sigma_1 = \begin{bmatrix} \Delta_1 & \gamma_1 C_1 \\ \gamma_1 C_1 & \Delta_1 \end{bmatrix}.$$

In what concerns the values of γ_{21}, γ_{22} and γ_{23}, we will have the hypothesis of independence among $\underline{X}_{21}$, $\underline{X}_{22}$ and $\underline{X}_{23}$ confirmed when all these three parameters are equal to zero, and we will be under the alternative hypothesis if at least one of them is different from zero, with

$$\begin{aligned} \gamma_{21} \neq 0 &\implies Cov(\underline{X}_{21}, \underline{X}_{22}) = \gamma_{21} C_2 \neq 0\,, \\ \gamma_{22} \neq 0 &\implies Cov(\underline{X}_{21}, \underline{X}_{23}) = \gamma_{22} C_2 \neq 0\,, \\ \gamma_{23} \neq 0 &\implies Cov(\underline{X}_{22}, \underline{X}_{22}) = \gamma_{23} C_2 \neq 0\,. \end{aligned}$$

In order to define a hierarchy of the triplets of values of these three parameters, we compute the determinant of the matrix

$$\Sigma_2 = \begin{bmatrix} \Delta_2 & \gamma_{21} C_2 & \gamma_{22} C_2 \\ \gamma_{21} C_2 & \Delta_2 & \gamma_{23} C_2 \\ \gamma_{22} C_2 & \gamma_{23} C_2 & \Delta_2 \end{bmatrix}.$$

Values for $|\Sigma_2|$, as a function of the values of γ_{21}, γ_{22} and γ_{23} are shown in Table 6. These values are listed in decreasing order of $|\Sigma_2|$ and they are used to define the values of the new parameter γ^*, with increasing values of γ^* corresponding to decreasing values of $|\Sigma_2|$. The parameter γ^* is then used ahead in Tables 7 and 8 and Figure 5.

Then, while for $\gamma^* = 1$ we will be under the null hypothesis $H^2_{02|1}$ of independence among the sets of variables $\underline{X}_{21}$, $\underline{X}_{22}$ and $\underline{X}_{23}$, for increasing values of γ^* we will be increasingly further away from this null hypothesis.

We may see by looking at Tables 7 and 8 how the values for power give the value of α for $\gamma_1 = 0$ and $\gamma^* = 1$, situation in which we are under the null hypothesis H_0 in (1.2), while for all other combinations of values of these two parameters the value of power increases as the values of γ^* and/or γ_1 increase, easily reaching 1.

Table 6: Definition of the values of γ^* for the different values of the parameters γ^*_{21}, γ^*_{22} and γ^*_{23}.

γ_{21}	0.0	3.0	3.5	4.0	3.0	3.0	3.5	4.5	4.0	4.5	3.5	4.95	4.95
γ_{22}	0.0	0.0	0.0	0.0	3.0	3.0	3.5	0.0	4.0	4.5	3.5	0.0	4.95
γ_{23}	0.0	0.0	0.0	0.0	3.0	0.0	3.5	0.0	4.0	4.5	0.0	0.0	4.95
$\|\Sigma_2\|$	5.359	3.405	2.706	1.904	1.843	1.479	1.121	1.001	0.534	0.142	0.105	0.105	0.001
γ^*	1	2	3	4	5	6	7	8	9	10	11	12	13

Table 7: Power values, rounded to three decimal places, for $\alpha = 0.05$ and different values of γ_1 and γ^*.

γ_1	γ^*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
0.0	0.050	0.153	0.217	0.362	0.379	0.483	0.605	0.672	0.864	0.994	1.000	1.000	1.000
1.0	0.108	0.249	0.346	0.511	0.523	0.632	0.732	0.793	0.923	0.998	1.000	1.000	1.000
1.5	0.311	0.523	0.629	0.773	0.773	0.854	0.901	0.938	0.981	1.000	1.000	1.000	1.000
1.75	0.666	0.833	0.892	0.951	0.947	0.975	0.984	0.993	0.998	1.000	1.000	1.000	1.000
1.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 8: Power values, rounded to three decimal places, for $\alpha = 0.01$ and different values of γ_1 and γ^*.

γ_1	γ^*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
0.0	0.010	0.039	0.070	0.144	0.161	0.222	0.338	0.383	0.659	0.967	0.996	0.996	1.000
1.0	0.027	0.086	0.137	0.247	0.265	0.351	0.474	0.531	0.771	0.984	0.999	0.999	1.000
1.5	0.115	0.257	0.350	0.509	0.521	0.626	0.725	0.785	0.916	0.997	1.000	1.000	1.000
1.75	0.378	0.591	0.689	0.818	0.816	0.887	0.923	0.953	0.986	1.000	1.000	1.000	1.000
1.95	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

In Figure 5 we present smoothed surface and line plots of the power values for the cases considered in Tables 6 and 7. From the plots in this Figure and the values in Tables 7 and 8 we may see how power attains the value 1 for the larger values of γ_1 and γ^*, as expected. We may also note that as in the case of the previous subsection, for $\gamma_1 = 0$ and $\gamma^* = 1$, in which case we are under the null hypothesis, the value of the power equals the α value considered, showing again the unbiasedness characteristic of the test.

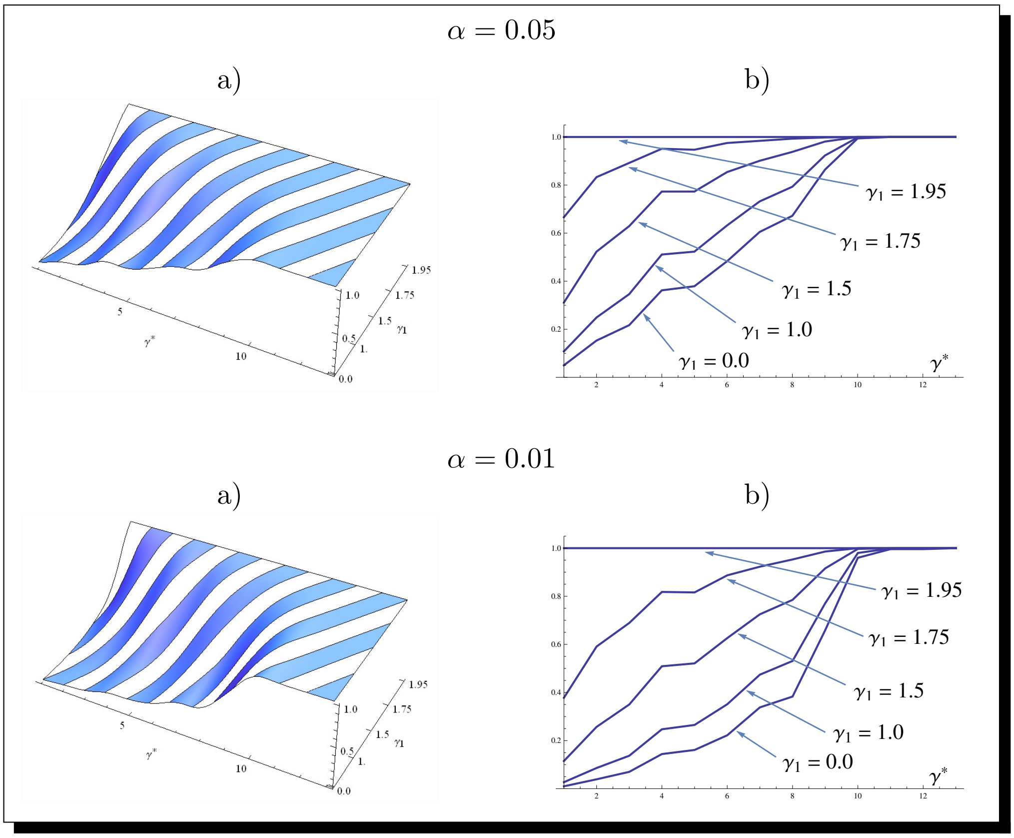

Figure 5: a) Smoothed surface plots; and b) non-smoothed profile plots, for the different values of γ_1 and running value of γ^*, for the values of power in Tables 7 and 8.

7. CONCLUSIONS

The procedure developed in this paper makes it possible to test elaborate covariance structures such as the HBM spherical structure through the use of very precise near-exact approximations. The testing procedure is based on an adequate decomposition of the overall null hypothesis into a sequence of sub-hypotheses, in our case the ones used to test the independence of several groups of variables and the equality of several covariance matrices in different sequences of covariance matrices. This decomposition of the null hypothesis allows us to obtain the likelihood ratio test statistic, the expression of its h-th moment and the expression of the characteristic function of its logarithm. Furthermore, the suitable decomposition of the null hypothesis also induces a factorization of this characteristic function which is the basis for the development of the near-exact approximations. These approximations can be easily implemented since there are already computational modules available in the internet for the two main distributions involved, which are the GNIG and GIG distributions.

The high precision of the near-exact distributions, which was assessed in the numerical studies section, makes them an efficient tool to obtain p-values and quantiles for the test statistic, even in cases where the sample size is very small and/or the number of variables is large.

Power studies conducted through simulations show the unbiased nature of the test as well as its good power properties, reaching rapidly powers close to 1 in the different scenarios considered.

The procedure developed may be very useful to address other, eventually, more complex structures. A natural extension of this framework is to consider the same global structure but for complex Normal random variables or even for quaternion random variables. Other possible extension is to consider specific structures for the block-diagonal covariance matrices such as the circular or the compound symmetry structures.

ACKNOWLEDGMENTS

This research was supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, project UID/MAT/00297/2013 (CMA/UNL).

The authors would also like to thank the referee and the Associate Editor for their comments which helped in improving the readability of the paper.

APPENDICES

APPENDIX A – Obtaining the expression of the l.r.t. statistic Λ, associated with the null hypothesis H_0 in (1.2), by using the definition of likelihood ratio statistic

Let us consider the vector $\underline{X} \sim N_p\left(\underline{\mu}, \Sigma\right)$ and let us suppose that we have a sample of size N from $\underline{X}$. The l.r.t statistic Λ associated with the HBM sphericity test is defined by

$$\Lambda = \frac{\sup L_0}{\sup L_1} \tag{A.1}$$

where L_0 is the likelihood function when the parameter space is under H_0 in (1.2) and L_1 is the likelihood function under the alternative hypothesis.

The likelihood function associated with the sample is

$$L\left(x_1,\ldots,x_N;\underline{\mu};\Sigma\right) = \frac{1}{(2\pi)^{\frac{Np}{2}}\,|\Sigma|^{\frac{N}{2}}}\,e^{-\frac{1}{2}tr\left(\left(\mathcal{X}-E_{N1}\underline{\mu}^T\right)\Sigma^{-1}\left(\mathcal{X}-E_{N1}\underline{\mu}^T\right)^T\right)}\,, \tag{A.2}$$

where E_{rs} denotes a matrix of 1's of dimension $r{\times}s$.

Let then $\mathcal{L}_0 = \mathcal{L}_0\left(x_1,\ldots,x_N;\underline{\mu};\Sigma_{|H_o}\right) = \log\left(L_0\left(x_1,\ldots,x_N;\Sigma_{|H_o}\right)\right)$. From (A.2) we have

$$\begin{aligned}\mathcal{L}_0 = -\frac{Np}{2}\log(2\pi) - \frac{N}{2}\log\left|\Sigma_{|H_o}\right| - \frac{1}{2}tr\left(\left(\mathcal{X}-E_{N1}\overline{\underline{x}}^T\right)\Sigma_{|H_o}^{-1}\left(\mathcal{X}-E_{N1}\overline{\underline{x}}^T\right)^T\right)\\ -\frac{1}{2}tr\left(\left(E_{N1}\overline{\underline{x}}^T - E_{N1}\underline{\mu}^T\right)\Sigma_{|H_o}^{-1}\left(E_{N1}\overline{\underline{x}}^T - E_{N1}\underline{\mu}^T\right)^T\right).\end{aligned}$$

As

$$tr\left(\left(\mathcal{X}-E_{N1}\overline{\underline{x}}^T\right)\Sigma_{|H_o}^{-1}\left(\mathcal{X}-E_{N1}\overline{\underline{x}}^T\right)^T\right) = tr\left(\Sigma_{|H_o}^{-1}\mathcal{A}\right)$$

where $\mathcal{A} = \left(\mathcal{X}-E_{N1}\overline{\underline{x}}^T\right)^T\left(\mathcal{X}-E_{N1}\overline{\underline{x}}^T\right) = \mathcal{X}^T\mathcal{X} - \frac{1}{N}\mathcal{X}^T E_{NN}\mathcal{X}$, and

$$\begin{aligned}tr\left(\left(E_{N1}\overline{\underline{x}}^T - E_{N1}\underline{\mu}^T\right)\Sigma_{|H_o}^{-1}\left(E_{N1}\overline{\underline{x}}^T - E_{N1}\underline{\mu}^T\right)^T\right) &= N tr\left(\Sigma_{|H_o}^{-1}\left(\overline{\underline{x}}-\underline{\mu}\right)\left(\overline{\underline{x}}-\underline{\mu}\right)^T\right)\\ &= N\left(\overline{\underline{x}}-\underline{\mu}\right)^T\Sigma_{|H_o}^{-1}\left(\overline{\underline{x}}-\underline{\mu}\right)\end{aligned}$$

the function $\mathcal{L}_0$ can be written as

$$\mathcal{L}_0 = -\frac{Np}{2}\log(2\pi) - \frac{N}{2}\log\left|\Sigma_{|H_o}\right| - \frac{1}{2}tr\left[\mathcal{A}\Sigma_{|H_o}^{-1}\right] - \frac{1}{2}N\left(\overline{\underline{x}}-\underline{\mu}\right)^T\Sigma_{|H_o}^{-1}\left(\overline{\underline{x}}-\underline{\mu}\right).$$

Given that

$$\begin{aligned}\left|\Sigma_{|H_o}\right| &= \left|bdiag\left(I_{k_\ell}\otimes\Delta_\ell,\ell=1,...,m\right)\right| = \prod_{\ell=1}^{m}\left|I_{k_\ell}\otimes\Delta_\ell\right| \\ &= \prod_{\ell=1}^{m}\left|I_{k_\ell}\right|^{p_\ell^*}\left|\Delta_\ell\right|^{k_\ell} = \prod_{\ell=1}^{m}\left|\Delta_\ell\right|^{k_\ell}\end{aligned}$$

where Δ_ℓ is a matrix of order p_ℓ^*, and

$$\begin{aligned}\Sigma_{|H_o}^{-1} &= \Big(bdiag\left(I_{k_\ell}\otimes\Delta_\ell,\ell=1,...,m\right)\Big)^{-1} \\ &= bdiag\Big(\left(I_{k_\ell}\otimes\Delta_\ell\right)^{-1},\ell=1,...,m\Big) \\ &= bdiag\left(I_{k_\ell}\otimes\Delta_\ell^{-1},\ell=1,...,m\right)\end{aligned}$$

with

$$\begin{aligned}tr\left(\mathcal{A}\Sigma_{|H_o}^{-1}\right) &= tr\Big(\mathcal{A}\,.\,bdiag\left(I_{k_\ell}\otimes\Delta_\ell^{-1},\ell=1,\ldots,m\right)\Big) \\ &= \sum_{\ell=1}^{m} tr\left(\mathcal{A}_{\ell\ell}\left(I_{k_\ell}\otimes\Delta_\ell^{-1}\right)\right) \\ &= \sum_{\ell=1}^{m}\sum_{v=1}^{k_\ell} tr\left(\mathcal{A}_{\ell\ell}^{v}\Delta_\ell^{-1}\right),\end{aligned}$$

we can write $\mathcal{L}_0$ as

$$\begin{aligned}\mathcal{L}_0 = -\frac{Np}{2}\log\left(2\pi\right) - \frac{N}{2}\log\left(\prod_{\ell=1}^{m}\left|\Delta_\ell\right|^{k_\ell}\right) - \frac{1}{2}\sum_{\ell=1}^{m}\sum_{v=1}^{k_\ell} tr\left(\mathcal{A}_{\ell\ell}^{v}\Delta_\ell^{-1}\right) \\ -\frac{1}{2}N\left(\underline{\overline{x}}-\underline{\mu}\right)^T\Sigma_{|H_o}^{-1}\left(\underline{\overline{x}}-\underline{\mu}\right).\end{aligned}$$

By solving the system of likelihood equations

$$\begin{aligned}&\begin{cases}\left.\frac{\partial\mathcal{L}_0}{\partial\underline{\mu}}\right|_{\underline{\mu}=\underline{\widehat{\mu}}}=0 \\ \left.\frac{\partial\mathcal{L}_0}{\partial\Delta_\ell}\right|_{\substack{\Delta_\ell=\widehat{\Delta}_\ell\\ \underline{\mu}=\underline{\widehat{\mu}}}}=0\,,\;\;(\ell=1,...,m)\end{cases} \Leftrightarrow \begin{cases}N\left(\underline{\overline{x}}-\underline{\widehat{\mu}}\right)\Sigma_{|H_o}^{-1}=0 \\ \\ \text{———}\end{cases} \\ &\Leftrightarrow \begin{cases}\underline{\widehat{\mu}}=\underline{\overline{x}} \\ -\frac{N}{2}k_\ell\widehat{\Delta}_\ell^{-1}+\frac{1}{2}\widehat{\Delta}_\ell^{-1}\sum_{v=1}^{k_\ell}A_{\ell\ell}^{v}\widehat{\Delta}_\ell^{-1}=0\,, \\ \qquad\qquad\qquad\qquad(\ell=1,...,m)\end{cases} \\ &\Leftrightarrow \begin{cases}\underline{\widehat{\mu}}=\underline{\overline{x}} \\ \widehat{\Delta}_\ell=\frac{1}{Nk_\ell}\sum_{v=1}^{k_\ell}A_{\ell\ell}^{v}\,,\quad(\ell=1,...,m)\end{cases}\end{aligned}$$

we obtain the maximum likelihood estimators of $\underline{\mu}$ and Σ under H_0, which are, $\widehat{\underline{\mu}} = \overline{\underline{X}}$ and

$$\widehat{\Sigma} = \begin{bmatrix} I_{k_1} \otimes \widehat{\Delta}_1 & & 0 \\ & \ddots & \\ 0 & & I_{k_m} \otimes \widehat{\Delta}_m \end{bmatrix} = \begin{bmatrix} I_{k_1} \otimes \frac{1}{Nk_1} A_1^* & & 0 \\ & \ddots & \\ 0 & & I_{k_m} \otimes \frac{1}{Nk_m} A_m^* \end{bmatrix}$$

where $A_\ell^* = \sum_{v=1}^{k_\ell} A_{\ell\ell}^v$ $(\ell = 1, ..., m)$.

Then we have

(A.3)
$$\begin{aligned}
&\sup L_0\left(X_1, \ldots, X_N; \underline{\mu}; \Sigma_{|H_o}\right) \\
&= (2\pi)^{-\frac{Np}{2}} \left\{\prod_{\ell=1}^{m} |\widehat{\Delta}_\ell|^{k_\ell}\right\}^{-\frac{N}{2}} e^{-\frac{1}{2}tr\left(A.bdiag\left(I_{k_\ell}\otimes\widehat{\Delta}_\ell^{-1},\ell=1,\ldots,m\right)\right)} \\
&= (2\pi)^{-\frac{Np}{2}} N^{\frac{pN}{2}} \prod_{\ell=1}^{m} k_\ell^{\frac{p_\ell N}{2}} \prod_{\ell=1}^{m} |A_\ell^*|^{-\frac{k_\ell N}{2}} e^{-\frac{N}{2}\sum_{\ell=1}^{m} k_\ell tr\left(A_{\ell\ell}\left(I_{k_\ell}\otimes\left(A_\ell^*\right)^{-1}\right)\right)} \\
&= (2\pi)^{-\frac{Np}{2}} N^{\frac{pN}{2}} \prod_{\ell=1}^{m} k_\ell^{\frac{p_\ell N}{2}} \prod_{\ell=1}^{m} |A_\ell^*|^{-\frac{k_\ell N}{2}} e^{-\frac{N}{2}\sum_{\ell=1}^{m} k_\ell \sum_{v=1}^{k_\ell} tr\left(A_{\ell\ell}^v\left(A_\ell^*\right)^{-1}\right)} \\
&= (2\pi)^{-\frac{Np}{2}} N^{\frac{pN}{2}} \prod_{\ell=1}^{m} k_\ell^{\frac{p_\ell N}{2}} \prod_{\ell=1}^{m} |A_\ell^*|^{-\frac{k_\ell N}{2}} e^{-\frac{N}{2}\sum_{\ell=1}^{m} k_\ell tr\left(A_\ell^*\left(A_\ell^*\right)^{-1}\right)} \\
&= (2\pi)^{-\frac{Np}{2}} N^{\frac{pN}{2}} \prod_{\ell=1}^{m} k_\ell^{\frac{p_\ell N}{2}} \prod_{\ell=1}^{m} |A_\ell^*|^{-\frac{k_\ell N}{2}} e^{-\frac{N}{2}\sum_{\ell=1}^{m} k_\ell tr\left(I_{p_\ell^*}\right)} \\
&= (2\pi)^{-\frac{Np}{2}} N^{\frac{pN}{2}} \prod_{\ell=1}^{m} k_\ell^{\frac{p_\ell N}{2}} \prod_{\ell=1}^{m} |A_\ell^*|^{-\frac{k_\ell N}{2}} e^{-\frac{N}{2}\sum_{\ell=1}^{m} p_\ell} \\
&= (2\pi)^{-\frac{Np}{2}} N^{\frac{pN}{2}} \prod_{\ell=1}^{m} k_\ell^{\frac{p_\ell N}{2}} \prod_{\ell=1}^{m} |A_\ell^*|^{-\frac{k_\ell N}{2}} e^{-\frac{Np}{2}}.
\end{aligned}$$

Under H_1, the likelihood function is given by

$$\begin{aligned}
L_1\left(x_1, \ldots, x_N; \underline{\mu}; \Sigma\right) &= (2\pi)^{-\frac{Np}{2}} |\Sigma|^{-\frac{N}{2}} e^{-\frac{1}{2}tr\left(\left(\mathcal{X}-E_{N1}\underline{\mu}^T\right)^T\left(\mathcal{X}-E_{N1}\underline{\mu}^T\right)\Sigma^{-1}\right)} \\
&= (2\pi)^{-\frac{Np}{2}} |\Sigma|^{-\frac{N}{2}} e^{-\frac{1}{2}tr\left(\mathcal{A}\Sigma^{-1}\right)}
\end{aligned}$$

and $\mathcal{L}_1 = \mathcal{L}_1\left(x_1, \ldots, x_N; \underline{\mu}; \Sigma\right) = \log\left(L_1\left(x_1, \ldots, x_N; \underline{\mu}; \Sigma\right)\right)$ is given by

$$\mathcal{L}_1 = -\frac{Np}{2}\log(2\pi) + \frac{N}{2}\log|\Sigma| - \frac{1}{2}tr\left[\mathcal{A}\Sigma^{-1}\right] - \frac{1}{2}N\,tr\left(\left(\underline{\overline{x}}-\underline{\mu}\right)^T\left(\underline{\overline{x}}-\underline{\mu}\right)\Sigma^{-1}\right).$$

By solving the system of likelihood equations

$$\begin{cases} \left.\frac{\partial\mathcal{L}_1}{\partial\underline{\mu}}\right|_{\underline{\mu}=\widehat{\underline{\mu}}} = 0 \\ \left.\frac{\partial\mathcal{L}_1}{\partial\Sigma}\right|_{\substack{\Sigma=\widehat{\Sigma} \\ \underline{\mu}=\widehat{\underline{\mu}}}} = 0 \end{cases} \Leftrightarrow \begin{cases} N\left(\underline{\overline{x}} - \widehat{\underline{\mu}}\right)\Sigma^{-1} = 0 \\ -\frac{N}{2}\widehat{\Sigma}^{-1} + \frac{1}{2}\widehat{\Sigma}^{-1}A\widehat{\Sigma}^{-1} = 0 \end{cases} \Leftrightarrow \begin{cases} \widehat{\underline{\mu}} = \underline{\overline{x}} \\ \widehat{\Sigma} = \frac{1}{N}A \end{cases}$$

we conclude that

$$
\begin{aligned}
\sup L_1\left(X_1,\dots,X_N;\underline{\mu};\Sigma\right) &= (2\pi)^{-\frac{Np}{2}}\,|\widehat{\Sigma}|^{-\frac{N}{2}}\,e^{-\frac{1}{2}tr\left(A\widehat{\Sigma}^{-1}\right)}\\
&= (2\pi)^{-\frac{Np}{2}}\left|\tfrac{1}{N}A\right|^{-\frac{N}{2}}\,e^{-\frac{1}{2}tr\left(NAA^{-1}\right)}\\
&= (2\pi)^{-\frac{Np}{2}}\,N^{\frac{Np}{2}}\,|A|^{-\frac{N}{2}}\,e^{-\frac{1}{2}N\,tr(I_p)}\\
&= (2\pi)^{-\frac{Np}{2}}\,N^{\frac{Np}{2}}\,|A|^{-\frac{N}{2}}\,e^{-\frac{1}{2}Np}
\end{aligned}
\tag{A.4}
$$

Then, from (A.1), (A.3) and (A.4) we have

$$
\begin{aligned}
\Lambda &= \frac{\sup L_0\left(x_N;\underline{\mu};\Sigma_{|H_o}\right)}{\sup L_1\left(x_N;\underline{\mu};\Sigma\right)}\\
&= \frac{(2\pi)^{-\frac{Np}{2}}\,N^{\frac{Np}{2}}\left\{\prod_{\ell=1}^{m}k_\ell^{\frac{Np_\ell}{2}}\right\}\left\{\prod_{\ell=1}^{m}|A_\ell^*|^{-\frac{k_\ell N}{2}}\right\}e^{-\frac{Np}{2}}}{(2\pi)^{-\frac{Np}{2}}\,N^{\frac{Np}{2}}\,|A|^{-\frac{N}{2}}\,e^{-\frac{Np}{2}}}\\
&= \left\{\prod_{\ell=1}^{m}k_\ell^{\frac{Np_\ell}{2}}\right\}\frac{|A|^{\frac{N}{2}}}{\prod_{\ell=1}^{m}\left|A_\ell^*\right|^{\frac{k_\ell N}{2}}}
\end{aligned}
\tag{A.5}
$$

where the matrix A is the maximum likelihood estimator of Σ, $A_{\ell\ell}$ is the ℓ-th diagonal block of order $p_\ell = k_\ell \times p_\ell^*$ of A $(\ell = 1, ..., m)$, with $p = \sum_{\ell=1}^{m} p_\ell$ and $A_\ell^* = A_{\ell\ell}^1 + \cdots + A_{\ell\ell}^{k_\ell}$, where $A_{\ell\ell}^v$ is the v-th $(v = 1, ..., k_\ell)$ diagonal block of order p_ℓ^* of $A_{\ell\ell}$. We should note how expression (A.5) is the same as expression (2.11).

Appendix B – Shape parameters

The shape parameters $r^\ell_{3;k}$ in (3.5) are given by

$$\text{(B.1)}\quad r^\ell_{3,k} = \begin{cases} r^{\ell *}_k, & k=1,...,p^*_\ell-1, \\ & k\neq p^*_\ell-1-2\alpha^\ell_1 \\ r^{\ell *}_k+(p^*_\ell \bmod 2)\left(\alpha^\ell_2-\alpha^\ell_1\right)\left(k_\ell-\frac{p^*_\ell-1}{2}+k_\ell\left\lfloor\frac{p^*_\ell}{2k_\ell}\right\rfloor\right), & k=p^*_\ell-1-2\alpha^\ell_1 \end{cases}$$

where

$$\alpha^\ell = \left\lfloor \frac{p^*_\ell-1}{k_\ell}\right\rfloor,\ \alpha^\ell_1 = \left\lfloor \frac{k_\ell-1}{k_\ell}\frac{p^*_\ell-1}{2}\right\rfloor,\ \alpha^\ell_2 = \left\lfloor \frac{k_\ell-1}{k_\ell}\frac{p^*_\ell+1}{2}\right\rfloor,$$

and

$$r^{\ell *}_k = \begin{cases} c^\ell_k\,, & k=1,...,\alpha^\ell+1 \\ k_\ell\left(\left\lfloor\frac{p^*_\ell}{2}\right\rfloor-\left\lfloor\frac{k}{2}\right\rfloor\right), & k=\alpha^\ell+2,...,\min\left(p^*_\ell-2\alpha^\ell_1, p^*_\ell-1\right) \\ & \text{and } k=2+p^*_\ell-2\alpha^\ell_1,...,2\left\lfloor\frac{p^*_\ell}{2}\right\rfloor-1, \textit{step } 2 \\ k_\ell\left(\left\lfloor\frac{p^*_\ell+1}{2}\right\rfloor-\left\lfloor\frac{k}{2}\right\rfloor\right), & k=1+p^*_\ell-2\alpha^\ell_1,...,p^*_\ell-1, \textit{step } 2 \end{cases}$$

with

$$c_k = \left\lfloor\frac{k_\ell}{2}\right\rfloor\left((k-1)\,k_\ell-2\left((k_\ell+1) \bmod 2\right)\left\lfloor\frac{k}{2}\right\rfloor\right)+\left\lfloor\frac{k_\ell}{2}\right\rfloor\left\lfloor\frac{k_\ell+k \bmod 2}{2}\right\rfloor,$$

for $k=1,...,\alpha^\ell$, and

$$c^\ell_{\alpha+1} = -\left(\left\lfloor\frac{p^*_\ell}{2}\right\rfloor-\alpha^\ell\left\lfloor\frac{k_\ell}{2}\right\rfloor\right)^2+k_\ell\left(\left\lfloor\frac{p^*_\ell}{2}\right\rfloor-\left\lfloor\frac{\alpha^\ell+1}{2}\right\rfloor\right)$$
$$+\,(k_\ell \bmod 2)\left(\alpha^\ell\left\lfloor\frac{p^*_\ell}{2}\right\rfloor+\frac{\left(\alpha^\ell \bmod 2\right)}{4}-\frac{\left(\alpha^\ell\right)^2}{4}-\left(\alpha^\ell\right)^2\left\lfloor\frac{k_\ell}{2}\right\rfloor\right).$$

For the derivation of the expressions for these parameters see [14] and references therein, making the necessary adjustments.

APPENDIX C – The Mathematica® modules

The modules described in this Appendix are available at the web-page https://sites.google.com/site/nearexactdistributions/hyper-block-matrix-sphericity and may be downloaded from this web-page.

C.1 - Computation of the p.d.f. and c.d.f. of the near-exact distributions

The modules made available for the computation of the p.d.f. and c.d.f. of the near-exact distributions for Λ are called, respectively, `NEpdf` and `NEcdf`. These modules have 4 mandatory arguments, which are:

- the sample size,
- a list with the values of p^*_ℓ ($\ell = 1, \ldots, m$),
- a list with the values of k_ℓ ($\ell = 1, \ldots, m$),
- the running value where the p.d.f. or c.d.f. is to be computed,

and which have to be given in this order; and 5 optional arguments, which are:

- `nm`: the number of exact moments to be matched by the near-exact distribution, that is, the value of m^+ in (4.1) and (4.4)–(4.9) (default value: 4),
- `prec1`: the number of digits used to print the value of the p.d.f. or c.d.f. (default value: 10),
- `prec2`: the number of precision digits used in the computation of the p.d.f. or c.d.f. (default value: 200),
- `prec3`: the number of precision digits used to store the m^+ exact moments of $W = -\log \Lambda$ computed (default value: 200),
- `prec4`: number of precision digits used in the computation of the m^+ exact moments of $W = -\log \Lambda$ by the module that does this computation (default value: 1500).

These optional arguments may be given in any order, but they will have to be called by their names, as it is exemplified below. If not used, they will assume their default values.

These modules use a number of other modules available on the same web-page which compute the weights π_j and the rate parameter θ in (4.1), the shape and rate parameters in $\Phi_1(t)$ as well as other shape and rate parameters involved in the expressions of the near-exact p.d.f. and c.d.f.. The module that computes the weights π_j uses another module which computes the exact moments of $W = -\log \Lambda$ by applying a numerical method to the exact c.f. of W in (3.1).

For example to compute the near-exact c.d.f. of Λ, on a value near the 0.05 quantile, for a case with the same parameters as those used for the examples for which we computed power in Section 6, which was a case with $m = 2$, $p_1^* = 5$, $p_2^* = 2$, $k_1 = 2$ and $k_2 = 3$, using the default values for all optional arguments, we would use the first command in Figure 6. The second command in that same figure uses the option `prec1` in order to obtain an output with more digits. The options named `prec2`, `prec3`, and `prec4`, will usually not be necessary, unless one suspects from lack of precision in the result obtained, which may happen in cases where the number of variables or the sample size are very large. This fact is illustrated with the third command in Figure 6, where although 500 precision digits are requested for the internal representation of the exact moments of W, the result obtained is exactly the same as the one obtained with the second command. The fourth command in Figure 6 illustrates, together with the third one that the order in which the optional arguments are given is arbitrary.

```
NEcdf[29, {5, 2}, {2, 3}, 5.914780554*10^-44]
0.05000000000
NEcdf[29, {5, 2}, {2, 3}, 5.914780554*10^-44, prec1 → 30]
0.04999999999993341511569283887674
NEcdf[29, {5, 2}, {2, 3}, 5.914780554*10^-44, prec1 → 30, prec3 → 500]
0.04999999999993341511569283887674
NEcdf[29, {5, 2}, {2, 3}, 5.914780554*10^-44, prec3 → 500, prec1 → 30]
0.04999999999993341511569283887674
```

Figure 6: Mathematica® commands to be used with the module `NEcdf`.

We remark that, for a given level α, we should reject the null hypothesis when the computed value of the l.r.t. statistic is lower than the α-quantile of the l.r.t. statistic. As such, the computation of the c.d.f. for the l.r.t. statistic also gives automatically the p-value.

We may note the extremely low values that these quantiles attain. This is due to the fact that we chose to use the 'complete' l.r.t. statistic, that is, the l.r.t. statistic with its exponent $N/2$. This is indeed the case why some authors chose to use l.r.t. statistics without this exponent, to make these values not so close to zero, what in some cases may cause some numerical problems. But indeed this poses absolutely no problems to the computation of the near-exact p.d.f.'s or c.d.f.'s.

The computation of quantiles is done with the module `Quant`. Given the sample size and the values for p_ℓ^* and k_ℓ $(\ell = 1, \ldots, m)$, the module generates, by default, 10 pseudo-random samples, under the null hypothesis of hyper-block sphericity in (1.2), using then the empirical α-quantile as a 'starting value' for a Newton-type method, which will find the approximate near-exact quantile using the values of the near-exact p.d.f. and c.d.f. computed on the successive iteration values.

This module has 5 mandatory arguments which first one is the α value for the quantile and which last 4 are exactly the same as the 4 mandatory arguments for the modules `NEpdf` and `NEcdf`, given in the exact same order. This module also has 8 optional arguments, which are:

- `nm`: the number of exact moments to be matched by the near-exact distribution, that is, as for `NEpdf` and `NEcdf`, the value of m^+ in (4.1) and (4.4)–(4.9) (default value: 4),
- `prec1`: the number of digits used to print the value of the quantile (default value: 10),
- `prec2`: the number of precision digits used in the computation of the p.d.f. or c.d.f. for the implementation of a Newton-type method (default value: 400),
- `prec3`: the number of precision digits used to store the m^+ exact moments of $W = -\log\Lambda$ computed (default value: 200),
- `prec4`: number of precision digits used in the computation of the m^+ exact moments of $W = -\log\Lambda$ by the module that does this computation (default value: 1500),
- `eps`: the value of the minimum upper-bound for two consecutive quantile approximations obtained from the Newton-type method; if those two consecutive approximations differ a quantity that is less than `eps`, the process stops, giving as result the last approximation found (default value: 10^{-6} times the 'starting value'),
- `nsamp`: the number of pseudo-random samples generated by the module to obtain the 'starting value' (default value: 10).

In Figure 7 we present a few commands that may be used with the module `Quant` to compute the 0.05 quantile of Λ for the same scenario considered in Figure 6. The first command uses all optional arguments with their default values, which will be adequate for most cases. The second command uses the optional argument `prec1` to request 20 digits, instead of 10, for the approximate 0.05 quantile. We may see that when this second command is repeated, as the third command in Figure 7, the result obtained is different. There is indeed no problem, and for the attentive reader there should be not much of a surprise. What happens is that since we use for `eps` its default value, the precision obtained for the approximation of the quantile should ensure at least 6 decimal digits correct. This is exactly what happens. Indeed it seems that at least 11 digits are correct. Then the fourth and fifth commands give the same result, which should be correct for all digits displayed. They illustrate the fact that the order in which the optional arguments are given is arbitrary and also that by giving the optional argument `eps` a small enough value, in this case a value which would ensure that at least 21 digits of the approximate quantile are correct, we will always get the same result.

```
Quant[5 / 100, 29, {5, 2}, {2, 3}]
5.914780554 × 10^-44
Quant[5 / 100, 29, {5, 2}, {2, 3}, prec1 → 20]
5.9147805544727498310 × 10^-44
Quant[5 / 100, 29, {5, 2}, {2, 3}, prec1 → 20]
5.9147805544731374828 × 10^-44
Quant[5 / 100, 29, {5, 2}, {2, 3}, prec1 → 20, eps → 10 ^ -65]
5.9147805544731417794 × 10^-44
Quant[5 / 100, 29, {5, 2}, {2, 3}, eps → 10 ^ -65, prec1 → 20]
5.9147805544731417794 × 10^-44
```

Figure 7: Mathematica® commands to be used with the module `Quant`.

There is also another module called `Lambda`, which may be used to compute the value of the statistic Λ in (2.11) for a given dataset. This dataset has to be given in a file, with observations defining the rows and variables the columns. This module has 3 mandatory arguments, which are:

- the name of the data file (including the path),
- a list with the values of p^*_ℓ ($\ell = 1, \ldots, m$),
- a list with the values of k_ℓ ($\ell = 1, \ldots, m$),

and which have to be given in this order.

Further details on these modules and their use are available at the web-page https://sites.google.com/site/nearexactdistributions/hyper-block-matrix-sphericity.

REFERENCES

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed., J. Wiley & Sons, New York.

[2] Box, G.E.P. (1949). A general distribution theory for a class of likelihood criteria, *Biometrika*, **36**(3/4), 317–346.

[3] Cardeño, L. and Nagar, D.K. (2001). Testing block sphericity of a covariance matrix, *Divulgaciones Matemáticas*, **9**, 25–34.

[4] Chao, C.C. and Gupta, A.K. (1991). Testing of homogeneity of diagonal blocks with blockwise independence, *Communications in Statistics – Theory and Methods*, **20**, 1957–1969.

[5] Chen, P.; Dong, L.; Chen, W. and Lin, J.-G. (2013). Outlier Detection in Adaptive Functional-Coefficient Autoregressive Models Based on Extreme Value Theory, *Mathematical Problems in Engineering*, **2013**, Article ID 910828, 9 pp..

[6] Coelho, C.A. (1998). The Generalized Integer Gamma distribution – a basis for distributions in Multivariate Statistics, *Journal of Multivariate Analysis*, **64**(1), 86–102.

[7] Coelho, C.A. (2004). The generalized near-integer Gamma distribution: a basis for near-exact approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables, *Journal of Multivariate Analysis*, **89**(2), 191–218.

[8] Coelho, C.A. and Marques, F.J. (2009). The advantage of decomposing elaborate hypotheses on covariance matrices into conditionally independent hypotheses in building near-exact distributions for the test statistics, *Linear Algebra and its Applications*, **430**(10), 2592–2606.

[9] Coelho, C.A. and Marques, F.J. (2010). Near-exact distributions for the independence and sphericity likelihood ratio test statistics, *Journal of Multivariate Analysis*, **101**(3), 583–593.

[10] Jensen, S.T. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance, *Annals of Statistics*, **16**(1), 302–322.

[11] Kopylev, L. (2012). Constrained Parameters in Applications: Review of Issues and Approaches, *ISRN Biomathematics*, **2012**, Article ID 872956, 6 pp..

[12] Marques, F.J. and Coelho, C.A. (2008). Near-exact distributions for the sphericity likelihood ratio test statistic, *Journal of Statistical Planning and Inference*, **138**(3), 726–741.

[13] Marques, F.J. and Coelho, C.A. (2012). The block sphericity test – exact and near-exact distributions for the likelihood ratio statistic, *Mathematical Methods in the Applied Sciences*, **35**(4), 373–383.

[14] Marques, F.J.; Coelho, C.A. and Arnold, B.C. (2011). A general near-exact distribution theory for the most common likelihood ratio test statistics used in Multivariate Analysis, *TEST*, **20**(1), 180–203.

[15] Marques, F.J.; Coelho, C.A. and Marques, P. (2013). *The Block-Matrix Sphericity Test: Exact and Near-exact Distributions for the Test Statistic*. In "Recent Developments in Modeling and Applications in Statistics – Studies in Theoretical and Applied Statistics: Selected Papers of the Statistical Societies" (P.E. Oliveira, M.G. Temido, C. Henriques and M. Vichi, Eds.), Springer, 169–177.

[16] MAUCHLY, J.W. (1940). Significance test for sphericity of a normal n-variate distribution, *Annals of Mathematical Statistics*, **11**, 204–209.

[17] MUIRHEAD, R.J. (2005). *Aspects of Multivariate Statistical Theory*, 2nd ed., J. Wiley & Sons, New York.

[18] NAGAR, D.K.; JAIN, S.K. and GUPTA, A.K. (1985). Distribution of LRC for testing sphericity of a complex multivariate Gaussian model, *International Journal of Mathematics and Mathematical Sciences*, **8**(3), 555–562.

[19] NAGARSENKER, B.N. and NAGARSENKER, P.B. (1981). Distribution of the likelihood ratio statistic for testing sphericity structure for a complex normal covariance matrix, *Sankhya*, Ser. B, **43**(3), 352–359.

[20] NANDAKUMAR, K.; CHEN, Y.; DASS, S.C. and JAIN, A. (2008). Likelihood Ratio-Based Biometric Score Fusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**(2), 342–347.

[21] PIRIE, A.; WOOD, A.; LUSH, M.; TYRER, J. and PHAROAH, P.D. (2015). The effect of rare variants on inflation of the test statistics in case-control analyses, *BMC Bioinformatics*, **16**(53), 5 pp.

[22] RENCHER, A.C. and CHRISTENSEN, W.F. (2012). *Methods of Multivariate Analysis*, 3rd ed., J. Wiley & Sons, New York.

[23] TRICOMI, F.G. and ERDÉLYI, A. (1951). The asymptotic expansion of a ratio of gamma functions, *Pacific Journal of Mathematics*, **1**(1), 133–142.

[24] WILKS, S.S. (1932). Certain generalizations in the analysis of variance, *Biometrika*, **24**, 471–494.

[25] WILKS, S.S. (1935). On the independence of k sets of normally distributed statistical variables, *Econometrica*, **3**, 309–326.

[26] WILKS, S.S. (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses, *The Annals of Mathematical Statistics*, **9**(1), 60–62.

[27] ZHANG, J.; ZOU, C. and WANG, Z. (2010). A control chart based on likelihood ratio test for monitoring process mean and variability, *Quality and Reliability Engineering International*, **26**(1), 63–73.