MPRA
Munich Personal RePEc Archive

On testing for the mean vector of a multivariate distribution with generalized and 2-inverses

Duchesne, Pierre and Francq, Christian

Université de Montréal, Université Lille 3

January 2010

Online at https://mpra.ub.uni-muenchen.de/19740/
MPRA Paper No. 19740, posted 07 Jan 2010 08:34 UTC

On testing for the mean vector of a multivariate distribution with generalized and { 2 }- inverses*

Pierre DUCHESNE†

Département de mathématiques et statistique

Université de Montréal

Christian FRANCQ

EQUIPPE-GREMARS

Université Charles de Gaulle - Lille 3

December 15, 2009

Abstract

Generalized Wald's method constructs testing procedures having chi-squared limiting distributions from test statistics having singular normal limiting distributions by use of generalized inverses. In this article, the use of {2}-inverses for that problem is investigated, in order to propose new test statistics with convenient asymptotic chi-square distributions. Alternatively, Imhof-based test statistics can also be defined, which converge in distribution to weighted sum of chi-square variables; The critical values of such procedures can be found using Imhof's (1961) algorithm. The asymptotic distributions of the test statistics under the null and alternative hypotheses are discussed. Under fixed and local alternatives, the asymptotic powers are compared theoretically. Simulation studies are also performed to compare the exact powers of the test statistics in finite samples. A data analysis on the temperature and precipitation variability in the European Alps illustrates the proposed methods.

Key words and phrases: {2}-inverses; generalized Wald's method; generalized inverses; multivariate analysis; singular normal distribution.

Mathematics subject classification codes (2000): primary 62H10; secondary 62F05, 62H15.

1. INTRODUCTION

Let $\mathbf{T}_n = (T_{n1}, \ldots, T_{np})^\top$, $n \geq 1$, be a sequence of statistics, and introduce $\mathbf{Z}_n(\boldsymbol{\mu}) = n^{1/2}(\mathbf{T}_n - \boldsymbol{\mu})$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^\top$. The classical testing problem confronts hypotheses H_0 and H_1:

$$\begin{aligned} H_0 &: \boldsymbol{\mu} = \boldsymbol{\mu}_0, \\ H_1 &: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0. \end{aligned} \tag{1}$$

*Abbreviated title: "Testing for the mean vector of a multivariate distribution".

Corresponding author: Pierre Duchesne, Université de Montréal; Département de mathématiques et de statistique; C.P. 6128 Succursale Centre-Ville; Montréal, Québec H3C 3J7; Canada. tel: (514) 343-7267; fax: (514) 343-5700; e-mail: duchesne@dms.umontreal.ca.

†This work was supported by a grant from the Natural Science and Engineering Research Council of Canada.

Under the null hypothesis, it is assumed that the sequence of test statistics converges toward a normal distribution: $\mathbf{Z}_n := \mathbf{Z}_n(\boldsymbol{\mu}_0) \xrightarrow{d} \mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma})$ as $n \to \infty$, where $\mathbf{0}_p = (0, \ldots, 0)^\top$ is the p-dimensional null vector. The covariance matrix $\boldsymbol{\Sigma} \neq \mathbf{0}$, possibly singular, is supposed to be unknown. Based on sample data, the hypothesis testing problem is to decide whether it seems likely that the null hypothesis to be true. To study limiting alternative distributions, under the fixed alternative $H_1 : \boldsymbol{\mu} = \boldsymbol{\mu}_1$, we suppose that $\mathbf{Z}_n(\boldsymbol{\mu}_1) \xrightarrow{d} \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$ as $n \to \infty$, and under any sequence of local alternative hypotheses $H_{1n} : \boldsymbol{\mu}_n = \boldsymbol{\mu}_0 + n^{-1/2}\boldsymbol{\Delta}_n$, $\boldsymbol{\Delta}_n \to \boldsymbol{\Delta}$, we suppose instead that $\mathbf{Z}_n(\boldsymbol{\mu}_0) \xrightarrow{d} \mathcal{N}_p(\boldsymbol{\Delta}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Delta} \neq \mathbf{0}$. The problem described above represents a general framework for many statistical problems coming from multivariate analysis, non-linear regression models, and time series analysis.

1.1 Leading examples of the testing problem

Let $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})^\top$, $i = 1, \ldots, n$, be a random sample from a multivariate distribution, where $E(\mathbf{X}_i) = \boldsymbol{\mu}$ and $\mathrm{var}(\mathbf{X}_i) = E\{(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^\top\} = \boldsymbol{\Sigma}$. The sequence of test statistics could be simply based on the sample means, $\mathbf{T}_n = \bar{\mathbf{X}}_n = n^{-1}\sum_{i=1}^n \mathbf{X}_i$, and the asymptotic normal distribution under the null hypothesis is found invoking the multivariate central limit theorem. That framework goes back to the seminal work of Wald (1954). In fact, under the normality assumption, that is $\mathbf{X}_i \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with the covariance matrix $\boldsymbol{\Sigma}$ unknown but supposed positive definite, Hotelling's T^2 test statistic represents the classical test procedure for problem (1) and Wald (1954) simply developed the large sample analog of Hotelling's method. See, e.g., Srivastava and Khatri (1979), Muirhead (1982) or Anderson (1984), among others. When the sample is obtained from a singular multivariate distribution, Bhimasankaram and Sengupta (1991) proposed a methodology similar to Hotelling's test statistic. If $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\det(\boldsymbol{\Sigma}) = 0$, it is well-known that $\mathbf{X} - \boldsymbol{\mu}$ belongs to the column space of rank r, say, of the covariance matrix $\boldsymbol{\Sigma}$ with probability one (w.p.1); visually, the data lies in the r-dimensional affine subspace of $\mathbb{R}^p$, $r < p$. Furthermore, a certain linear transformation of $\mathbf{X} - \boldsymbol{\mu}$ follows a nonsingular normal distribution (see, e.g., Bilodeau and Brenner (1999, p. 62) or Eaton (2007)). Thus the initial problem can be reformulated in a smaller dimension using a non-singular normal distribution. However, to work with the data in the transformed scale may be seen as a disadvantage from a practical point of view, and, more importantly, the rank of $\boldsymbol{\Sigma}$ must be known a priori, which can be a restrictive assumption. Wald's method has been generalized by Moore (1977) to sequences of test statistics having singular normal distributions by a natural use of generalized inverses.

Multivariate sampling is just a simple example and the study of other test statistics may result in asymptotic singular normal distributions. In parametric models, when $\boldsymbol{\mu}$ is the vector of cell probabilities in a multinomial model and $\mathbf{T}_n$ represents the vector of observed relative frequencies, the covariance matrix $\boldsymbol{\Sigma}$ is singular; that example has been studied in detail by Moore (1977). In non-linear regression models, under certain conditions, the asymptotic distribution of the regression parameters in non-linear regression models is a singular normal distribution, see Robinson (1972). See also Hadi and Wells (1990), who give several examples of non-linear models with singular information matrices. Another example is taken from time series analysis, where a central problem is to test for serial correlation. It is well-known that the asymptotic distribution of a vector of fixed length of residual autocorrelations is approximately a singular normal distribution, see Box and Pierce (1970), Li and McLeod (1981) and Ljung (1986), amongst others.

1.2 Testing procedures

In this paper, we consider several test procedures which can be used whether $\boldsymbol{\Sigma}$ is singular or non-singular, without assuming normality. More specifically, we study the general class of test statistics:

$$\mathcal{Q}_n(\mathbf{W}_n) = \mathbf{Z}_n^\top \mathbf{W}_n \mathbf{Z}_n, \tag{2}$$

where $\mathbf{W}_n$ is a weight matrix. For the testing problem (1), the null is rejected for large values of $\mathcal{Q}_n(\mathbf{W}_n)$. We discuss in detail three testing procedures corresponding to the weighting matrices: (i) $\mathbf{W}_n = \mathbf{I}_p$, where $\mathbf{I}_p$ denotes the $p \times p$ identity matrix, (ii) $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$, where $\boldsymbol{\Sigma}_n^{-k}$ represents a $\{2\}$-inverse of $\boldsymbol{\Sigma}_n$, and finally (iii) $\mathbf{W}_n = \boldsymbol{\Sigma}_n^-$, where $\boldsymbol{\Sigma}_n^-$ is the Moore-Penrose inverse (or pseudo-inverse) of $\boldsymbol{\Sigma}_n$. The estimator $\boldsymbol{\Sigma}_n$ is assumed to be strongly consistent for $\boldsymbol{\Sigma}$, that is $\boldsymbol{\Sigma}_n \to \boldsymbol{\Sigma}$, almost surely. Note that the strong consistency of $\boldsymbol{\Sigma}_n$ is assumed to be true under the null hypothesis, under sequences of local alternatives and for fixed alternatives. Our framework is general, and the test statistics can be applied for all the testing problems described above. For example, in multivariate sampling, natural candidates would be procedures based on sample means for $\mathbf{Z}_n$, and $\boldsymbol{\Sigma}_n$ could be the sample covariance matrix $\mathbf{S}_n = n^{-1}\sum_{i=1}^n(\mathbf{X}_i - \bar{\mathbf{X}}_n)(\mathbf{X}_i - \bar{\mathbf{X}}_n)^\top$. The test statistics in class (i) are simply based on the usual Euclidian norm of $\mathbf{Z}_n$. The test procedures in class (ii) consider to use $\{2\}$-inverses for the weight matrix $\mathbf{W}_n$. The literature on generalized and pseudo-inverses is considerable, see, e.g., Rao and Mitra (1971) and Rao (1973), among others. On the other hand, the important role of $\{2\}$-inverses in statistics, more particularly in the study of quadratic forms, seems less well documented. Applications of $\{2\}$-inverses in statistics are described in Getson and Hsuan (1988). Finally, the class of test statistics (iii) is composed of the generalized Wald's test statistics introduced in Moore (1977). See also Andrews (1987) and Hadi and Wells (1990). Duchesne and Francq (2008) investigated diagnostic checking time series models with portmanteau test statistics relying on generalized inverses and $\{2\}$-inverses. In their applications, $\mathbf{Z}_n$ was based on a vector of sample autocorrelations and $\boldsymbol{\Sigma}_n$ was a certain consistent estimator of the asymptotic covariance matrix of the sample autocorrelations. Here the framework is considerably more general, and we investigate the theoretical and empirical properties of the test statistics $\mathcal{Q}_n(\mathbf{W}_n)$ under fixed and local alternatives.

The paper is organized as follows. In Section 2, we discuss the asymptotic distributions of the test statistics $\mathcal{Q}_n(\mathbf{I}_d)$, $\mathcal{Q}_n(\boldsymbol{\Sigma}_n^{-k})$ and $\mathcal{Q}_n(\boldsymbol{\Sigma}_n^-)$ under null and local hypotheses. The asymptotic powers of these test statistics are compared in Sections 3 and 4, under fixed and local alternatives, respectively. In Section 5, some simulation experiments are conducted. A data analysis is presented in Section 6 on the monthly temperature and precipitation variability in the European Alps for the period 1659-1999. Concluding remarks are offered in Section 7. An Appendix gives some technical details concerning the construction of the test statistics based on $\{2\}$-inverses.

2. ASYMPTOTIC DISTRIBUTIONS UNDER THE NULL HYPOTHESIS

For a reason that will be transparent in the next section, the test statistic based on $\mathcal{Q}_n(\mathbf{I}_d)$ will be called Imhof-based test.

2.1 The Imhof-based test statistic (case $\mathbf{W}_n = \mathbf{I}_p$)

A simple and natural test procedure leads to the study of the norm of $\mathbf{Z}_n$, namely $||\mathbf{Z}_n||^2$. The asymptotic distribution of $\mathcal{Q}_n(\mathbf{I}_p)$ under the null hypothesis H_0 follows easily invoking the multivariate central limit theorem and a spectral decomposition of $\mathbf{\Sigma}$. Consider the spectral decomposition $\mathbf{\Sigma} = \mathbf{P\Lambda P}^\top = \sum_{i=1}^p \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$ with $\mathbf{P}^\top\mathbf{P} = \mathbf{I}_p$, where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_p)$ and the columns $\mathbf{v}_1, \ldots, \mathbf{v}_p$ of the matrix $\mathbf{P}$ constitute an orthonormal basis of $\mathbb{R}^p$. The weights $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ and $\lambda_{r+1} = \ldots = \lambda_p = 0$ represent the eigenvalues of $\mathbf{\Sigma}$ and thus $r = \text{rank}(\mathbf{\Sigma})$. The chi-squared distribution with k degrees of freedom and non-centrality parameter c is noted $\chi_k^2(c)$. The vector $\mathbf{1}_p = (1, \ldots, 1)^\top$ corresponds to a p-dimensional vector composed only of ones.

Proposition 1. *If $\mathbf{Z}_n = \mathbf{Z}_n(\boldsymbol{\mu}_0) \xrightarrow{d} \mathcal{N}_p(\mathbf{\Delta}, \mathbf{\Sigma})$ and $\mathbf{\Sigma}_n \to \mathbf{\Sigma}$ in probability as $n \to \infty$, then*

$$\mathbf{Z}_n^\top \mathbf{Z}_n \xrightarrow{d} \sum_{i=1}^r \lambda_i \chi_1^2(\theta_i^{*2}) + \boldsymbol{\theta}_0^\top \boldsymbol{\theta}_0 := \sum_{i=1}^r \lambda_i N_i^{*2} + \boldsymbol{\theta}_0^\top \boldsymbol{\theta}_0,$$

where $\boldsymbol{\theta}_0 = diag\left(\mathbf{0}_r^\top, \mathbf{1}_{p-r}^\top\right) \mathbf{P}^\top \mathbf{\Delta}$. The non-centrality parameters θ_i^* satisfy $\boldsymbol{\theta}^* = (\theta_1^*, \ldots, \theta_p^*)^\top = \mathbf{\Lambda}^{1/2-} \mathbf{P}^\top \mathbf{\Delta}$ with $\mathbf{\Lambda}^{1/2-} = diag\left(\lambda_1^{-1/2}, \ldots, \lambda_r^{-1/2}, \mathbf{0}_{p-r}^\top\right)$, and the normal random variables N_i^* are components of the singular normal distribution $\mathbf{N}^* = (N_1^*, \ldots, N_p^*)^\top \sim \mathcal{N}_p(\boldsymbol{\theta}^*, \mathbf{I}_p^{-r})$ with $\mathbf{I}_p^{-r} = diag\left(\mathbf{1}_r^\top, \mathbf{0}_{p-r}^\top\right)$.

Proof: The proposition is a well known consequence of the continuous mapping theorem and of standard results on quadratic forms of Gaussian vectors. See, e.g., Rao and Mitra (1971) or Rao (1973) for distributions of quadratic functions in normal random variables. More precisely, it is obtained by noting that $\mathbf{N}_n^* := \mathbf{\Lambda}^{1/2-} \mathbf{P}^\top \mathbf{Z}_n \xrightarrow{d} \mathbf{N}^* = \mathcal{N}_p(\boldsymbol{\theta}^*, \mathbf{I}_p^{-r})$ and that $\|\mathbf{Z}_n\|^2 = \|\mathbf{P}^\top \mathbf{Z}_n\|^2 = \|\mathbf{\Lambda}^{1/2} \mathbf{N}_n^* + \boldsymbol{\theta}_0\|^2$. $\square$

Under the null hypothesis $\mathbf{\Delta} = \mathbf{0}_p$ in Proposition 1 and it follows that:

$$\mathcal{Q}_n(\mathbf{I}_p) = ||\mathbf{Z}_n||^2 \xrightarrow{d} \sum_{i=1}^r \lambda_i N_i^2, \tag{3}$$

where $N_1, \ldots, N_r$ correspond to independent $\mathcal{N}(0,1)$ random variables. In practice, it is possible to evaluate the distribution of the Gaussian quadratic form in (3) by means of the algorithm of Imhof (1961). More precisely, the test procedure based on the weight matrix $\mathbf{W}_n = \mathbf{I}_p$ relies on the following steps: 1) Compute the eigenvalues $\hat{\lambda}_1, \ldots, \hat{\lambda}_p$ of $\mathbf{\Sigma}_n$, which provides a consistent estimator $\mathbf{\Sigma}$. 2) Evaluate the $(1-\alpha)$-quantile $c_\alpha(\hat{\lambda}_1, \ldots, \hat{\lambda}_p)$ of $\sum_{i=1}^p \hat{\lambda}_i N_i^2$ using Imhof's algorithm, and finally 3) the null hypothesis is rejected when $\mathcal{Q}_n(\mathbf{I}_p) \geq c_\alpha(\hat{\lambda}_1, \ldots, \hat{\lambda}_p)$. An interesting advantage of that procedure is that the rank of $\mathbf{\Sigma}$ does not need to be known and the procedure is in all points the same whether $\mathbf{\Sigma}$ is singular or non-singular.

In a time series framework, a similar procedure has been considered in Francq, Roy and Zakoïan (2005) in testing for null autocorrelations in the residuals from autoregressive-moving-average (ARMA) models. In their context, they specified $\mathbf{Z}_n = n^{1/2}\hat{\boldsymbol{\rho}}$, with $\hat{\boldsymbol{\rho}} = (\hat{\rho}(1), \ldots, \hat{\rho}(p))^\top$, where $\hat{\rho}(h)$ denotes the lag-h residual autocorrelation, $h = 1, \ldots, p$. In order to test null autocorrelations, their test statistic reduced to the so-called Box-Pierce test statistic $||\mathbf{Z}_n||^2 = n \sum_{h=1}^{p} \hat{\rho}^2(h)$ proposed by Box and Pierce (1970), which relies on the sum of squared residual autocorrelations. See for example Li (2004), among others. In their procedure, the quantiles were found estimating the eigenvalues of a consistent estimator of the asymptotic covariance matrix of the residual autocorrelations. That strategy has been adapted in Duchesne and Francq (2008) for diagnostic checking non-linear time series models. Note that the use of the test statistic $\mathcal{Q}_n(\mathbf{I}_p)$ in the context of multivariate sampling seems to be absent from the literature, probably due to the non-standard limiting distribution of that test procedure.

2.2 The test statistic based on $\{2\}$-inverses (case $\mathbf{W}_n = \mathbf{\Sigma}_n^{-k}$)

A generalized inverse (g-inverse) of $\mathbf{\Sigma}$ is a matrix $\tilde{\mathbf{\Sigma}}$ satisfying the condition:

$$\mathbf{\Sigma}\tilde{\mathbf{\Sigma}}\mathbf{\Sigma} = \mathbf{\Sigma}. \tag{4}$$

It is also called a $\{1\}$-inverse since (4) is usually the first of the four conditions defining the (unique) Moore-Penrose inverse of $\mathbf{\Sigma}$ (see, e.g., Getson and Hsuan (1988)). On the other hand, a $\{2\}$-inverse of $\mathbf{\Sigma}$ is any matrix $\mathbf{\Sigma}^*$ satisfying the second relation defining the Moore-Penrose inverse of $\mathbf{\Sigma}$, that is:

$$\mathbf{\Sigma}^*\mathbf{\Sigma}\mathbf{\Sigma}^* = \mathbf{\Sigma}^*. \tag{5}$$

When requirements (4) and (5) are satisfied, the resulting matrix is called a $\{1,2\}$-inverse or reflexive g-inverse (see, e.g., Rao (1973, p. 25)).

Note that the matrix $\mathbf{\Sigma}^- = \mathbf{P}\mathbf{\Lambda}^-\mathbf{P}^\top$ is the Moore-Penrose inverse (or pseudo-inverse) of $\mathbf{\Sigma}$, where $\mathbf{\Lambda}^- = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_r^{-1}, \mathbf{0}_{p-r}^\top)$. For $k = 1, \ldots, r$, define the matrix $\mathbf{\Sigma}^{-k} = \mathbf{P}\mathbf{\Lambda}^{-k}\mathbf{P}^\top$, where $\mathbf{\Lambda}^{-k} = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_k^{-1}, \mathbf{0}_{p-k}^\top)$. The matrix $\mathbf{\Sigma}^{-k}$ is always a $\{2\}$-inverse, but this is not a g-inverse of $\mathbf{\Sigma}$ when $k < r$.

Empirical versions are easily constructed. Since $\mathbf{\Sigma}_n \to \mathbf{\Sigma}$ almost surely, as $n \to \infty$, a natural estimator of $\mathbf{\Sigma}^{-k}$ relies on $\mathbf{\Sigma}_n^{-k}$, where the spectral decomposition of $\mathbf{\Sigma}_n$ is given by $\mathbf{\Sigma}_n = \mathbf{P}_n\mathbf{\Lambda}_n\mathbf{P}_n^\top$.

When all the non null eigenvalues of $\mathbf{\Sigma}$ are distinct, the matrix $\mathbf{\Sigma}^{-k}$ is uniquely defined. However, when some eigenvalues display multiplicities, $\mathbf{\Sigma}^{-k}$ is not uniquely defined, because it depends on the particular choice of the orthonormal basis in the spectral decompositions of $\mathbf{\Sigma}$. That caveat is fixed using projections and the Gram-Schmidt orthogonalization process. Consider an arbitrary basis $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ of $\mathbb{R}^p$. For each eigenvalue $\lambda_k = \lambda_k(\mathbf{\Sigma})$ of multiplicity $m_k(\mathbf{\Sigma})$, let $\mathcal{V} = \mathcal{V}_k(\mathbf{\Sigma})$ be the associated eigenspace with $\dim(\mathcal{V}) = m_k(\mathbf{\Sigma})$. Single eigenvalues do not pose problem; consequently suppose $k > 1$. The projection on $\mathcal{V}$ is denoted $\mathbf{P}_\mathcal{V}$, which is uniquely defined. The vectors $\mathbf{P}_\mathcal{V}(\mathbf{u}_1), \ldots, \mathbf{P}_\mathcal{V}(\mathbf{u}_p)$ span $\mathcal{V}$ since any vector $\mathbf{v} \in \mathcal{V}$ can be expressed as $\mathbf{v} = \sum_{i=1}^{p} c_i\mathbf{u}_i = \mathbf{P}_\mathcal{V}(\mathbf{v}) = \sum_{i=1}^{p} c_i\mathbf{P}_\mathcal{V}(\mathbf{u}_i)$. From the vectors $\mathbf{P}_\mathcal{V}(\mathbf{u}_1), \ldots, \mathbf{P}_\mathcal{V}(\mathbf{u}_p)$, a basis $B_\mathcal{V} = \{\mathbf{P}_\mathcal{V}(\mathbf{u}_{i_1}), \ldots, \mathbf{P}_\mathcal{V}(\mathbf{u}_{i_k})\}$ of $\mathcal{V}$ is extracted as follows: let i_1 be the smallest index of $\{1, \ldots, p\}$ such that $\mathbf{P}_\mathcal{V}(\mathbf{u}_{i_1}) \neq \mathbf{0}$ and for

$\ell \in \{2, \ldots, k\}$, i_ℓ represents the smallest index of $\{i_{\ell-1}+1, \ldots, p\}$ such that $\mathbf{P}_{\mathcal{V}}(\mathbf{u}_{i_\ell})$ is not spanned by $\{\mathbf{P}_{\mathcal{V}}(\mathbf{u}_{i_1}), \ldots, \mathbf{P}_{\mathcal{V}}(\mathbf{u}_{i_{\ell-1}})\}$. Using the Gram-Schmidt process, the basis $B_{\mathcal{V}}$ is transformed in an orthonormal basis of $\mathcal{V}$. This process allows to define a unique common basis of eigenvectors for the spectral composition of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^{-k}$. More precisely, we define a unique matrix $\mathbf{P}_B$ such that $\boldsymbol{\Sigma} = \mathbf{P}_B \boldsymbol{\Lambda} \mathbf{P}_B^\top$ and $\boldsymbol{\Sigma}^{-k} = \mathbf{P}_B \boldsymbol{\Lambda}^{-k} \mathbf{P}_B^\top$ for all $k \leq r$. The matrix $\mathbf{P}_B$ will be called the *B-eigenvector matrix* of $\boldsymbol{\Sigma}$.

A similar construction holds for $\boldsymbol{\Sigma}_n$ with eigenvalues $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_p$. Since the B-eigenvector matrices $\mathbf{P}_B$ and $\mathbf{P}_{n,B}$ of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_n$ are uniquely defined, the $\{2\}$-inverses $\boldsymbol{\Sigma}^{-k} = \boldsymbol{\Sigma}_B^{-k}$ and $\boldsymbol{\Sigma}_n^{-k} = \boldsymbol{\Sigma}_{n,B}^{-k}$ are now uniquely defined by $\boldsymbol{\Sigma}_B^{-k} = \mathbf{P}_B \boldsymbol{\Lambda}^{-k} \mathbf{P}_B^\top$ and $\boldsymbol{\Sigma}_{n,B}^{-k} = \mathbf{P}_{n,B} \boldsymbol{\Lambda}_n^{-k} \mathbf{P}_{n,B}^\top$. An algorithm is given in the Appendix on the construction of the matrices $\boldsymbol{\Sigma}_B^{-k}$ and $\boldsymbol{\Sigma}_{n,B}^{-k}$. In practice, a tolerance is needed to estimate the rank of $\boldsymbol{\Sigma}$, the multiplicities, and to distinguish null and non null eigenvalues. The proposed algorithm defines a function, noted $\mathfrak{A}_{B,k,\epsilon}(\cdot)$, based on a tolerance ϵ. The following Assumption A(ϵ) is necessary in order to specify the minimum distance between the different eigenvalues of $\boldsymbol{\Sigma}$.

Assumption A(ϵ). Let $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an arbitrary basis of $\mathbb{R}^p$. The tolerance $\epsilon > 0$ is such that:

C_1: $\min\{|\lambda_i(\boldsymbol{\Sigma}) - \lambda_j(\boldsymbol{\Sigma})| : \lambda_i(\boldsymbol{\Sigma}) \neq \lambda_j(\boldsymbol{\Sigma})\} > \epsilon$,

C_2: For $k \leq \mathrm{rank}(\boldsymbol{\Sigma})$, the application $\mathfrak{A}_{B,k,\epsilon}$ such that $\mathfrak{A}_{B,k,\epsilon}(\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}_B^{-k}$ is continuous at $\boldsymbol{\Sigma}$.

Condition C_1 in Assumption A(ϵ) ensures that the multiplicities are consistently estimated if ϵ is chosen small enough. It can be seen that C_2 is satisfied for all but a finite number of basis B. The following lemma is useful for establishing the asymptotic distribution of $\mathcal{Q}_n(\mathbf{W}_n)$ in the case $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$.

Proposition 2. *Suppose that $\mathbf{Z}_n = \mathbf{Z}_n(\boldsymbol{\mu}_0) \xrightarrow{d} \mathcal{N}_p(\boldsymbol{\Delta}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}_n \to \boldsymbol{\Sigma}$ almost surely, as $n \to \infty$. Let $B = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ be a basis of $\mathbb{R}^p$. Under Assumption A(ϵ), if $k \leq$ rank($\boldsymbol{\Sigma}$), it follows that:*

$$\mathbf{Z}_n^\top \boldsymbol{\Sigma}_{n,B}^{-k} \mathbf{Z}_n \xrightarrow{d} \chi_k^2(\theta_1^{(k)}),$$

where $\theta_1^{(k)} = \boldsymbol{\Delta}^\top \mathbf{P}_B \boldsymbol{\Lambda}^{-k} \mathbf{P}_B^\top \boldsymbol{\Delta}$, and $\mathbf{P}_B$ is the B-eigenvector matrix of $\boldsymbol{\Sigma}$.

Proof: Assumption C_2 in $\mathbf{A}(\epsilon)$ and the almost sure convergence of $\boldsymbol{\Sigma}_n$ to $\boldsymbol{\Sigma}$ give:

$$\boldsymbol{\Sigma}_{nB}^{-k} = \mathfrak{A}_{B,k,\epsilon}(\boldsymbol{\Sigma}_n) \to \boldsymbol{\Sigma}_B^{-k} = \mathfrak{A}_{B,k,\epsilon}(\boldsymbol{\Sigma}) \ a.s., \text{ as } n \to \infty.$$

The continuous mapping theorem then entails that $\mathbf{Z}_n^\top \boldsymbol{\Sigma}_{n,B}^{-k} \mathbf{Z}_n \xrightarrow{d} \mathbf{Z}^\top \boldsymbol{\Sigma}_B^{-k} \mathbf{Z}$. The Ogasawara-Takahashi theorem establishes the chi-square limiting distribution (see, e.g., Rao and Mitra (1971) or Rao (1973)): Since $\boldsymbol{\Sigma}_B^{-k}$ is a $\{2\}$-inverse, the conditions $(\boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k})^3 = (\boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k})^2$ and $\boldsymbol{\Delta}^\top \boldsymbol{\Sigma}_B^{-k} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_B^{-k} \boldsymbol{\Delta} = \boldsymbol{\Delta}^\top \boldsymbol{\Sigma}_B^{-k} \boldsymbol{\Delta}$ hold trivially. Finally, $\boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k}\boldsymbol{\Delta}$ belongs to the column space of $\boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k}\boldsymbol{\Sigma}$, since $\boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k}\boldsymbol{\Delta} = \boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k}\boldsymbol{\Sigma}\boldsymbol{\eta}$, $\boldsymbol{\eta} = \mathbf{C}_k \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_k^{-1}) \mathbf{C}_k^\top \boldsymbol{\Delta}$, where $\mathbf{P}_B = (\mathbf{C}_k \ \mathbf{C}_{p-k})$, with $\mathbf{C}_k$ and $\mathbf{C}_{p-k}$ of dimensions $p \times k$ and $p \times (p-k)$, respectively. The number of degrees of freedom is $k = \mathrm{rank}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}_B^{-k})$ with non-centrality parameter $\boldsymbol{\Delta}^\top \boldsymbol{\Sigma}_B^{-k} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_B^{-k} \boldsymbol{\Sigma} \boldsymbol{\Sigma}_B^{-k} \boldsymbol{\Delta} = \boldsymbol{\Delta}^\top \mathbf{P}_B \boldsymbol{\Lambda}^{-k} \mathbf{P}_B^\top \boldsymbol{\Delta}$. This concludes the proof. $\square$

Note that the condition $k \leq r = \mathrm{rank}(\boldsymbol{\Sigma})$ appears to be essential (see Duchesne and Francq (2008)). It follows immediately from Proposition 2 that, when $k \leq r$, the asymptotic distribution of

$\mathcal{Q}_n(\mathbf{\Sigma}_n^{-k})$ is χ_k^2 under the null. The tests based on the $\{2\}$-inverses are thus defined by the critical regions $\left\{\mathcal{Q}_n(\mathbf{\Sigma}_n^{-k}) > \chi^2_{k,1-\alpha}\right\}$, where $\chi^2_{k,\alpha}$ denotes the α-quantile of the χ_k^2 distribution.

2.3 The test statistic based on the generalized inverse (case $\mathbf{W}_n = \mathbf{\Sigma}_n^-$)

In order to test the null hypothesis, Hotelling's type procedure is based on the weighting $\mathbf{W}_n = \mathbf{\Sigma}_n^{-1}$ and thus the test statistic $\mathcal{Q}_n(\mathbf{\Sigma}_n^{-1})$, provided that $\mathbf{\Sigma}_n$ is invertible. When $\mathbf{\Sigma}$ is non-singular, $\mathbf{\Sigma}_n$ will be non-singular when n is chosen large enough. In the singular case, the weighting $\mathbf{W}_n = \mathbf{\Sigma}_n^-$ offers a natural choice, since the Moore-Penrose inverse is uniquely defined. Recall that the generalized inverse is obtained by inverting the non zero eigenvalues. In practice, as for the $\{2\}$-inverses, a tolerance $\epsilon > 0$ is required for assessing the non zero eigenvalues. The following proposition clarifies the role of ϵ and gives the asymptotic distribution of $\mathcal{Q}_n(\mathbf{\Sigma}_n^-)$.

Proposition 3. *Suppose that $\mathbf{Z}_n = \mathbf{Z}_n(\boldsymbol{\mu}_0) \xrightarrow{d} \mathcal{N}_p(\mathbf{\Delta}, \mathbf{\Sigma})$ and $\mathbf{\Sigma}_n \to \mathbf{\Sigma}$ almost surely, as $n \to \infty$. Let the spectral decomposition $\mathbf{\Sigma}_n = \mathbf{P}_n \mathbf{\Lambda}_n \mathbf{P}_n^\top$ where $\mathbf{\Lambda}_n = diag\{\lambda_1(\mathbf{\Sigma}_n), \ldots, \lambda_p(\mathbf{\Sigma}_n)\}$. For any $\epsilon > 0$, let $\mathbf{\Sigma}_{n,\epsilon} = \mathbf{P}_n \mathbf{\Lambda}_{n,\varepsilon} \mathbf{P}_n^\top$, where $\mathbf{\Lambda}_{n,\epsilon}$ is the matrix obtained by replacing by zero the elements of $\mathbf{\Lambda}_n$ which are less than ϵ. If ϵ is sufficiently small, so that:*

$$Pr\left\{rank(\mathbf{\Sigma}_{n,\varepsilon}) = rank(\mathbf{\Sigma})\right\} \to 1, \tag{6}$$

as $n \to \infty$, it follows that:

$$\mathbf{Z}_n^\top \mathbf{\Sigma}_{n,\varepsilon}^- \mathbf{Z}_n \xrightarrow{d} \chi_r^2(\theta_2),$$

where $r = rank(\mathbf{\Sigma})$ and $\theta_2 = \mathbf{\Delta}^\top \mathbf{\Sigma}^- \mathbf{\Delta}$.

Using similar arguments as in Proposition 2, Proposition 3 follows. See also the asymptotic distribution theory in Moore (1977, 1978) or Tyler (1981, Lemma 2.4). The proposition shows that, if ϵ is chosen sufficiently small, under the null hypothesis $\mathcal{Q}_n(\mathbf{\Sigma}_n^-) = \mathbf{Z}_n^\top \mathbf{\Sigma}_{n,\varepsilon}^- \mathbf{Z}_n + o_P(1) \xrightarrow{d} \chi_r^2$. Because $\mathbf{\Sigma}^- = \mathbf{\Sigma}^{-r}$, under the assumptions $\mathbf{A}(\epsilon)$ and (6), for ϵ sufficiently small we have

$$Pr\left\{\mathcal{Q}_n(\mathbf{\Sigma}_n^-) = \mathcal{Q}_n(\mathbf{\Sigma}_n^{-r})\right\} \to 1, \tag{7}$$

as $n \to \infty$. In this sense, the generalized inverse statistic can be considered as a particular $\{2\}$-inverse statistic. Note that in the case where $\mathbf{\Sigma}_n^-$ is only supposed to be a $\{1\}$-inverse, additional hypotheses are required in order to have the stated result (in particular $\mathbf{\Delta}$ must be in the column space of $\mathbf{\Sigma}$). In view of Theorem 9.2.3 of Rao and Mitra (1971, p. 173), the conclusion of Proposition 3 is also true when the estimator $\mathbf{\Sigma}_n^-$ is a given symmetric reflexive g-inverse of the matrix $\mathbf{\Sigma}_n$ and when $\mathbf{\Sigma}$ is a continuity point of the application which gives this particular symmetric reflexive g-inverse.

The test statistic $\mathcal{Q}_n(\mathbf{\Sigma}_n^-)$ has been considered in a time series context for testing null autocorrelations in time series analysis. In fact, Li (1992) investigated the use of $\mathcal{Q}_n(\mathbf{\Sigma}_n^{-1})$ in non-linear time series analysis. However, in the time series framework, it seems difficult to formulate precise conditions which guarantee the invertibility of the asymptotic covariance matrix $\mathbf{\Sigma}$. For example, the non-linear time series model of Li (1992) rules out linear models such as the ARMA models, in which case it is known that the asymptotic covariance matrix of the residual autocorrelations is essentially singular. From the simulation results in Duchesne and Francq (2008), to invert an approximately singular covariance matrix may result in empirical levels far from the nominal levels. In order to have a test statistic well-defined in linear and non-linear models, Duchesne and Francq (2008) investigated the use of $\mathcal{Q}_n(\mathbf{\Sigma}_n^-)$, with $\mathbf{\Sigma}_n^-$ the Moore-Penrose inverse of a certain estimator of the asymptotic covariance matrix of the residual autocorrelations.

3. POWER UNDER FIXED ALTERNATIVES

We now examine the asymptotic powers under fixed alternatives, adopting the approach of Bahadur (1960). In this approach, the efficiency of a test statistic is measured by its slope, defined as the rate of convergence of its p-value under a fixed alternative hypothesis $H_1 : \boldsymbol{\mu} = \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_0$. Using the notation in (3), let

$$S^I(t) = \mathbb{P}\left(\sum_{i=1}^{r} \lambda_i N_i^2 > t\right), \; S^{-k}(t) = \mathbb{P}\left(\chi_k^2 > t\right), \text{ and } S^-(t) = \mathbb{P}\left(\chi^2_{\mathrm{rank}(\mathbf{\Sigma})} > t\right)$$

be the respective asymptotic survival functions of the test statistics $\mathcal{Q}_n(\mathbf{I}_p)$, $\mathcal{Q}_n(\mathbf{\Sigma}_n^{-k})$ and $\mathcal{Q}_n(\mathbf{\Sigma}_n^-)$ under the null hypothesis H_0. Denote by $\mathcal{V}(\{\lambda_{i_1}, \ldots, \lambda_{i_j}\})$ the linear vector space generated by the columns $i_1, \ldots, i_j$ of $\mathbf{P} = \mathbf{P}_B$. When, with obvious conventions, $\lambda_{i_1-1} < \lambda_{i_1} \leq \cdots \leq \lambda_{i_j} < \lambda_{i_j+1}$ then $\mathcal{V}(\{\lambda_{i_1}, \ldots, \lambda_{i_j}\})$ denotes the eigenvector space associated to the eigenvalues $\lambda_{i_1}, \ldots, \lambda_{i_j}$ of $\mathbf{\Sigma}$.

Proposition 4. *Let a basis B and a tolerance ϵ such that $A(\epsilon)$ and (6) hold true, and let $k \leq$ rank$(\mathbf{\Sigma})$. Under the alternative $H_1 : \boldsymbol{\mu} = \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_0$, the (approximate) Bahadur slopes of the test procedure (2) with the weightings $\mathbf{W}_n = \mathbf{I}_p$, $\mathbf{W}_n = \mathbf{\Sigma}_n^{-k}$ and $\mathbf{W}_n = \mathbf{\Sigma}_n^-$ are given by:*

$$c^I = \lim_{n\to\infty} -\frac{2}{n} \log S^I\left\{\mathcal{Q}_n\left(\mathbf{I}_p\right)\right\} = \frac{\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0\|^2}{\lambda_1}, \tag{8}$$

$$c^{-k} = \lim_{n\to\infty} -\frac{2}{n} \log S^{-k}\left\{\mathcal{Q}_n\left(\mathbf{\Sigma}_n^{-k}\right)\right\} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^\top \mathbf{\Sigma}^{-k}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0), \tag{9}$$

$$c^- = \lim_{n\to\infty} -\frac{2}{n} \log S^-\left\{\mathcal{Q}_n\left(\mathbf{\Sigma}_n^-\right)\right\} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^\top \mathbf{\Sigma}^-(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0), \tag{10}$$

respectively, where the convergence in probability (9) and (10) hold, respectively, when:

$$\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \notin \mathcal{V}(\{\lambda_{k+1}, \ldots, \lambda_r, 0\}), \tag{11}$$

$$\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \notin \mathcal{V}(\{0\}). \tag{12}$$

The convergence in probability in formula (8) stands without further restriction.

Note that, even when the variance is known, *i.e.* when $\mathbf{\Sigma}_n = \mathbf{\Sigma}$, the test statistic based on the $\{2\}$-inverse (resp. the generalized inverse) is not consistent when (11) (resp. (12)) does not hold true. Indeed, when $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \in \mathcal{V}(\{\lambda_{k+1}, \ldots, \lambda_r, 0\})$ we have $\mathbf{\Sigma}^{-k}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) = \mathbf{0}_p$, and it follows that:

$$\begin{aligned} \mathcal{Q}_n\left(\mathbf{\Sigma}^{-k}\right) &= \mathbf{Z}_n^\top(\boldsymbol{\mu}_0)\mathbf{\Sigma}^{-k}\mathbf{Z}_n(\boldsymbol{\mu}_0), \\ &= \left\{\mathbf{Z}_n(\boldsymbol{\mu}_1) + n^{1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\right\}^\top \mathbf{\Sigma}^{-k} \left\{\mathbf{Z}_n(\boldsymbol{\mu}_1) + n^{1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\right\}, \\ &= \mathbf{Z}_n^\top(\boldsymbol{\mu}_1)\mathbf{\Sigma}^{-k}\mathbf{Z}_n(\boldsymbol{\mu}_1) \not\to \infty \text{ in probability,} \end{aligned}$$

as $n \to \infty$. Similarly, when $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \in \mathcal{V}(\{0\})$ the statistic $\mathcal{Q}_n\left(\mathbf{\Sigma}^-\right) = \mathbf{Z}_n^\top(\boldsymbol{\mu}_1)\mathbf{\Sigma}^-\mathbf{Z}_n(\boldsymbol{\mu}_1)$ does not diverge under the alternative hypothesis. Note also that, in view of the right-hand sides of (9) and

(10), the previous derivations show that one can set, by continuity, $c^{-k} = 0$ when (11) is not satisfied, and $c^{-} = 0$ when the relation (12) does not hold. The following corollary presents a comprehensive comparison of the Bahadur slopes of the test statistics $\mathcal{Q}_n(\mathbf{I}_p)$, $\mathcal{Q}_n\left(\mathbf{\Sigma}_n^{-k}\right)$ and $\mathcal{Q}_n\left(\mathbf{\Sigma}_n^{-}\right)$.

Corollary 1. *Under the assumptions of Proposition 4, the following comparisons can be made:*

*i) The Imhof-based test is always consistent (*i.e. *we always have $c^I > 0$);*

*ii) The test based on the $\{2\}$-inverse $\mathbf{\Sigma}_n^{-k}$ is consistent (*i.e. *$c^{-k} > 0$) if and only if (11);*

*iii) The generalized-inverse based test is consistent (*i.e. *$c^{-} > 0$) if and only if (12);*

iv) For all $k \leq r :=$ rank$(\mathbf{\Sigma})$ we have $c^{-} = c^{-r} \geq c^{-k} \geq c^{-k-1} \geq \cdots \geq c^{-1}$ with $c^{-k} = c^{-k-1}$ iff $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \in \mathcal{V}(\{\lambda_k\}^c)$;

v) When $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \in \mathcal{V}(\{\lambda_1, \ldots, \lambda_k\})$ with $k > 1$ we have $c^{-k} \geq c^I$, with strict inequality iff there exists k' such that $1 < k' \leq k$, $\lambda_{k'} < \lambda_1$ and $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \notin \mathcal{V}\left(\{\lambda_{k'}\}^c\right)$;

vi) When $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \notin \mathcal{V}(\{\lambda_1, \ldots, \lambda_k\})$ we have $c^I > c^{-k}$;

vii) The Bahadur slope of Imhof-based test statistic is always larger than the one of the test statistic based on $\mathbf{\Sigma}^{-1}$, that is $c^I \geq c^{-1}$, with equality iff $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \in \mathcal{V}(\{\lambda_1\})$.

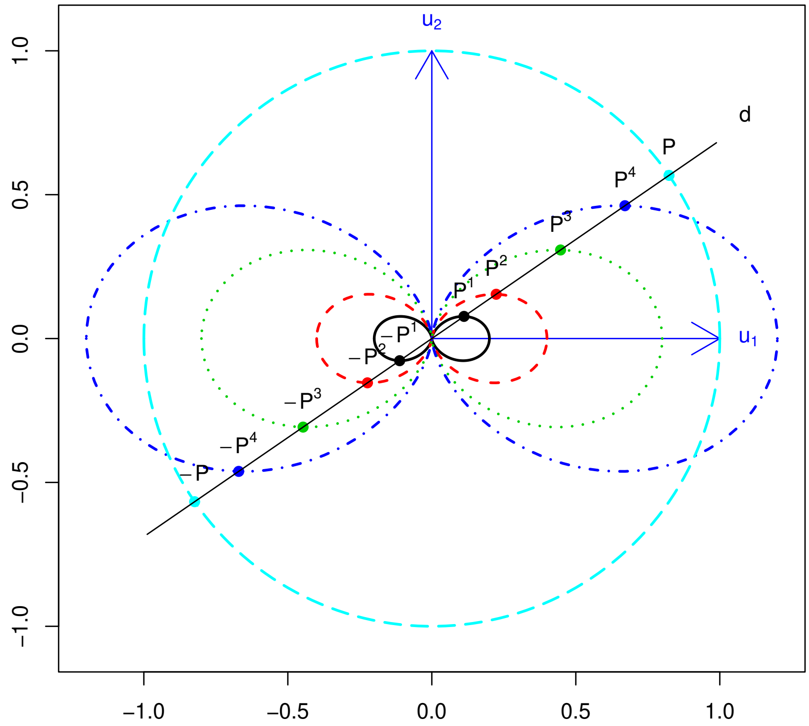
The most noticeable result of this corollary is that, contrary to the other test procedures, the Imhof-based test offers a strictly positive Bahadur slope for all $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \neq \mathbf{0}_p$. From that point of view, $\mathcal{Q}_n(\mathbf{I}_p)$ represents the only omnibus test statistic with non trivial power under all alternative hypotheses, and is in the spirit of the so-called portmanteau test statistics in the time series literature. However, for an alternative hypothesis in the non-zero eigenspace of $\mathbf{\Sigma}$, the slope of the Imhof-based test is smaller than that of the test based on the generalized inverse $\mathbf{\Sigma}_n^{-}$. Note also that, in term of the Bahadur slope, the $\{2\}$-inverse test statistic based on $\mathbf{\Sigma}_n^{-k}$ dominates the test statistic based on $\mathbf{\Sigma}_n^{-k'}$ when $k > k'$.

Figure 1 displays the Bahadur slopes of the different tests when $\mathbf{\Sigma} = \text{diag}(1, 1, 1/2, 1/2, 0, 0)$ and when $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ is a unit vector with direction d in the plane containing $\mathbf{u}_1 = 1/\sqrt{5}\,(1, 1, 1, 1, 0, 1)^\top$ and $\mathbf{u}_2 = (0, 0, 0, 0, 1, 0)^\top$. The length of the vector going from the origin to the curve P^k in the direction d gives the Bahadur slope of the test statistic $\mathcal{Q}_n(\mathbf{\Sigma}_n^{-k})$. In this example, the Bahadur slope of $\mathcal{Q}_n(\mathbf{I}_6)$ is always one, since $\lambda_1 = 1$ and the length of $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ is normalized to one. Since $\mathbf{u}_2 \in \mathcal{V}(\{0\})$, only the Imhof-based test is powerful for alternatives in the direction of $\mathbf{u}_2$ and the slopes of the other tests cancel for that alternative hypothesis. This figure thus illustrates the points i)-iv) of Corollary 1. Figure 2 illustrates other points. In particular, in this figure, $c^{-3} = c^{-}$ because the direction d of the alternative belongs to $\mathcal{V}(\{\lambda_4\}^c)$ and $c^{-3} \geq c^I$ because d belongs to $\mathcal{V}(\{\lambda_1, \lambda_2, \lambda_3\})$.

Proof of Proposition 4: Under the null hypothesis, $n^{-1/2}\mathbf{Z}_n = n^{-1/2}\mathbf{Z}_n(\boldsymbol{\mu}_0) \xrightarrow{p} \mathbf{0}$ and under the alternative $H_1 : \boldsymbol{\mu} = \boldsymbol{\mu}_1$ we have $n^{-1/2}\mathbf{Z}_n = n^{-1/2}\{\mathbf{Z}_n(\boldsymbol{\mu}_1) + n^{1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\} \xrightarrow{p} \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$. A large deviation result yields:

$$\log P\left(\sum_{i=1}^{r} \lambda_i N_i^2 > x\right) \sim \frac{-x}{2\lambda_1},$$

FIGURE 1. The Bahadur slopes c^{-k} of $\mathcal{Q}_n(\boldsymbol{\Sigma}_n^{-k})$ and c^I of $\mathcal{Q}_n(\mathbf{I}_p)$ when $\boldsymbol{\Sigma} = \mathrm{diag}(1,1,1/2,1/2,0,0)$, for alternatives in the direction $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 = \mathrm{d} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$, $||\mathrm{d}|| = 1$, where $\mathbf{u}_1 = 1/\sqrt{5}\,(1,1,1,1,0,1)^\top$ and $\mathbf{u}_2 = (0,0,0,0,1,0)^\top$. The Bahadur slope c^{-k} corresponds to the length of the vector going from the origin to the curve P^k in the direction d. The slope of the Imhof-based test describes a circle P because it is constant.



as $x \to \infty$, where $N_1, \dots, N_r$ are independent $\mathcal{N}(0,1)$ random variables (see Zolotarev (1961)). The Bahadur slope of the first test is thus given by:

$$c^I = \lim_{n\to\infty} -\frac{2}{n} \log S^I\left(\|\mathbf{Z}_n\|^2\right) = \lim_{n\to\infty} \frac{||\mathbf{Z}_n||^2}{n\lambda_1} = \frac{||\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0||^2}{\lambda_1}.$$

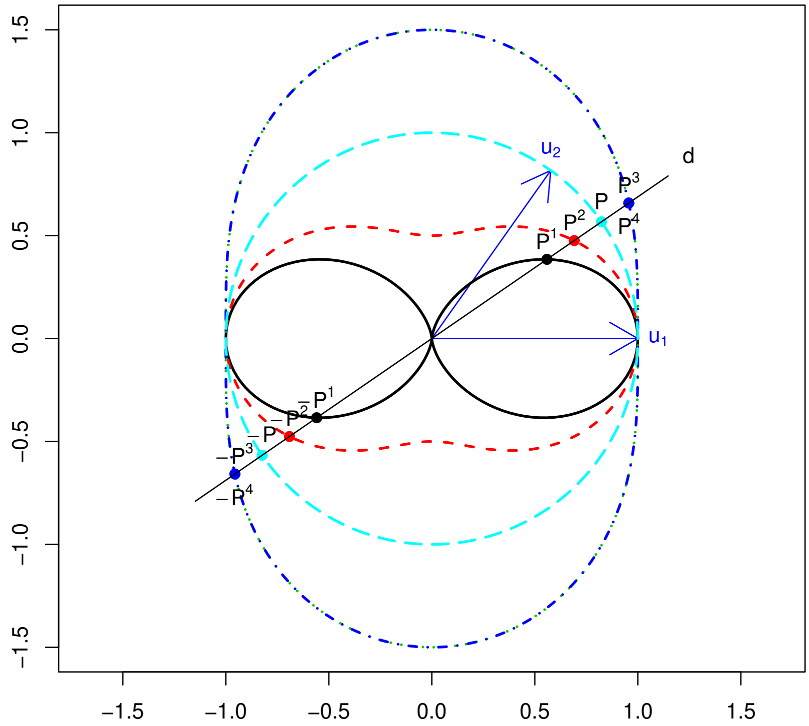
Now note that

$$\lim_{n\to\infty} \mathcal{Q}_n\left(\boldsymbol{\Sigma}_n^{-k}\right) = \lim_{n\to\infty} n(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^\top \boldsymbol{\Sigma}^{-k}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) = \infty \text{ in probability}$$

under the condition (11), and that $\mathcal{Q}_n\left(\boldsymbol{\Sigma}_n^{-}\right) \to \infty$ in probability under the condition (12). Using the large deviation result $\log P(\chi^2_k > x) \sim -x/2$ as $x \to \infty$, under the condition (11) (resp. (12)) the Bahadur slope c^{-k} (resp. c^{-}) is then obtained by the arguments used to compute c^I. □

Proof of Corollary 1: Points $i) - iii)$ are direct consequences of Proposition 4. To show $iv)$, consider the spectral decomposition $\boldsymbol{\Sigma} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^\top$ where $\mathbf{P}^\top\mathbf{P} = \mathbf{I}_p$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_p)$, and

FIGURE 2. The Bahadur slopes c^{-k} of $\mathcal{Q}_n(\mathbf{\Sigma}_n^{-k})$ and c^I of $\mathcal{Q}_n(\mathbf{I}_p)$ when $\mathbf{\Sigma} = \text{diag}(1, 1, 1/2, 1/2, 0, 0)$, for alternatives in the direction $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 = \mathrm{d} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$, $||\mathrm{d}|| = 1$, where $\mathbf{u}_1 \in \mathcal{V}(\{\lambda_1\})$ and $\mathbf{u}_2 = 1/\sqrt{3}(1, 1, 1, 0, 0, 0)^\top \in \mathcal{V}(\{\lambda_1, \lambda_2, \lambda_3\})$. The Bahadur slope c^{-k} corresponds to the length of the vector from the origin to the curve P^k in the direction d. The slope of the Imhof-based test describes a circle P because it is constant.



note that for $1 \leq k < k' \leq r$ the difference $c^{-} - c^{-k}$ is non-negative:

$$c^{-} - c^{-k} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^\top \mathbf{P}\text{diag}\left(\mathbf{0}_k^\top, \lambda_{k+1}^{-1}, \ldots, \lambda_{k'}^{-1}, \mathbf{0}_{p-r}^\top\right)\mathbf{P}^\top(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) \geq 0,$$

with equality if and only if $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0 \in \mathcal{V}(\{\lambda_{k+1}, \ldots, \lambda_{k'}\}^c)$. Points $v) - vii)$ come from

$$c^{-k} - c^I = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^\top \mathbf{P}\text{diag}\left(0, \lambda_2^{-1} - \lambda_1^{-1}, \ldots, \lambda_k^{-1} - \lambda_1^{-1}, -\lambda_1^{-1}, \ldots, -\lambda_1^{-1}\right)\mathbf{P}^\top(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0).$$

This shows the result. □

Other comparisons between the test statistics can be performed under local alternatives. In the next section, the so-called Pitman's approach is investigated.

4. POWER UNDER LOCAL ALTERNATIVES

Consider a sequence of local alternatives of the form $H_{1n} : \boldsymbol{\mu}_n = \boldsymbol{\mu}_0 + n^{-1/2}\boldsymbol{\Delta}_n$, where $\boldsymbol{\Delta}_n \to \boldsymbol{\Delta} \neq \mathbf{0}_p$. The following proposition gives the Asymptotic Local Power (ALP) of the test procedures.

Proposition 5. *When $\mathbf{Z}_n(\boldsymbol{\mu}_0) \stackrel{d}{\to} \mathcal{N}_p(\boldsymbol{\Delta}, \boldsymbol{\Sigma})$, and with the notations and assumptions of Propositions 1-3, the ALP of the test procedure (2) with the weighting $\mathbf{W}_n = \mathbf{I}_p$, $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$ and $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-}$ under the local alternatives H_{1n} are given by:*

$$\begin{aligned} \mathrm{ALP}^{I}(\boldsymbol{\Delta}) &= Pr\left(\sum_{i=1}^{r} \lambda_i \chi_1^2(\theta_i^{*2}) + \boldsymbol{\theta}_0^\top \boldsymbol{\theta}_0 > c_\alpha(\lambda_1, \ldots, \lambda_r)\right), && (13)\\ \mathrm{ALP}^{-k}(\boldsymbol{\Delta}) &= Pr\left(\chi_k^2(\theta_1^{(k)}) > \chi_{k,1-\alpha}^2\right), && (14)\\ \mathrm{ALP}^{-}(\boldsymbol{\Delta}) &= Pr\left(\chi_r^2(\theta_2) > \chi_{r,1-\alpha}^2\right). && (15) \end{aligned}$$

The proof of Proposition 5 represents a direct consequence of Propositions 1-3 and therefore it is omitted. The following corollary compares the ALP of the different test procedures.

Corollary 2. *Under the assumption of Proposition 5:*

*i) The Imhof-based test is always locally asymptotically powerful (*i.e. $\mathrm{ALP}^{I}(\boldsymbol{\Delta}) > \alpha\ \forall \boldsymbol{\Delta}$*);*

*ii) The test based on the $\{2\}$-inverse $\boldsymbol{\Sigma}_n^{-k}$ is locally asymptotically powerful (*i.e. $\mathrm{ALP}^{-k}(\boldsymbol{\Delta}) > \alpha$*) if and only if $\boldsymbol{\Delta} \notin \mathcal{V}(\{\lambda_{k+1}, \ldots, \lambda_r, 0\})$;*

*iii) The generalized-inverse based test is locally asymptotically powerful (*i.e. $\mathrm{ALP}^{-}(\boldsymbol{\Delta}) > \alpha$*) if and only if $\boldsymbol{\Delta} \notin \mathcal{V}(\{0\})$;*

iv) When $\boldsymbol{\Delta} \in \mathcal{V}(\{\lambda_1, \ldots, \lambda_k\})$ with $k < r$ then $\mathrm{ALP}^{-k}(\boldsymbol{\Delta}) > \mathrm{ALP}^{-}(\boldsymbol{\Delta})$;

v) When $\boldsymbol{\Delta} \in \mathcal{V}(\{\lambda_1\})$ we have $\mathrm{ALP}^{-1}(\boldsymbol{\Delta}) > \mathrm{ALP}^{I}(\boldsymbol{\Delta})$;

vi) When $\boldsymbol{\Delta} \in \mathcal{V}(\{\lambda_1\})$ we have $\mathrm{ALP}^{-1}(\boldsymbol{\Delta}) \geq \mathrm{ALP}^{-2}(\boldsymbol{\Delta}) \geq \cdots \geq \mathrm{ALP}^{-r-1}(\boldsymbol{\Delta}) \geq \mathrm{ALP}^{-}(\boldsymbol{\Delta})$.

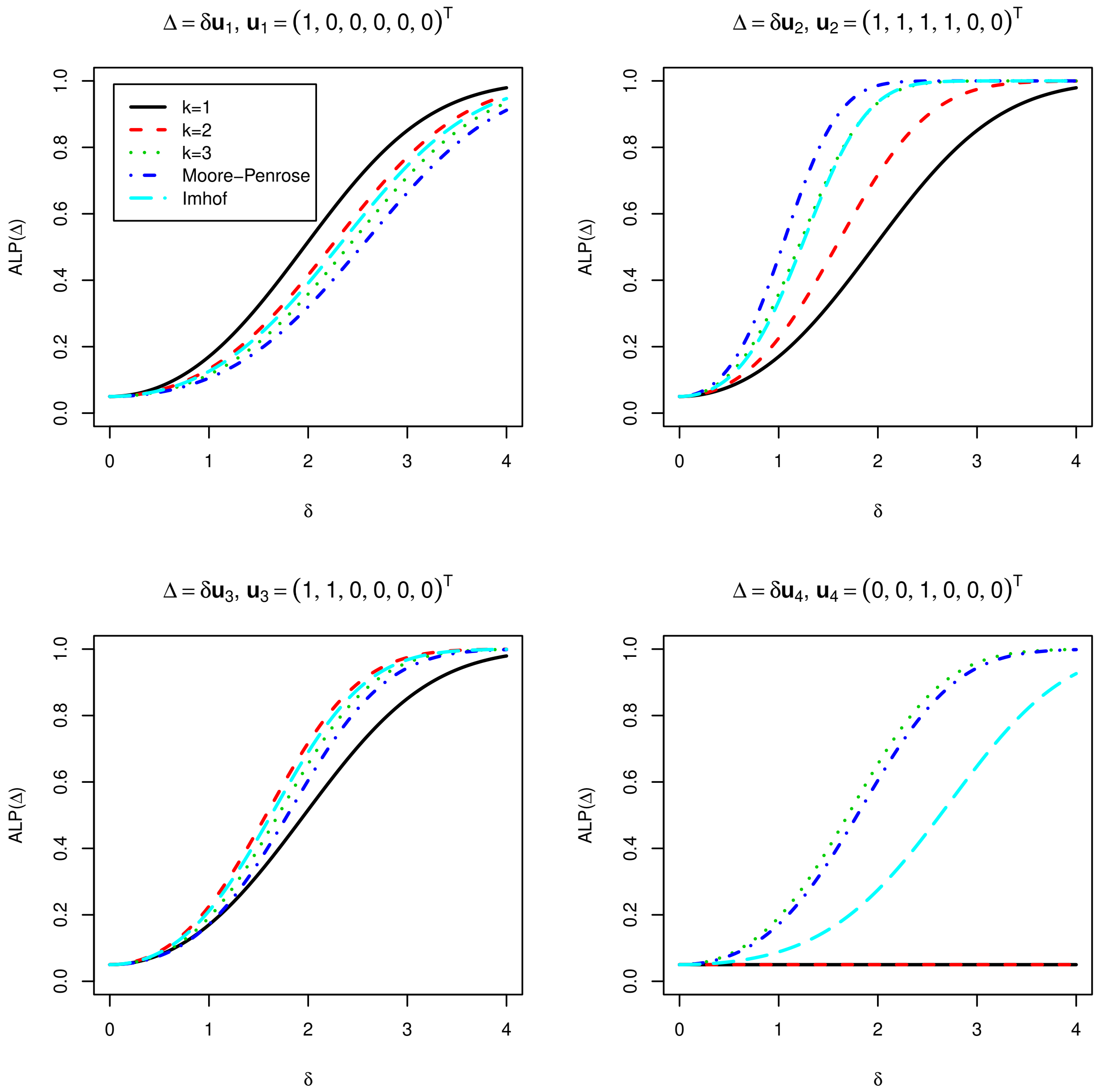
This corollary shows that, as for the Bahadur slopes, the ranking of the local asymptotic powers of the different tests depends on the position of the alternative with respect to the eigenspaces of $\boldsymbol{\Sigma}$. However, compared to Bahadur's approach, Pitman's approach highlights the relative merits of the test procedures with a different viewpoint. In particular, in term of ALP, the performance of the test statistics based of the $\{2\}$-inverse does not necessarily increases with k (compare *iv*) in Corollary 1 with *vi*) in Corollary 2).

Figure 3 displays the ALP's for several directions $\boldsymbol{\Delta}$ of the local alternative, and for the same matrix $\boldsymbol{\Sigma}$ as that used in Figures 1 and 2. From this figure, writing $\mathrm{ALP}^{-k} \equiv \mathrm{ALP}^{-k}(\boldsymbol{\Delta})$, the following relations are satisfied:

$$\begin{aligned} &\mathrm{ALP}^{-1} > \mathrm{ALP}^{-2} > \mathrm{ALP}^{I} > \mathrm{ALP}^{-3} > \mathrm{ALP}^{-}, && \text{when } \boldsymbol{\Delta} \propto (1,0,0,0,0,0)^\top,\\ &\mathrm{ALP}^{-} > \mathrm{ALP}^{-3} \simeq \mathrm{ALP}^{I} > \mathrm{ALP}^{-2} > \mathrm{ALP}^{-1}, && \text{when } \boldsymbol{\Delta} \propto (1,1,1,1,0,0)^\top,\\ &\mathrm{ALP}^{-2} > \mathrm{ALP}^{I} > \mathrm{ALP}^{-3} > \mathrm{ALP}^{-} > \mathrm{ALP}^{-1}, && \text{when } \boldsymbol{\Delta} \propto (1,1,0,0,0,0)^\top,\\ &\mathrm{ALP}^{-3} > \mathrm{ALP}^{-} > \mathrm{ALP}^{I} > \mathrm{ALP}^{-2} = \mathrm{ALP}^{-1} \equiv \alpha, && \text{when } \boldsymbol{\Delta} \propto (0,0,1,0,0,0)^\top. \end{aligned}$$

To summarize, in term of ALP, the test based on $\boldsymbol{\Sigma}^{-k}$ is very powerful for alternatives close to the direction of the first k eigenvectors of $\boldsymbol{\Sigma}$, but may be completely powerless for orthogonal alternatives.

FIGURE 3. Asymptotic local powers of the tests when $\mathbf{\Sigma} = \mathrm{diag}(1, 1, 1/2, 1/2, 0, 0)$, for alternatives in different directions.



The Imhof-based test statistic and the one relying on the generalized inverse offer power for more alternatives, but it appears that none test is dominated by another one.

Proof of Corollary 2: To show $i)$ we note that $\|\boldsymbol{\theta}^*\|^2 = \mathbf{\Delta}^\top \mathbf{\Sigma} \mathbf{\Delta} = 0$ iff $\mathbf{\Delta} \in \mathcal{V}(\{0\})$, and that $\|\boldsymbol{\theta}_0\|^2 = 0$ iff $\mathbf{\Delta} \in \mathcal{V}(\{\lambda_1, \ldots, \lambda_r\})$. The points $ii)$ and $iii)$ are obtained similarly. To show $iv)$, first note that

$$\theta_2 - \theta_1^{(k)} = \mathbf{\Delta}^\top \mathbf{P}_B \mathrm{diag}\left(\mathbf{0}_k^\top, \lambda_{k+1}^{-1}, \ldots, \lambda_r^{-1}, \mathbf{0}_{p-r}^\top\right) \mathbf{P}_B^\top \mathbf{\Delta} \geq 0,$$

with equality iff $\mathbf{\Delta} \in \mathcal{V}(\{\lambda_{k+1}, \ldots, \lambda_r\}^c)$. Thus $\theta_2 = \theta_1^{(k)} > 0$ when $\mathbf{\Delta} \in \mathcal{V}(\{\lambda_1, \ldots, \lambda_k\})$. Now $iv)$ is implied by the fact that $P\left(\chi_k^2(\theta) > \chi_{k,1-\alpha}^2\right)$ strictly decreases with k for all $\theta > 0$ and all $\alpha \in (0,1)$ (see Theorem 2 in Ghosh (1973)).

In order to show $v)$ and $vi)$, we adapt arguments called upon by Ghosh (1973). Let X_1 be a $\chi_k^2(\nu)$-distributed random variable, and let X_2 be a random variable independent of X_1. Consider the testing problem $H_0 : \nu = 0$ against $H_0 : \nu > 0$ based on the observations (X_1, X_2). Assume that the distribution of X_2 is the same under the null and alternative hypotheses. Using the Neyman-Pearson lemma, the most powerful test statistic of $H_0 : \nu = 0$ against $H_0 : \nu = \nu_1$, with $\nu_1 > 0$, rejects the null hypothesis if the likelihood ratio is large. A straightforward but tedious computation shows that this likelihood ratio is given by:

$$\frac{L(X_1, X_2; \nu_1)}{L(X_1, X_2; 0)} = \Gamma(k/2)\exp(-\nu_1/2)\sum_{i=0}^{\infty}(\nu_1 X_1)^i/\{4^i i!\Gamma(i + k/2)\}.$$

Note that this ratio is an increasing function of X_1. Consequently the critical region of the uniformly most powerful (UMP) test is given by $\{X_1 \geq \chi_{k,1-\alpha}^2\}$. We now apply this result when $\mathbf{\Delta} \in \mathcal{V}(\{\lambda_1\})$, setting $\nu = \theta_1^{*2}$, $X_1 = (N_1 + \nu)^2$ and $X_2 = \sum_{i=2}^r \lambda_i/\lambda_1 N_i^2$, with the notations of Proposition 1 and $(N_1, \ldots, N_r)^\top \sim \mathcal{N}_r(\mathbf{0}, \mathbf{I}_r)$. Noting also that $\theta_1^{(1)} = \theta_1^{*2}$ when $\mathbf{\Delta} \in \mathcal{V}(\{\lambda_1\})$, we then obtain:

$$\begin{aligned}
\mathrm{ALP}^{-1}(\mathbf{\Delta}) &= P\left\{(N_1 + \nu)^2 > \chi_{1,1-\alpha}^2\right\}, \\
&> P\left\{(N_1 + \nu)^2 + \sum_{i=2}^{r} \lambda_i/\lambda_1 N_i^2 > \frac{c_\alpha(\lambda_1, \ldots, \lambda_r)}{\lambda_1}\right\} = \mathrm{ALP}^I(\mathbf{\Delta}),
\end{aligned}$$

which shows $v)$. Setting $X_1 = (N_1 + \nu)^2 + \sum_{i=2}^k N_i^2$ and $X_2 = N_{k+1}^2$, the same argument entails:

$$\begin{aligned}
\mathrm{ALP}^{-k}(\mathbf{\Delta}) &= P\left\{(N_1 + \nu)^2 + \sum_{i=2}^{k} N_i^2 > \chi_{k,1-\alpha}^2\right\}, \\
&> P\left\{(N_1 + \nu)^2 + \sum_{i=2}^{k+1} N_i^2 > \chi_{k+1,1-\alpha}^2\right\} = \mathrm{ALP}^{-k+1}(\mathbf{\Delta}),
\end{aligned}$$

and point $vi)$ follows. □

In the next section, further comparisons and undertaken using Monte Carlo experiments.

5. SIMULATION EXPERIMENTS

In the previous sections, we have presented the asymptotic null distributions of three classes of test statistics, and we have given some asymptotic properties under fixed and local alternatives. It is natural to inquire about their finite sample properties, in particular their exact levels and powers. Furthermore, the theoretical results obtained in Sections 3 and 4 need to be completed empirically. The power comparisons between the weighting $\mathbf{W}_n = \mathbf{\Sigma}_n^-$ and $\mathbf{W}_n = \mathbf{I}_p$ seem also of particular interest. To partially answer these considerations, some Monte Carlo experiments were conducted. The main computer code for the experiments described below has been written using the R language, and Imhof's (1961) algorithm has been implemented in the FORTRAN 90 language.

5.1 Description of the simulation experiments

In order to compare the test statistics, we considered multivariate sampling from the multivariate normal distribution, for several choices of the covariance matrix $\mathbf{\Sigma}$. The test statistics included in our simulation experiments used the weighting $\mathbf{W}_n = \mathbf{I}_p$, $\mathbf{W}_n = \mathbf{\Sigma}_n^{-k}$, $k \leq r$, $r = \mathrm{rank}(\mathbf{\Sigma})$, and $\mathbf{W}_n = \mathbf{\Sigma}_n^{-}$, where $\mathbf{\Sigma}_n = \mathbf{S}_n$ represents the sample covariance matrix. For each random sample of size $n = 100$, we examined the empirical frequencies of rejection of the null hypothesis $H_0 : \boldsymbol{\mu} = \mathbf{0}$ when the latter was true by using test statistics with three nominal levels (1, 5 and 10%). Multivariate sampling appears particularly convenient to study the power of the test procedures, given the analytical results demonstrated in the previous sections. Several fixed alternatives have been included in the study, which have been chosen by examining the spectral decomposition of the covariance matrix $\mathbf{\Sigma}$ and their associated eigenspaces.

TABLE 1. Mean vectors $\boldsymbol{\mu}_{ij}^{(k)}$ in multivariate sampling from the normal distribution $\mathcal{N}_3(\boldsymbol{\mu}_{ij}^{(k)}, \mathbf{\Sigma}_{ij}^{(k)})$, $i = I, II$, $j = 1, 2$, $k = 1, \ldots, 4$.

Experiment $i = I$

$$\boldsymbol{\mu}_{I,1}^{(1)} = (\tfrac{3}{10}, 0, 0)^\top, \quad \boldsymbol{\mu}_{I,1}^{(2)} = (\tfrac{3}{10}, \tfrac{3}{10}, 0)^\top, \quad \boldsymbol{\mu}_{I,1}^{(3)} = (\tfrac{3}{10}, \tfrac{3}{10}, \tfrac{1}{10})^\top, \quad \boldsymbol{\mu}_{I,1}^{(4)} = (\tfrac{3}{10}, 0, \tfrac{1}{10})^\top,$$
$$\boldsymbol{\mu}_{I,2}^{(1)} = (\tfrac{4}{5}, 0, 0)^\top, \quad \boldsymbol{\mu}_{I,2}^{(2)} = (\tfrac{3}{10}, \tfrac{3}{10}, 0)^\top, \quad \boldsymbol{\mu}_{I,2}^{(3)} = (\tfrac{3}{10}, \tfrac{3}{10}, \tfrac{3}{10})^\top, \quad \boldsymbol{\mu}_{I,2}^{(4)} = (\tfrac{4}{5}, 0, \tfrac{3}{10})^\top;$$

Experiment $i = II$

$$\boldsymbol{\mu}_{II,1}^{(1)} = (\tfrac{3}{10}, 0, 0)^\top, \quad \boldsymbol{\mu}_{II,1}^{(2)} = (0, -\tfrac{2}{25}, \tfrac{2}{5})^\top, \quad \boldsymbol{\mu}_{II,1}^{(3)} = (\tfrac{1}{10}, -\tfrac{1}{5}, \tfrac{1}{50})^\top, \quad \boldsymbol{\mu}_{II,1}^{(4)} = (\tfrac{1}{10}, -\tfrac{3}{10}, \tfrac{1}{5})^\top,$$
$$\boldsymbol{\mu}_{II,2}^{(1)} = (-\tfrac{1}{10}, -\tfrac{1}{10}, \tfrac{1}{5})^\top, \quad \boldsymbol{\mu}_{II,2}^{(2)} = (\tfrac{1}{10}, -\tfrac{1}{10}, 0)^\top, \quad \boldsymbol{\mu}_{II,2}^{(3)} = (\tfrac{1}{10}, \tfrac{1}{10}, \tfrac{1}{10})^\top, \quad \boldsymbol{\mu}_{II,2}^{(4)} = (\tfrac{1}{10}, -\tfrac{1}{10}, \tfrac{1}{10})^\top;$$

TABLE 2. Mean vectors $\boldsymbol{\mu}_{ij}^{(k)}$ in multivariate sampling from the normal distribution $\mathcal{N}_6(\boldsymbol{\mu}_{ij}^{(k)}, \mathbf{\Sigma}_{ij}^{(k)})$, $i = III, IV$, $j = 1, 2, 3$, $k = 1, \ldots, 4$.

Experiment $i = III$

$$\boldsymbol{\mu}_{III,1}^{(1)} = (\tfrac{3}{10}, 0, 0, 0, 0, 0)^\top, \quad \boldsymbol{\mu}_{III,1}^{(2)} = (\tfrac{3}{10}, \tfrac{3}{10}, 0, 0, 0, 0)^\top,$$
$$\boldsymbol{\mu}_{III,1}^{(3)} = (\tfrac{3}{10}, \tfrac{3}{10}, \tfrac{3}{10}, 0, 0, 0)^\top, \quad \boldsymbol{\mu}_{III,1}^{(4)} = (\tfrac{3}{10}, \tfrac{3}{10}, \tfrac{1}{10}, \tfrac{1}{10}, 0, 0)^\top,$$
$$\boldsymbol{\mu}_{III,2}^{(1)} = (\tfrac{4}{5}, 0, 0, 0, 0, 0)^\top, \quad \boldsymbol{\mu}_{III,2}^{(2)} = (\tfrac{4}{5}, \tfrac{4}{5}, 0, 0, 0, 0)^\top,$$
$$\boldsymbol{\mu}_{III,2}^{(3)} = (\tfrac{4}{5}, \tfrac{4}{5}, \tfrac{2}{5}, 0, 0, 0)^\top, \quad \boldsymbol{\mu}_{III,2}^{(4)} = (\tfrac{3}{5}, \tfrac{3}{5}, \tfrac{3}{10}, \tfrac{3}{10}, 0, 0)^\top,$$
$$\boldsymbol{\mu}_{III,3}^{(1)} = (\tfrac{4}{5}, 0, 0, 0, 0, 0)^\top, \quad \boldsymbol{\mu}_{III,3}^{(2)} = (\tfrac{2}{5}, \tfrac{4}{5}, 0, 0, 0, 0)^\top,$$
$$\boldsymbol{\mu}_{III,3}^{(3)} = (0, 0, \tfrac{3}{5}, 0, 0, 0)^\top, \quad \boldsymbol{\mu}_{III,3}^{(4)} = (\tfrac{3}{5}, \tfrac{3}{5}, \tfrac{3}{10}, \tfrac{3}{10}, 0, 0)^\top;$$

Experiment $i = IV$

$$\boldsymbol{\mu}_{IV,1}^{(1)} = (0, \tfrac{3}{10}, -\tfrac{1}{5}, -\tfrac{3}{10}, 0, \tfrac{1}{5})^\top, \quad \boldsymbol{\mu}_{IV,1}^{(2)} = (0, -\tfrac{1}{10}, -\tfrac{1}{20}, \tfrac{3}{100}, \tfrac{1}{2}, \tfrac{1}{20})^\top,$$
$$\boldsymbol{\mu}_{IV,1}^{(3)} = (0.37, -0.24, -0.18, -0.12, -0.06, -0.03)^\top, \quad \boldsymbol{\mu}_{IV,1}^{(4)} = (0.22, 0.18, 0.14, 0.09, 0.05, 0.02)^\top,$$
$$\boldsymbol{\mu}_{IV,2}^{(1)} = (0.18, -0.17, 0.00, 0.35, -0.18, -0.18)^\top, \quad \boldsymbol{\mu}_{IV,2}^{(2)} = (-0.19, -0.34, 0.00, 0.14, 0.19, 0.19)^\top,$$
$$\boldsymbol{\mu}_{IV,2}^{(3)} = (-0.19, 0.20, -0.23, 0.20, -0.19, 0.21)^\top, \quad \boldsymbol{\mu}_{IV,2}^{(4)} = (-0.20, 0.08, -0.22, 0.08, -0.19, 0.08)^\top,$$
$$\boldsymbol{\mu}_{IV,3}^{(1)} = (0.00, 0.00, 0.00, 0.00, -0.35, 0.35)^\top, \quad \boldsymbol{\mu}_{IV,3}^{(2)} = (0, 0, \tfrac{1}{2}, 0, 0, 0)^\top,$$
$$\boldsymbol{\mu}_{IV,3}^{(3)} = (0, \tfrac{1}{5}, 0, \tfrac{1}{10}, 0, 0)^\top, \quad \boldsymbol{\mu}_{IV,3}^{(4)} = (0, 0, 0, \tfrac{1}{5}, -\tfrac{\sqrt{2}}{10}, \tfrac{\sqrt{2}}{10})^\top.$$

TABLE 3. Relative frequency of rejection of $H_0 : \boldsymbol{\mu} = \mathbf{0}$ (in percentage) for the test based on the statistics $\mathcal{Q}_n(\mathbf{W}_n)$ defined by (2), using $\mathbf{W}_n = \mathbf{I}_3$, $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$, $k = 1, 2$, and $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-}$, with $\mathbf{Z}_n = \sqrt{n}\bar{\mathbf{X}}_n$ and $\boldsymbol{\Sigma}_n = \mathbf{S}_n$, where the mean vectors are given in Table 1.

$(p, r) = (3, 2)$	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{I,1} = \mathrm{diag}(1, 1, 0)$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,1}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,1}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,1}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,1}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_3)$	0.9	4.8	10.6	55.0	77.0	85.0	89.7	97.1	98.4	92.4	97.7	99.4	61.5	82.3	88.6
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.3	5.8	11.0	57.4	77.0	85.0	90.3	96.9	98.7	90.3	96.9	98.7	57.4	77.0	85.0
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	0.4	<u>**2.9**</u>	<u>**7.3**</u>	27.6	45.3	54.0	50.5	65.3	71.5	50.5	65.3	71.5	27.6	45.3	54.0
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	1.3	5.8	11.0	57.4	77.0	85.0	90.3	96.9	98.7	90.3	96.9	98.7	57.4	77.0	85.0
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{I,2} = \mathrm{diag}(10, 1, 0)$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,2}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,2}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,2}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{I,2}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_3)$	1.2	5.6	11.3	50.0	71.5	81.0	8.3	24.4	39.4	12.3	35.5	57.3	56.3	79.4	86.3
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.3	5.8	11.0	39.7	61.6	73.4	59.9	82.6	89.0	59.9	82.6	89.0	39.7	61.6	73.4
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	0.9	5.6	11.0	50.1	71.1	80.9	6.6	17.4	26.5	6.6	17.4	26.5	50.1	71.1	80.9
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	1.3	5.8	11.0	39.7	61.6	73.4	59.9	82.6	89.0	59.9	82.6	89.0	39.7	61.6	73.4

For our investigations, we considered multivariate normal distributions of dimensions $p = 3, 6$, where $\boldsymbol{\Sigma}$ was singular and non-singular. The definitions of the covariance matrices $\boldsymbol{\Sigma}$ are given in Tables 3-6. In Tables 3 and 5, the covariance matrices are exactly singular and they are diagonal. We investigated situations where the non null eigenvalues are both equal to one, and when they are different. In the case $p = 6$, we investigated four unit eigenvalues, multiplicities of dimension two and a situation where the non null eigenvalues are distinct. In Tables 4 and 6 the covariance matrices are of the form $\boldsymbol{\Sigma} = \mathbf{I}_p - c\mathbf{A}\mathbf{A}^\top$, where c is a real value and $\mathbf{A}$ is a $p \times r$ matrix; they are precisely defined in the Tables. It is easily shown that these matrices have at least $p - r$ unit eigenvalues. In time series, the asymptotic covariance matrices of residual autocovariances exhibit similar forms, see Li (2004, Chapter 2), among others. The covariance matrices in Tables 4 and 6 are non-singular, but the ratio of the largest to the smallest eigenvalue is large; consequently they are approximately singular. The values of the non null mean vectors under the alternative hypotheses are given in Tables 1 and 2. For each case, 1000 independent realizations have been generated.

5.2 Discussion of the Monte Carlo results

In Tables 3-6, the results for the level study correspond to the column $\boldsymbol{\mu} = \mathbf{0}$. For the nominal level $\alpha = 5\%$, the empirical size over the 1000 independent replications should belong to the interval $[3.6\%, 6.4\%]$ with probability 95% (at the nominal levels $\alpha = 1\%$ and 10%, the intervals are $[0.4\%, 1.6\%]$ and $[8.1\%, 11.9\%]$, respectively). When the relative rejection frequencies are outside the 95% significance limits, they are displayed in bold in the Tables. When the relative rejection frequencies are outside the 99% significant limits, they are underlined. At the nominal levels $\alpha = 1\%$, 5% and 10%, the 99% significance intervals are $[0.2\%, 1.8\%]$, $[3.2\%, 6.8\%]$ and $[7.6\%, 12.4\%]$, respectively.

TABLE 4. Relative frequency of rejection of $H_0 : \boldsymbol{\mu} = \mathbf{0}$ (in percentage) for the test based on the statistics $\mathcal{Q}_n(\mathbf{W}_n)$ defined by (2), using $\mathbf{W}_n = \mathbf{I}_3$, $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$, $k = 1, 2, 3$, and $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-}$, with $\mathbf{Z}_n = \sqrt{n}\bar{\mathbf{X}}_n$ and $\boldsymbol{\Sigma}_n = \mathbf{S}_n$, where the mean vectors are given in Table 1.

$(p, r) = (3, 3)$	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II,1} = \mathbf{I}_3 - 0.79\mathbf{x}_3\mathbf{x}_3^\top$, $\mathbf{x}_3 = (1, \frac{1}{2}, \frac{1}{10})^\top$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,1}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,1}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,1}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,1}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_3)$	1.0	4.8	10.7	62.0	99.9	100.0	86.5	94.9	97.5	26.9	48.8	62.2	79.5	92.3	96.0
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.2	5.4	11.3	100.0	100.0	100.0	83.5	93.0	96.1	24.9	44.2	57.9	98.3	99.8	99.9
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	0.4	**2.9**	**7.3**	3.7	12.8	21.8	47.9	62.5	68.4	14.1	29.3	40.1	41.6	58.3	67.0
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	1.3	5.8	11.0	8.1	21.8	32.1	86.3	95.2	97.9	27.3	50.3	63.3	80.4	91.8	95.6
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	1.2	5.4	11.3	100.0	100.0	100.0	83.5	93.0	96.1	24.9	44.2	57.9	98.3	99.8	99.9
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II,2} = \mathbf{I}_3 - 0.333\mathbf{X}\mathbf{X}^\top$, $\mathbf{X} = (\mathbf{c}_1, \mathbf{c}_2), \mathbf{c}_1 = \mathbf{1}_3, \mathbf{c}_2 = (1, -1, 0)^\top$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,2}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,2}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,2}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{II,2}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_3)$	1.3	5.8	10.3	45.8	67.7	77.5	6.3	26.9	46.2	6.1	40.3	84.9	15.3	41.4	60.1
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.2	5.4	11.3	31.6	54.3	66.6	29.4	50.6	65.3	100.0	100.0	100.0	100.0	100.0	100.0
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	1.1	5.1	10.9	45.8	68.7	78.8	1.2	6.2	10.9	1.0	5.2	10.8	4.9	14.5	22.4
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	1.3	5.8	11.0	36.8	60.0	71.2	34.4	58.2	70.5	1.4	5.9	11.0	37.9	63.0	75.8
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	1.2	5.4	11.3	31.6	54.3	66.6	29.4	50.6	65.3	100.0	100.0	100.0	100.0	100.0	100.0

We discuss the results presented in Table 3. In general, the errors of the first kind of the test statistics are well controlled, except for $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ at the 5% and 10% nominal levels when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{I,1}$. For the alternatives $\boldsymbol{\mu} = \boldsymbol{\mu}_{I,1}^{(1)}, \boldsymbol{\mu}_{I,1}^{(2)}$, the mean vectors belong to the eigenspace generated by the eigenvalue $\lambda_1 = 1$; it appears preferable to specify $k = 2$ than $k = 1$ in that situation, which may be explained by the multiplicity of that unit eigenvalue. Interestingly, $\mathcal{Q}_n(\mathbf{S}_n^{-2})$ delivers higher power than $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ under $\boldsymbol{\mu}_{I,1}^{(1)}$, even if that alternative belongs to the vector space generated by $\mathbf{v}_1 = (1, 0, 0)^\top$; this is explained by the fact that in finite samples $\boldsymbol{\mu}_{I,1}^{(2)}$ does not belong exactly to the vector space of the first column of $\mathbf{P}_n$ in the spectral decomposition of $\mathbf{S}_n$. Under the alternatives $\boldsymbol{\mu} = \boldsymbol{\mu}_{I,1}^{(3)}, \boldsymbol{\mu}_{I,1}^{(4)}$, the empirical powers of $\mathcal{Q}_n(\mathbf{I}_3)$ and $\mathcal{Q}_n(\mathbf{S}_n^{-2})$ are very similar when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{I,1}$, with a slight advantage for the weighting $\mathbf{W}_n = \mathbf{I}_3$. These alternatives do not lie in a specific eigenspace, and $\mathcal{Q}_n(\mathbf{I}_3)$ offers high power. When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{I,2}$, all the eigenvalues are different and the covariance matrix is singular. Since $\boldsymbol{\mu}_{I,2}^{(1)} \in \mathcal{V}(\{10\})$, $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ is very powerful, but the weighting $\mathbf{W}_n = \mathbf{I}_3$ delivers a similar power. The differences in powers between $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ and $\mathcal{Q}_n(\mathbf{S}_n^{-2}) = \mathcal{Q}_n(\mathbf{S}_n^{-})$ is significant. For $\boldsymbol{\mu}_{I,2}^{(2)} \in \mathcal{V}(\{10, 1\})$, the weighting $\mathbf{W}_n = \mathbf{I}_3$ gives low power, and to use a generalized inverse provides the best empirical power. When $\boldsymbol{\mu} = \boldsymbol{\mu}_{I,2}^{(3)}$, that alternative does not lie in a specific eigenspace; consequently all test statistics offer some power, but the generalized inverse appears the most powerful. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}_{I,2}^{(4)} \in \mathcal{V}(\{10, 0\})$; the most powerful test statistics are $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ and $\mathcal{Q}_n(\mathbf{I}_3)$.

TABLE 5. Relative frequency of rejection of $H_0 : \boldsymbol{\mu} = \mathbf{0}$ (in percentage) for the test based on the statistics $\mathcal{Q}_n(\mathbf{W}_n)$ defined by (2), using $\mathbf{W}_n = \mathbf{I}_6$, $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$, $k = 1, 2, 3, 4$, and $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-}$, with $\mathbf{Z}_n = \sqrt{n}\bar{\mathbf{X}}_n$ and $\boldsymbol{\Sigma}_n = \mathbf{S}_n$, where the mean vectors are given in Table 2.

$(p, r) = (6, 4)$															
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III,1} = \text{diag}(1, 1, 1, 1, 0, 0)$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,1}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,1}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,1}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,1}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_6)$	0.7	4.1	9.3	38.9	64.5	76.9	81.9	94.4	97.1	96.6	99.3	99.9	87.2	96.4	98.7
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.1	5.0	11.3	42.9	67.8	79.6	85.4	94.8	97.2	97.2	99.4	99.8	88.1	96.9	98.7
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	**0.3**	3.6	**7.7**	10.3	22.6	31.8	23.7	37.4	46.8	34.1	48.2	55.6	25.8	40.4	49.1
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	**0.3**	3.7	**7.9**	17.9	36.2	46.9	43.9	61.7	70.5	59.0	73.3	79.6	45.9	64.1	73.4
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	0.5	4.2	**7.8**	29.3	51.2	63.1	62.9	79.4	86.8	80.8	89.4	93.1	67.7	82.7	88.3
$\mathcal{Q}_n(\mathbf{S}_n^{-4})$	1.1	5.0	11.3	42.9	67.8	79.6	85.4	94.8	97.2	97.2	99.4	99.8	88.1	96.9	98.7
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III,2} = \text{diag}(10, 10, 1, 1, 0, 0)$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,2}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,2}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,2}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,2}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_6)$	0.7	4.6	9.6	34.6	60.2	72.0	73.1	90.5	95.0	82.7	95.4	98.1	53.6	79.3	89.4
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.1	5.0	11.3	27.7	49.9	64.4	64.9	83.9	91.1	98.2	99.6	99.8	94.8	98.6	99.6
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	**0.2**	**<u>2.8</u>**	**7.8**	16.8	33.2	44.1	41.5	59.2	66.9	41.6	59.6	67.5	21.9	41.7	51.2
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	0.9	5.1	10.7	36.0	61.5	73.3	73.6	90.3	94.8	73.6	90.2	94.8	42.2	66.1	75.9
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	0.6	5.2	9.8	30.8	54.2	67.0	68.3	87.5	92.4	86.2	93.6	97.2	72.7	86.3	91.1
$\mathcal{Q}_n(\mathbf{S}_n^{-4})$	1.1	5.0	11.3	27.7	49.9	64.4	64.9	83.9	91.1	98.2	99.6	99.8	94.8	98.6	99.6
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III,3} = \text{diag}(8, 4, 2, 1, 0, 0)$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,3}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,3}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,3}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{III,3}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_6)$	0.8	4.5	10.5	56.1	77.5	84.9	75.7	93.9	97.2	27.0	62.7	78.9	82.0	96.0	99.0
$\mathcal{Q}_n(\mathbf{S}_n^{-})$	1.1	5.0	11.3	38.4	62.8	74.5	85.0	95.0	97.5	83.0	92.5	96.9	96.9	99.5	99.9
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	0.9	5.1	9.9	57.7	79.0	87.7	14.4	29.7	40.9	1.1	4.8	10.6	31.4	54.5	67.2
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	0.8	5.2	10.6	48.4	70.1	80.2	89.3	97.3	98.8	1.8	7.3	14.8	76.2	90.5	94.9
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	0.5	5.2	10.9	42.6	66.5	75.8	88.3	96.8	98.0	84.5	94.6	97.2	86.8	94.8	97.6
$\mathcal{Q}_n(\mathbf{S}_n^{-4})$	1.1	5.0	11.3	38.4	62.8	74.5	85.0	95.0	97.5	83.0	92.5	96.9	96.9	99.5	99.9

From the results presented in Table 4, the empirical sizes were rather satisfactory, except for $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ at the 5% and 10% nominal levels, when $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II,1}$ which underrejected and displayed rejection rates outside the 99% significance limits. When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II,1}$, the spectral decomposition gives $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 4.6 \times 10^{-3}$. A basis for the eigenspace is composed of the vectors $\{\mathbf{v}_1 = (0.0, -0.2, 1.0)^\top, \mathbf{v}_2 = (0.5, -0.9, -0.2)^\top, \mathbf{v}_3 = (0.9, 0.4, 0.1)^\top\}$. Clearly $\boldsymbol{\mu}_{II,1}^{(1)}$ and $\mathbf{v}_1$ are orthogonal and as a result the empirical powers of $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k \leq 2$ were low. The best empirical powers have been observed for the $\{2\}$-inverse with $k = 3$. The weighting $\mathbf{W}_n = \mathbf{I}_3$ offered less power. The alternative $\boldsymbol{\mu}_{II,1}^{(2)}$ belongs to $\mathcal{V}(\mathbf{v}_1)$. The best empirical powers have been observed by $\mathcal{Q}_n(\mathbf{S}_n^{-2})$ and $\mathcal{Q}_n(\mathbf{S}_n^{-3})$. The weighting $\mathbf{W}_n = \mathbf{I}_3$ offered high power. The vector $\boldsymbol{\mu}_{II,1}^{(3)} \in \mathcal{V}(\mathbf{v}_1, \mathbf{v}_2)$;

TABLE 6. Relative frequency of rejection of $H_0 : \boldsymbol{\mu} = \mathbf{0}$ (in percentage) for the test based on the statistics $\mathcal{Q}_n(\mathbf{W}_n)$ defined by (2), using $\mathbf{W}_n = \mathbf{I}_6$, $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-k}$, $k = 1, 2, 3, 4, 5, 6$, and $\mathbf{W}_n = \boldsymbol{\Sigma}_n^{-}$, with $\mathbf{Z}_n = \sqrt{n}\bar{\mathbf{X}}_n$ and $\boldsymbol{\Sigma}_n = \mathbf{S}_n$, where the mean vectors are given in Table 2.

$(p,r) = (6,6)$	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{IV,1} = \mathbf{I}_6 - \frac{9}{20}\mathbf{x}_6\mathbf{x}_6^\top, \mathbf{x}_6 = (1, \frac{4}{5}, \frac{3}{5}, \frac{2}{5}, \frac{1}{5}, \frac{1}{10})^\top$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,1}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,1}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,1}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,1}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_6)$	0.9	4.7	8.7	94.3	98.9	99.6	95.5	99.1	99.7	91.9	98.5	99.2	49.4	98.2	100.0
$\mathcal{Q}_n(\mathbf{S}_n^-)$	1.6	$\underline{\mathbf{7.0}}$	**12.0**	97.0	99.5	99.7	95.8	99.2	99.5	92.6	98.3	99.2	100.0	100.0	100.0
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	**0.3**	**3.2**	$\underline{\mathbf{6.7}}$	26.0	40.4	48.6	24.6	38.2	46.7	23.8	39.2	47.3	0.4	3.3	6.7
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	**0.2**	$\underline{\mathbf{2.7}}$	$\underline{\mathbf{6.6}}$	45.7	63.0	71.2	45.7	61.6	69.6	46.6	63.7	71.7	0.2	2.5	6.6
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	0.5	**3.5**	$\underline{\mathbf{7.0}}$	65.6	79.5	84.9	65.2	77.5	83.2	64.2	79.8	85.7	0.5	3.4	6.9
$\mathcal{Q}_n(\mathbf{S}_n^{-4})$	0.7	4.1	8.8	83.4	92.1	94.8	82.6	91.9	94.4	82.4	92.2	95.2	0.7	3.9	8.3
$\mathcal{Q}_n(\mathbf{S}_n^{-5})$	1.5	6.1	11.4	95.5	99.2	99.6	96.8	99.4	99.6	93.7	98.6	99.6	1.7	6.1	11.7
$\mathcal{Q}_n(\mathbf{S}_n^{-6})$	1.6	$\underline{\mathbf{7.0}}$	**12.0**	97.0	99.5	99.7	95.8	99.2	99.5	92.6	98.3	99.2	100.0	100.0	100.0
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{IV,2} = \mathbf{I}_6 - 0.16\mathbf{X}\mathbf{X}^\top, \mathbf{X} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4), \mathbf{c}_1 = \mathbf{1}_6, \mathbf{c}_2 = (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})^\top,$ $\mathbf{c}_3 = (\frac{1}{2}, 0, 0, 0, 0, \frac{1}{2})^\top, \mathbf{c}_4 = (0, 0, \frac{1}{2}, 0, 0, 0)^\top$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_6)$	0.8	4.7	9.1	94.5	99.0	99.5	94.3	98.9	99.5	96.7	99.5	99.7	74.3	90.9	96.8
$\mathcal{Q}_n(\mathbf{S}_n^-)$	1.6	$\underline{\mathbf{7.0}}$	**12.0**	93.1	98.4	99.3	93.7	98.4	99.4	98.8	99.8	99.8	100.0	100.0	100.0
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	**0.2**	**3.2**	$\underline{\mathbf{7.0}}$	34.3	50.4	58.0	30.1	45.8	54.5	4.5	13.9	21.7	2.1	7.3	14.1
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	**0.2**	**3.3**	**7.7**	56.4	72.8	79.3	53.6	71.1	77.8	14.5	28.2	39.2	4.9	14.1	22.8
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	0.4	**3.3**	$\underline{\mathbf{7.3}}$	76.3	88.1	92.2	74.3	85.8	90.8	30.6	46.6	56.0	13.4	26.6	35.6
$\mathcal{Q}_n(\mathbf{S}_n^{-4})$	**0.3**	3.8	8.4	89.8	95.8	98.4	88.8	96.3	97.9	58.9	72.1	78.1	30.8	48.5	58.9
$\mathcal{Q}_n(\mathbf{S}_n^{-5})$	1.5	6.1	11.4	94.5	99.1	99.6	94.7	98.7	99.4	99.3	99.8	99.8	76.1	88.6	93.2
$\mathcal{Q}_n(\mathbf{S}_n^{-6})$	1.6	$\underline{\mathbf{7.0}}$	**12.0**	93.1	98.4	99.3	93.7	98.4	99.4	98.8	99.8	99.8	100.0	100.0	100.0
	$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{IV,3} = \mathbf{I}_6 - 0.49\mathbf{X}\mathbf{X}^\top, \mathbf{X} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_5), \mathbf{c}_1 = \mathbf{c}_2 = (1, 0, 0, 0, 0, 0)^\top,$ $\mathbf{c}_3 = (0, 1, 0, 0, 0, 0)^\top, \mathbf{c}_4 = (0, 0, 0, 1, 0, 0)^\top, \mathbf{c}_5 = (0, 0, 0, 0, 1, 1)^\top$														
	$\boldsymbol{\mu} = \mathbf{0}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(1)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(2)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(3)}$			$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(4)}$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\mathcal{Q}_n(\mathbf{I}_6)$	0.4	4.5	9.2	97.3	99.4	99.8	97.9	99.6	99.8	23.2	49.9	65.1	45.7	71.1	81.3
$\mathcal{Q}_n(\mathbf{S}_n^-)$	1.6	$\underline{\mathbf{7.0}}$	**12.0**	92.7	98.1	99.2	94.2	98.6	99.3	45.3	66.1	76.0	56.4	74.9	84.1
$\mathcal{Q}_n(\mathbf{S}_n^{-1})$	**0.2**	$\underline{\mathbf{2.5}}$	$\underline{\mathbf{7.2}}$	63.0	73.7	77.4	57.4	69.9	76.0	0.3	3.8	9.0	11.6	25.7	35.7
$\mathcal{Q}_n(\mathbf{S}_n^{-2})$	0.7	4.6	9.7	96.9	99.0	99.7	97.7	99.6	99.7	1.8	6.9	13.4	19.4	38.7	51.4
$\mathcal{Q}_n(\mathbf{S}_n^{-3})$	0.6	4.8	9.3	96.6	99.1	99.6	96.6	99.3	99.7	20.7	39.0	48.5	36.2	57.6	69.0
$\mathcal{Q}_n(\mathbf{S}_n^{-4})$	1.1	5.0	11.1	95.6	99.1	99.6	96.3	99.2	99.6	49.2	70.7	79.7	61.3	79.8	87.0
$\mathcal{Q}_n(\mathbf{S}_n^{-5})$	1.2	4.6	9.9	94.2	98.6	99.5	95.5	98.5	99.3	45.5	67.0	77.3	57.9	76.0	84.5
$\mathcal{Q}_n(\mathbf{S}_n^{-6})$	1.6	$\underline{\mathbf{7.0}}$	**12.0**	92.7	98.1	99.2	94.2	98.6	99.3	45.3	66.1	76.0	56.4	74.9	84.1

the best power has been observed for the $\{2\}$-inverse with $k = 2$. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}_{II,1}^{(4)}$ does not lie in a specific eigenspace; the best power has been observed with a $\{2\}$-inverse with $k = 3$.

When $\mathbf{\Sigma} = \mathbf{\Sigma}_{II,2}$, $\lambda_1 = 1.000, \lambda_2 = 0.334, \lambda_3 = 10^{-3}$. The first alternative $\boldsymbol{\mu}^{(1)}_{II,2}$ belongs to the eigenspace associated to λ_1; it appears that $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ is very powerful. The second alternative $\boldsymbol{\mu}^{(2)}_{II,2}$ belongs to the eigenspace associated to λ_2, and $\mathcal{Q}_n(\mathbf{S}_n^{-2})$ dominates whereas, as expected, $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ has no power. The third alternative lies in $\mathcal{V}(\{\lambda_3\})$; the test statistics based on the $\{2\}$-inverses with $k = 1, 2$ offered no power and the best power has been attained at $k = 3$. The alternative $\boldsymbol{\mu}^{(4)}_{II,2}$ does not lie in a specific eigenspace. Low power has been observed for $k = 1$ and $k = 2$, and the best empirical power has been observed when $k = 3$. Note that when $\boldsymbol{\mu} = \boldsymbol{\mu}^{(2)}_{II,2}, \boldsymbol{\mu}^{(3)}_{II,2}, \boldsymbol{\mu}^{(4)}_{II,2}$, the weighting $\mathbf{W}_n = \mathbf{I}_3$ displayed low power and it appeared preferable to use a generalized inverse or a $\{2\}$-inverse.

In Table 5, multivariate sampling from a singular normal distribution of dimension six and rank four was performed. In general the rejection rates were satisfactory under the null, except for $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ when $\mathbf{\Sigma} = \mathbf{\Sigma}_{III,1}, \mathbf{\Sigma}_{III,2}$, which underrejects slightly. In the case $\mathbf{\Sigma} = \mathbf{\Sigma}_{III,1}$, slight underrejection occurred at the 10% nominal level for $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 1, 2, 3$. When $\mathbf{\Sigma} = \mathbf{\Sigma}_{III,1}$, the empirical powers of $\mathcal{Q}_n(\mathbf{I}_6)$ and $\mathcal{Q}_n(\mathbf{S}_n^{-}) = \mathcal{Q}_n(\mathbf{S}_n^{-4})$ were similar for the alternatives included in the study. All the alternatives belong to the eigenspace $\mathcal{V}(\{1\})$. Under these alternatives the best powers were attained by $\mathcal{Q}_n(\mathbf{S}_n^{-}) = \mathcal{Q}_n(\mathbf{S}_n^{-4})$. When $\mathbf{\Sigma} = \mathbf{\Sigma}_{III,2}$, the best power was observed for $k = 2$ when the alternative was $\boldsymbol{\mu}^{(1)}_{III,2}$, which belongs to $\mathcal{V}(\{10\})$. The best $\{2\}$-inverse was the one with $k = 2$ under the alternative $\boldsymbol{\mu} = \boldsymbol{\mu}^{(2)}_{III,2}$, which is also in the eigenspace $\mathcal{V}(\{10\})$. The differences with the weighting $\mathbf{W}_n = \mathbf{I}_6$ were rather small. When $\boldsymbol{\mu} = \boldsymbol{\mu}^{(3)}_{III,2}, \boldsymbol{\mu}^{(4)}_{III,2}$, the generalized inverse offered the highest power. When $\mathbf{\Sigma} = \mathbf{\Sigma}_{III,3}$, the best power was reached by the test statistic $\mathcal{Q}_n(\mathbf{S}_n^{-k})$ based on a $\{2\}$-inverse with $k = 1$, which can be explained because $\boldsymbol{\mu}^{(1)}_{III,3} \in \mathcal{V}(\{8\})$ and the dimension of that eigenspace is one. When $\boldsymbol{\mu} = \boldsymbol{\mu}^{(2)}_{III,3}$, the best power is observed with $\mathcal{Q}_n(\mathbf{S}_n^{-2})$. The test statistic $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ offered low power under that particular alternative. On the other hand, the power differences between $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k \geq 2$ were rather small. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}^{(3)}_{III,3}$ belongs to $\mathcal{V}(\{2\})$. Consequently, $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k \leq 2$, have no power. The best empirical power is obtained when $k = 3$ for the $\{2\}$-inverse, but the generalized inverse exhibits also high power. The weighting $\mathbf{W}_n = \mathbf{I}_6$ offers some power, but that procedure was significantly less powerful than the generalized inverse. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}^{(4)}_{III,3}$ belongs to the eigenspace generated by the non null eigenvalues. Consequently the best power is attained with the $\{2\}$-inverse with $k = 4$.

Finally, we analyze the results in Table 6. In general the rejection rates were reasonable under the null. When $\mathbf{\Sigma} = \mathbf{\Sigma}_{IV,1}, \mathbf{\Sigma}_{IV,2}$, some underrejection has been observed for $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 1, 2, 3$, which seemed more pronounced at the 10% nominal level. Overrejection occurred for $\mathcal{Q}_n(\mathbf{S}_n^{-}) = \mathcal{Q}_n(\mathbf{S}_n^{-6})$ at the 5% and 10% nominal levels. Some underrejection has been observed for $\mathcal{Q}_n(\mathbf{S}_n^{-1})$ when sampling from a normal distribution with covariance matrix $\mathbf{\Sigma} = \mathbf{\Sigma}_{IV,3}$. The eigenvalues of $\mathbf{\Sigma} = \mathbf{\Sigma}_{IV,1}$ are $\lambda_1 = \ldots = \lambda_5 = 1$ and $\lambda_6 = 5.5 \times 10^{-3}$. The alternatives $\boldsymbol{\mu} = \boldsymbol{\mu}^{(1)}_{IV,1}, \boldsymbol{\mu}^{(2)}_{IV,1}, \boldsymbol{\mu}^{(3)}_{IV,1}$ are all in the orthogonal complement of the eigenspace associated to λ_6. For all these alternatives, the empirical power increase with k, and the best powers are attained by $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 5, 6$. In general the differences between the weighting $\mathbf{W}_n = \mathbf{I}_6$ and $\mathbf{W}_n = \mathbf{S}_n^{-}$ were rather small. When $\mathbf{\Sigma} = \mathbf{\Sigma}_{IV,2}$, $\lambda_1 = \lambda_2 = 1, \lambda_3 = 0.97, \lambda_4 = 0.94, \lambda_5 = 0.75$ and $\lambda_6 = 5.3 \times 10^{-3}$. The alternatives

$\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(1)}, \boldsymbol{\mu}_{IV,2}^{(2)}$ belong to the eigenspace associated to the unit eigenvalue. However, since the eigenvalues λ_i, $i = 1, \ldots, 5$ are close, the best power are offered by $\mathcal{Q}_n(\mathbf{S}_n^{-k})$ with a large k. The alternatives $\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(5)}$ lie in the eigenspace associated to λ_5. The best power is observed for $\mathcal{Q}_n(\mathbf{S}_n^{-5})$. In general the differences between the weighting $\mathbf{W}_n = \mathbf{I}_6$ and $\mathbf{W}_n = \mathbf{S}_n^{-}$ were rather small under the alternatives $\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(1)}, \boldsymbol{\mu}_{IV,2}^{(2)}, \boldsymbol{\mu}_{IV,2}^{(3)}$. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,2}^{(4)}$ belong to the eigenspace associated to λ_5 and λ_6. Low power is observed for $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k \leq 4$. The best power are observed for $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 5, 6$. In general the differences between the weighting $\mathbf{W}_n = \mathbf{I}_6$ and $\mathbf{W}_n = \mathbf{S}_n^{-5}$ were small under that alternative but substantially lower than the $\{2\}$-inverse with $k = 6$. When $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{IV,3}$, the eigenvalues are $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = \lambda_4 = 0.51$ and $\lambda_5 = \lambda_6 = 0.02$. The alternatives $\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(1)}, \boldsymbol{\mu}_{IV,3}^{(2)}$ belong to the eigenspace associated to λ_1. Consequently $\mathcal{Q}_n(\mathbf{S}_n^{-2})$ was the most powerful. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(3)}$ belongs to the eigenspace associated to λ_3. Consequently $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 1, 2$ had no power. The best empirical power has been observed when $\mathcal{Q}_n(\mathbf{S}_n^{-4})$. There were slight differences between $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 4, 5, 6$. The alternative $\boldsymbol{\mu} = \boldsymbol{\mu}_{IV,3}^{(4)}$ belongs to the eigenspace associated to λ_1 and λ_3. The test statistic $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k = 1, 2, 3$, offered some power under that alternative, but the best power has been observed with the test statistic $\mathcal{Q}_n(\mathbf{S}_n^{-4})$.

6. TESTING FOR CLIMATE CHANGES

The three classes of test procedures are now illustrated on a set of monthly reconstructions of temperatures and precipitations‡ (see Casty *et al.* 2005). These spatio-temporal data extend from January 1659 to December 1999 and cover a gridded area of 197 points over the whole European Alp region (note that the data file contains a gridded area of 275 points, but for each observation the same 197 points are measured). Our first aim is to compare the mean temperature over the last 40 years with the mean temperature over the whole period in order to test for a significant change. We constructed the 12-dimensional multivariate time series of temperatures, denoted $\mathbf{t}_t = (\mathbf{t}_t(1), \ldots, \mathbf{t}_t(12))^\top$, such that $\mathbf{t}_t(i)$ corresponds to the monthly average at time t and month i over the 197 grid points, $i = 1, \ldots, 12$, $t = 1659, \ldots, 1999$. The monthly mean are represented in Figure 4.

For our testing problem, we defined the time series $\mathbf{X}_t = \frac{1}{40}\sum_{i=1}^{40} \mathbf{t}_{t+1657+i}$ for $t = 1, \ldots, n = 302$ and the test statistic $\mathbf{Z}_n = \mathbf{X}_n - \frac{1}{n-1}\sum_{t=1}^{n-1} \mathbf{X}_t$. We supposed that the series of temperatures $\{\mathbf{X}_t\}$ constituted a stationary sequence with constant mean $\boldsymbol{\mu}_{\mathbf{X}}$, variance $\boldsymbol{\Sigma}_{\mathbf{X}}$, and autocovariance function $\boldsymbol{\Gamma}_{\mathbf{X}}(\cdot)$, that we presumed to be absolutely summable, that is $\sum_{h=1}^{\infty} \|\boldsymbol{\Gamma}_{\mathbf{X}}(h)\| < \infty$. Let $E(\mathbf{Z}_n) = \boldsymbol{\mu}_{\mathbf{Z}}$. Under the null hypothesis:

$$H_0 : \boldsymbol{\mu}_{\mathbf{Z}} = \mathbf{0}, \text{ against } H_1 : \boldsymbol{\mu}_{\mathbf{Z}} \neq \mathbf{0}, \tag{16}$$

and the variance of the test statistic $\mathbf{Z}_n$ is given by:

$$\boldsymbol{\Sigma} = \frac{n}{n-1}\boldsymbol{\Sigma}_{\mathbf{X}} - (n-1)^{-1}\sum_{h=1}^{n-1}\left\{\boldsymbol{\Gamma}_{\mathbf{X}}(h) + \boldsymbol{\Gamma}_{\mathbf{X}}^\top(h)\right\} + (n-1)^{-2}\sum_{h=1}^{n-2}(n-1-h)\left\{\boldsymbol{\Gamma}_{\mathbf{X}}(h) + \boldsymbol{\Gamma}_{\mathbf{X}}^\top(h)\right\}.$$

‡Casty, C., et al. 2008. European Gridded Monthly Temperature, Precipitation and 500hPa Reconstructions. IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series # 2008-023. NOAA/NCDC Paleoclimatology Program, Boulder CO, USA.

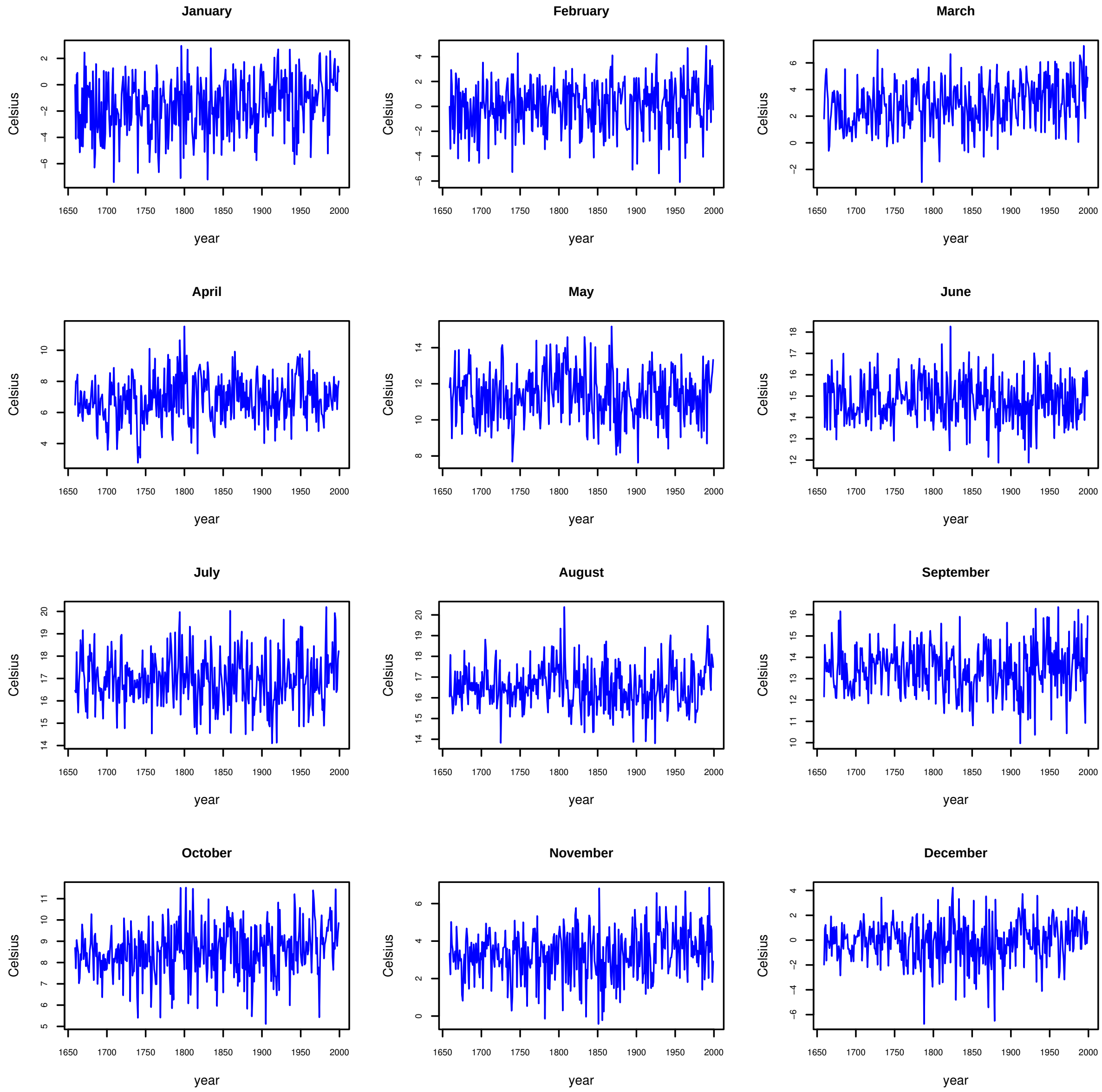


FIGURE 4. Monthly mean Alpine temperature from 1659 to 1999.

Since the autocovariance function is assumed absolutely summable, it follows that $\mathbf{\Sigma} \to \mathbf{\Sigma_X}$ almost surely, as $n \to \infty$. In order to estimate consistently the variance $\mathbf{\Sigma}$, a simple estimator is given by the empirical variance $\mathbf{S}_n$ of $\mathbf{X}_1, \ldots, \mathbf{X}_n$. An analysis of the eigenvalues of $\mathbf{S}_n$ revealed that the smallest (resp. largest) eigenvalue was 2.13×10^{-3} (resp. 5.36×10^{-1}), suggesting that the sample covariance matrix was relatively close to a singular matrix.

The test statistics $\mathcal{Q}_n(\mathbf{I}_{12})$, $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k \in \{1, \ldots, 11\}$ and $\mathcal{Q}_n(\mathbf{S}_n^{-})$ were computed. Since $\mathbf{S}_n$ is invertible, the generalized inverse is in fact the inverse, and the test statistic $\mathcal{Q}_n(\mathbf{S}_n^{-})$ is the classical Wald test procedure, that is $\mathcal{Q}_n(\mathbf{S}_n^{-}) = \mathcal{Q}_n(\mathbf{S}_n^{-1})$. From (v) in our Corollary 1, the $\mathcal{Q}_n(\mathbf{S}_n^{-})$ test statistic is likely to have a larger Bahadur slope than any $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $1 \le k \le 12$, and also than the Imhof-based test statistic $\mathcal{Q}_n(\mathbf{I}_{12})$ (but we cannot compare directly $\mathcal{Q}_n(\mathbf{I}_{12})$ and $\mathcal{Q}_n(\mathbf{S}_n^{-k})$). Using the local power analysis, any $\{2\}$-inverse is asymptotically locally more powerful than $\mathcal{Q}_n(\mathbf{S}_n^-)$ (and even $\mathcal{Q}_n(\mathbf{I}_{12})$) in certain directions. From our results, we cannot conclude which one is best between the Imhof-based test and $\mathcal{Q}_n(\mathbf{S}_n^-)$. The p-values of the different test procedures are displayed under the column labelled 'Temperature' in Table 7. At the nominal 5% level, the null hypothesis is not rejected by the test statistics based on the $\{2\}$-inverses with $k \in \{2, \dots, 7\}$, but is rejected by $\mathcal{Q}_n(\mathbf{S}_n^{-1})$, and $\mathcal{Q}_n(\mathbf{S}_n^{-k})$, $k \in \{8, \dots, 11\}$. As we have seen in the previous sections, the $\{2\}$-inverse test statistics $\mathcal{Q}_n(\mathbf{S}_n^{-k})$ and $\mathcal{Q}_n(\mathbf{S}_n^-)$ may have low powers in certain directions of the alternative hypothesis, that may explain that several test statistics do not reject the null hypothesis. By comparison, the Imhof-based test, which enjoys power in all directions, rejects the null at the usual 5% level. Moreover, the conservative Bonferroni procedure (consisting in rejecting if the minimal p-value multiplied by the number of tests is less than a given level) also tends to reject the assumption that the Alpine temperature of the 40 last years be stochastically similar to that of the period of reference. This is in accordance with many empirical studies exhibiting an accumulation of extremes positive temperature during the recent past (see *e.g.* Casty *et al.* 2005).

The same exercise has been performed, replacing the series of temperatures $\{\mathbf{t}_t\}$ by a multivariate time series composed of precipitations. Following the same procedure that described previously, a multivariate time series of monthly precipitations, denoted $\{\mathbf{p}_t\}$, $t = 1659, \dots, 1999$, has been created. The smallest (resp. largest) eigenvalue was 1.45 (resp. 131.0). The results are under the column entitled 'Precipitation' in Table 7. The results suggest that the test statistics do not detect any evidence against the null hypothesis that the average precipitations of the last 40 years are stochastic similar to those of the whole period. That conclusion is in accordance with studies showing that the precipitation dynamics does not exhibit a particular trend over the last period, but shows 'a clear cyclic variability on a timescale of 40-60 years' (see *e.g.* Casty *et al.* 2007 for more details).

7. DISCUSSION AND CONCLUSION

Generalized Wald's method constructs testing procedures having chi-squared limiting distributions from test statistics having singular normal limiting distributions by use of generalized inverses. In this article, we investigated the use of $\{2\}$-inverses for that problem and we proposed new test statistics with convenient asymptotic chi-square distributions. Imhof-based test statistics have also be studied, which converge in distribution to weighted sum of chi-square variables. We discussed the asymptotic null distributions of the test statistics, and we performed a power analysis under fixed and local alternatives. Simulation studies have been performed to study the exact levels in finite samples, and the exact powers have been compared empirically in a simulation study.

In general the test statistics offered satisfactory empirical levels; the test statistics based on the $\{2\}$-inverses with small values of k offered some underrejection, but generally in the 99% significance limits and reasonably close to the 95% significance intervals. From our theoretical and empirical results, the spectral decomposition of the covariance matrix plays an important role on the power properties. If an alternative lied in a specific eigenspace, powerful test procedures were constructed based on $\{2\}$-inverses with orders chosen large enough such that the associated eigenspaces included that alternative. This was expected from our theoretical results (see Corollaries 1 and 2) and confirmed in

TABLE 7. Testing stability against climate changes: p-values (in percentage) of the $\mathcal{Q}_n(\cdot)$-tests defined by (2) for the testing problem (16) when the $\mathbf{X_t}$'s correspond to temperature averages or precipitation averages.

		Temperature	Precipitation
Imhof $\mathcal{Q}_n(\mathbf{I}_{12})$		3.3	66.4
	$k=1$	2.0	96.9
	$k=2$	6.6	85.7
	$k=3$	12.3	45.5
	$k=4$	13.7	48.0
	$k=5$	13.1	57.4
$\{2\}$-inverse $\mathcal{Q}_n(\mathbf{S}_n^{-k})$	$k=6$	6.0	45.2
	$k=7$	5.6	39.5
	$k=8$	4.1	36.9
	$k=9$	0.4	23.4
	$k=10$	0.6	18.4
	$k=11$	1.0	21.4
Moore-Penrose $\mathcal{Q}_n(\mathbf{S}_n^{-})$		1.5	24.6

the simulation experiments. In practice, a spectral decomposition of the covariance matrix appears thus useful in specifying the order k: If the alternative of interest belongs to a particular eigenspace, it dictates the choice of k. When the covariance matrix was singular or approximately singular, and when the alternative lied in the eigenspace associated to the non-null eigenvalues, test statistics based on $\{2\}$-inverses with an order equal to the estimated rank of the covariance matrix were particularly powerful test procedures. If an alternative was orthogonal to the eigenspace associated to the eigenvalues used to construct a test statistic based on a $\{2\}$-inverse, low power has been observed (see also Corollary 2, ii)). The weighting based on a generalized inverse offered high power in several cases, and the omnibus weighting $\mathbf{W}_n = \mathbf{I}_p$ provided also interesting power, and in fact was very powerful for the alternatives which were in the eigenspace generated by the null eigenvalues.

The test procedures have been illustrated in the data analysis on the monthly temperature and precipitation variability in the European Alps. In comparing the monthly temperature of the last 40 years with the whole period under study, a significant difference has been found using the Imhof-based test, using the test statistics relying on $\{2\}$-inverses with $k=1$, $7 \le k \le 11$, and also for the test statistic using the generalized inverse. No significant difference has been found for the precipitation time series. Since the $\{2\}$-inverses may offer high power in certain directions, and low power in others, our data analysis contributed to explain the directions of the alternative hypothesis which entailed rejection of the null hypothesis of equal mean temperature.

APPENDIX. CONSTRUCTION OF THE $\{2\}$-INVERSE.

In order to compute $\mathbf{\Sigma}_B^{-k}$, with $k < r$, $r = \mathrm{rank}(\mathbf{\Sigma})$, we describe an algorithm, which has been used in Sections 5 and 6. Given a fixed tolerance $\epsilon >$, a basis $B = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ of $\mathbb{R}^r$, a symmetric semi-definite matrix $\boldsymbol{\Sigma}$ and an integer $1 \leq k < r$, the following steps are performed.

1. First, compute the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r$ of $\boldsymbol{\Sigma}$;
2. Find the largest integer $\underline{k} \leq k$ such that $\lambda_{\underline{k}} > \lambda_{k+1} + \epsilon$;
3. Define the following set: $\mathfrak{s}(k) = \{i, \ |\lambda_k - \lambda_i| < \epsilon\}$;
4. Compute the cardinal m_k of $\mathfrak{s}(k)$;
5. Calculate the matrix $\boldsymbol{M}_{\mathfrak{s}(k)}$ such that the m_k columns constitute an orthonormal basis of the eigenspace $\mathcal{V}_{\mathfrak{s}(k)}$ associated to the eigenvalues such that the indices are in $\mathfrak{s}(k)$;
6. The orthogonal projection on $\mathcal{V}_{\mathfrak{s}(k)}$ can be computed, and is given by $\boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}} = \boldsymbol{M}_{\mathfrak{s}(k)}\boldsymbol{M}_{\mathfrak{s}(k)}^\top$;
7. Define the generator $\boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}}(B) = \left\{\boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}}(\boldsymbol{u}_1), \ldots, \boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}}(\boldsymbol{u}_m)\right\}$;
8. A basis given by $\left\{\boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}}(\boldsymbol{u}_{i_1}), \ldots, \boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}}(\boldsymbol{u}_{i_{m_k}})\right\}$ is calculated, taking the m_k vectors of $\boldsymbol{P}_{\mathcal{V}_{\mathfrak{s}(k)}}(B)$ such that the norm is larger than ϵ and such that their distance is superior to ϵ of the space generated by the preceding vectors of the system (if that operation does not provide a basis, ϵ was chosen too large; thus ϵ is divided by two and the algorithm returns to step 2);
9. An orthonormal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m_k}\}$ is determined, applying the Gram-Schmidt process on the basis obtained in the preceding step;
10. if $\underline{k} > 0$, a matrix $\boldsymbol{M}_{\{1,\ldots,\underline{k}\}}$ such that the $\underline{k}$ columns constitute an orthonormal basis of the eigenspace associated to the eigenvalues $\lambda_1, \ldots, \lambda_{\underline{k}}$;
11. The matrices $\boldsymbol{\Lambda}_B^{-\underline{k}} = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_{\underline{k}}^{-1}, \boldsymbol{0}_{m-\underline{k}}^\top)$, $\boldsymbol{\Sigma}_B^{-\underline{k}} = \boldsymbol{M}_{\{1,\ldots,\underline{k}\}}\boldsymbol{\Lambda}_B^{-\underline{k}}\boldsymbol{M}_{\{1,\ldots,\underline{k}\}}^\top$ if $\underline{k} > 0$, adopting the convention $\boldsymbol{\Sigma}_B^{-\underline{k}} = 0$ if $\underline{k} = 0$, and $\boldsymbol{\Sigma}_B^{-k} = \boldsymbol{\Sigma}_B^{-\underline{k}} + \sum_{j=1}^{k-\underline{k}} \lambda_k^{-1}\boldsymbol{v}_i\boldsymbol{v}_i^\top$.

The algorithm defined by steps 1-11 gives a function $\mathfrak{A}_{B,k,\epsilon}$ such that $\mathfrak{A}_{B,k\epsilon}(\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}_B^{-k}$.

REFERENCES

Andrews, D. W. K. (1987), 'Asymptotic results for generalized Wald tests', *Econometric Theory* **3**, 348–358.

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.

Bahadur, R. R. (1960), 'Comparisons of tests', *The Annals of Mathematical Statistics* **31**, 276–295.

Bhimasankaram, P. and Sengupta, D. (1991), 'Testing for the mean vector of a multivariate normal distribution with a possibly singular dispersion matrix and related results', *Statistics & Probability Letters* **11**, 473–478.

Bilodeau, M. and Brenner, D. (1999), *Theory of Multivariate Statistics*, Springer, New York.

Box, G. E. P. and Pierce, D. A. (1970), 'Distribution of residual autocorrelations in autoregressive-integrated moving average time series models', *Journal of the American Statistical Association* **65**, 1509–1526.

Casty, C., Wanner, H., Luterbacher, J., Esper, J. and Böhm,R. (2005), 'Temperature and precipitation variability in the European Alps since 1500', *International Journal of Climatology* **25**, 1855–1880.

Casty, C., Raible, C.C., Stocker, T.F., Luterbacher, J. and H. Wanner. (2007), 'A European pattern climatology 1766-2000', *Climate Dynamics*, DOI: 10.1007/s00382-007-0257-6.

Duchesne, P. and Francq, C. (2008), 'On diagnostic checking time series models with portmanteau test statistics based on generalized inverses and { 2 }-inverses', *COMPSTAT 2008, Proceedings in Computational Statistics*, 143–154.

Eaton, M. L. (2007), *Multivariate Statistics. A Vector Space Approach*, Institute of Mathematical Statistics, Lecture Notes - Monograph Series, vol. 53, Ohio.

Francq, C., Roy, R. and Zakoïan, J.-M. (2005), 'Diagnostic checking in ARMA models with uncorrelated errors', *Journal of the American Statistical Association* **100**, 532–544.

Getson, A. J. and Hsuan, F. C. (1988), {2}*-Inverses and Their Statistical Application*, Lecture Notes in Statistics 47, Springer-Verlag, New York.

Ghosh, B. K. (1973), 'Some monotonicity theorems for χ^2, F and t distributions with applications', *Journal of the Royal Statistical Society, Series B* **35**, 480–492.

Hadi, A. S. and Wells, M. T. (1990), 'A note on generalized Wald's method', *Metrika* **37**, 309–315.

Imhof, J. P. (1961), 'Computing the distribution of quadratic forms in normal variables' *Biometrika* **48**, 419–426.

Li, W. K. (2004), *Diagnostic Checks in Time Series.* Chapman & Hall/CRC, New York.

Li, W. K. (1992), 'On the asymptotic standard errors of residual autocorrelations in nonlinear time series modelling', *Biometrika* **79**, 435–437.

Li, W. K. and McLeod, A. I. (1981), 'Distribution of the residual autocorrelations in multivariate ARMA time series models', *Journal of the Royal Statistical Society, Series B* **43**, 231–239.

Ljung, G. M. (1986), 'Diagnostic testing of univariate time series models', *Biometrika* **73**, 725–730.

Moore, D. S. (1978), 'Chi-square tests', in *Studies in Statistics, vol. 19*, Hogg, R. V. (Editor), 66–106, The Mathematical Association of America, Washington DC.

Moore, D. S. (1977), 'Generalized inverses, Wald's method, and the construction of chi-squared tests of fit', *Journal of the American Statistical Association* **72**, 131–137.

Muirhead, R. J. (1982), *Aspects of Multivariate Statistical Theory*, Wiley, New York.

Rao, C. R. and Mitra, S. K. M. (1971), *Generalized Inverse of Matrices and its Applications*, Wiley, New York.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, Wiley, New York.

Robinson, P. M. (1972), 'Non-linear regression for multiple time-series', *Journal of Applied Probability* **9**, 758–768.

Srivastava, M. S. and Khatri, C. G. (1979), *An Introduction to Multivariate Statistics*, North-Holland, New York.

Tyler, D. E. (1981), 'Asymptotic inference for eigenvectors', *The Annals of Statistics* **4**, 725–736.

Wald, A. (1943), 'Tests of statistical hypothesis concerning several parameters when the number of observations is large', *Transaction of the American Mathematical Society* **54**, 426–482.

Zolotarev, V. M. (1961), 'Concerning a certain probability problem', *Theory of Probability and its Applications* **6**, 201–204.