On testing the equality of latent roots of scatter matrices under ellipticity

Gaspard Bernard[a,*], Thomas Verdebout[b]

[a] Mathematics Department, University of Luxembourg, Campus Belval, Maison du Nombre, Avenue de la Fonte 6, L-4364 Esch-sur-Alzette, Luxembourg

[b] ECARES and Mathematics Department, Université libre de Bruxelles (ULB), Campus Plaine, Boulevard du Triomphe CP210, B-1050 Brussels, Belgium

Abstract

In the present paper, we tackle the problem of testing $\mathcal{H}_{0q} : \lambda_q > \lambda_{q+1} = \cdots = \lambda_p$, where $\lambda_1, \ldots, \lambda_p$ are the scatter matrix eigenvalues of an elliptical distribution on $\mathbb{R}^p$. This is a classical problem in multivariate analysis which is very useful in dimension reduction. We analyse the problem using the Le Cam asymptotic theory of experiments and show that contrary to the testing problems on eigenvalues and eigenvectors of a scatter matrix tackled in [7], the non-specification of nuisance parameters has an asymptotic cost for testing $\mathcal{H}_{0q}$. We moreover derive signed-rank tests for the problem that enjoy the property of being asymptotically distribution-free under ellipticity. The van der Waerden rank test uniformly dominates the classical pseudo-Gaussian procedure for the problem. Numerical illustrations show the nice finite-sample properties of our tests.

Keywords:
Elliptical distributions, Hypothesis testing, Latent roots.
2020 MSC: Primary 62H15, Secondary 62F05

1. Introduction

Many classical models in multivariate analysis include scatter/shape parameters. A very common example is the elliptical model characterised by a p-vector $\mathbf{X}$ whose characteristic function is of the form $\mathbf{t} \mapsto e^{i\mathbf{t}'\boldsymbol{\mu}}\phi((\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})^{1/2})$ for some characteristic generator ϕ. Here, the p-vector $\boldsymbol{\mu}$ is a location parameter and the $p \times p$ symmetric and positive definite matrix $\boldsymbol{\Sigma}$ is a scatter parameter. A popular elliptical distribution is of course the p-variate Gaussian distribution that is obtained with $\phi(s) = \exp(-s^2/2)$. Inference on the scatter parameter in the elliptical model has been the subject of many contributions: to cite only a few, [5, 14, 26] studied the asymptotic properties of robust estimators of $\boldsymbol{\Sigma}$, [2] provided several properties of the Minimum Covariance Determinant estimator of $\boldsymbol{\Sigma}$ while [25] computed the influence functions of empirical canonical correlation coefficients.

The general topic of the present paper is the well-known dimension reduction technique called Principal Component Analysis (PCA). In the elliptical framework, PCA based on estimators of $\boldsymbol{\Sigma}$ has been considered in [8, 23], [4, 27] studied estimators of the eigenvalues of $\boldsymbol{\Sigma}$, [3] computed influence functions of robust estimators of eigenvalues and eigenvectors of scatter matrices while [7] considered locally and asymptotically optimal tests for some problems involving also eigenvectors and eigenvalues of scatter/shape matrices. Sphericity tests have recently been studied, e.g., in [6, 12, 13]. Many testing problems in PCA including the one we consider in this paper are invariant with respect to scale transformations so that the formulation of the problem does not change when $\boldsymbol{\Sigma}$ is replaced with $c\boldsymbol{\Sigma}$ for any $c > 0$. It is therefore reasonable to write the problem in terms of the shape matrix $\mathbf{V} := \boldsymbol{\Sigma}/(\det\boldsymbol{\Sigma})^{1/p}$ associated with $\boldsymbol{\Sigma}$ (see [20]). Note that if the underlying distribution has finite second order moments, its covariance matrix is also proportional to $\mathbf{V}$.

Consider the spectral decomposition $\mathbf{V} = \boldsymbol{\beta}\boldsymbol{\Lambda}\boldsymbol{\beta}^\top$ of the shape matrix $\mathbf{V}$, where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p)$ is the matrix of orthonormal eigenvectors and $\boldsymbol{\Lambda} := \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ is the matrix of strictly positive and ordered ($\lambda_1 \geq \cdots \geq \lambda_p$)

*Corresponding author. Email address: gaspard.bernard@uni.lu

eigenvalues. The problem we consider throughout this paper is the problem of testing the equality of several latent roots. More precisely, fixing $q < p-1$, we consider the problem of testing the equality of the $p-q$ smallest eigenvalues $\mathcal{H}_{0q} : \lambda_q > \lambda_{q+1} = \cdots = \lambda_p$ of $\mathbf{V}$. This problem has been investigated a lot in the multivariate statistics literature since the main objective of PCA is often to extract a low-dimensional signal from the data. If the smallest $p-q$ eigenvalues cannot be distinguished, the corresponding components can be considered as noise. Methods for determining the number of components to be captured can be traced back to the work of Lawley [11], who developed Gaussian likelihood ratio tests to check the equality of the smallest eigenvalues. A pseudo-Gaussian test that is valid under elliptical assumptions has been proposed in [29]. The local asymptotic powers of robust tests have been obtained in [27] while other procedures for the problem have been investigated in [15], [16] and [17] to cite a few. Finally, high-dimensional tests have been studied in [24] and more recently in [28].

In the present paper, the objective is twofold: (i) we derive locally and asymptotically optimal tests for the problem within the class of elliptical distributions and (ii) we show how they can be turned into tests that combine nice asymptotic power properties (including local and asymptotic optimality) and natural robustness to outliers. The backbone of our approach is the Le Cam asymptotic theory of experiments. In a very similar context, [7] derived a Local and Asymptotic Normality (LAN) property assuming that the ordered eigenvalues $\lambda_1 > \cdots > \lambda_p$ of $\mathbf{V}$ are distinct. The Fisher Information matrix associated with this LAN property is block diagonal; one block for the location parameter $\boldsymbol{\mu}$, one block for the scale parameter $\sigma^2 := (\det \boldsymbol{\Sigma})^{1/p}$, one block for the eigenvectors $\boldsymbol{\beta}$ of the shape matrix and one block for the eigenvalues $\boldsymbol{\Lambda}$ of $\mathbf{V}$. This block diagonal structure implies that the non-specification of nuisance parameters $\boldsymbol{\mu}, \sigma$ and $\boldsymbol{\beta}$ when inference on $\boldsymbol{\Lambda}$ is performed has no asymptotic cost. This is precisely what is obtained in the testing problems considered in [7]. The situation we consider in this paper is very different since under the null hypothesis $\mathcal{H}_{0q} : \lambda_q > \lambda_{q+1} = \cdots = \lambda_p$, the smallest $p-q$ eigenvalues are equal so that the LAN property derived in [7] does not hold. We show that for the testing problem considered here, the non-specification of the nuisance parameter has an asymptotic cost. The various developments also make it possible to obtain the form of sign and signed-rank tests which, to the best of our knowledge, have not been derived for the problem yet.

The paper is organised as follows. In Section 2, we define the various notations used in the sequel, derive a LAN property and obtain the shape of locally and asymptotically optimal tests when an arbitrary value of $\boldsymbol{\beta}$ is specified. The particular Gaussian case is discussed in Section 3 where we also consider the unspecified nuisance case. In Section 4, we discuss how the obtained parametric tests can be turned into robust tests including sign and signed-rank tests.

2. Local asymptotic normality and optimal tests

For the sake of convenience, we collect here the notations that will be used in the paper. Throughout, $\mathbf{e}_\ell$ will denote the ℓth vector of the canonical basis of $\mathbb{R}^p$, so that $\mathbf{K}_p := \sum_{i,j=1}^{p}(\mathbf{e}_i\mathbf{e}_j^\top) \otimes (\mathbf{e}_j\mathbf{e}_i^\top)$ is the usual commutation matrix. Denoting as $\text{vec}\,(\mathbf{A})$ the vector obtained by stacking the columns of the matrix $\mathbf{A}$ on top of each other, we put $\mathbf{J}_p := \text{vec}\,(\mathbf{I}_p)\text{vec}^\top(\mathbf{I}_p)$, where $\mathbf{I}_\ell$ is the ℓ-dimensional identity matrix. We will write $\text{dvec}\,(\mathbf{A}) =: (\mathbf{A}_{11}, (\overset{\circ}{\text{dvec}}\,(\mathbf{A}))^\top)^\top$ for the p-dimensional vector obtained by stacking the diagonal elements of $\mathbf{A}$: $\overset{\circ}{\text{dvec}}\,(\mathbf{A})$ thus is $\text{dvec}\,(\mathbf{A})$ deprived of its first component. Let $\mathbf{H}_p$ be the $p \times p^2$ matrix such that $\mathbf{H}_p\text{vec}\,(\mathbf{A}) = \text{dvec}\,(\mathbf{A})$. Note that $\mathbf{H}_p\mathbf{H}_p^\top = \mathbf{I}_p$ and that if $\mathbf{A}$ is diagonal, then $\mathbf{H}_p^\top\text{dvec}\,(\mathbf{A}) = \text{vec}\,(\mathbf{A})$. Write $\text{diag}(\mathbf{B}_1, \ldots, \mathbf{B}_m)$ for the block-diagonal matrix with blocks $\mathbf{B}_1, \ldots, \mathbf{B}_m$ and $\mathbf{A}^{\otimes 2} := \mathbf{A} \otimes \mathbf{A}$. We also write $\overset{\circ}{\text{vech}}\,(\mathbf{A})$, the vector obtained by stacking the upper diagonal entries of $\mathbf{A}$ deprived of its first component. For a symmetric and positive definite matrix $\mathbf{B}$, we will denote by $\mathbf{B}^{1/2}$ its symmetric and positive definite square root.

Throughout the paper, we assume that the observations at hand $\mathbf{X}_1, \ldots, \mathbf{X}_n$ form a random sample from the absolutely continuous p-variate elliptical distribution with location $\boldsymbol{\mu}$, scale parameter σ^2, shape matrix $\mathbf{V} = \boldsymbol{\beta}\boldsymbol{\Lambda}\boldsymbol{\beta}^\top$ (with $p \times p$ eigenvectors matrix $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p)$ and eigenvalues $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_p)$) and positive radial function f_1; the corresponding density is of the form

$$\mathbf{x} \to c_{p,\sigma^2,f_1} f_1\Big(\frac{1}{\sigma}((\mathbf{x}-\boldsymbol{\mu})^\top\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu}))^{1/2}\Big),$$

where c_{p,σ^2,f_1} is a normalizing constant. In the sequel, we will tacitly assume that the location parameter $\boldsymbol{\mu}$ is known and fix $\boldsymbol{\mu} = \mathbf{0}$ for the sake of clarity and simplicity. The tests we will derive below are based on the random matrices (5) and (12) used in [6]. It directly follows from the same [6] that replacing the unknown location parameter by any root-n consistent estimator will have no asymptotic effect on testing procedures based on the latter matrices for the shape parameter. More precisely, all the results obtained below can be stated in exactly the same way by replacing the observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$ with $\mathbf{X}_1 - \hat{\boldsymbol{\mu}}, \ldots, \mathbf{X}_n - \hat{\boldsymbol{\mu}}$, where $\hat{\boldsymbol{\mu}}$ is a root-n consistent estimator of $\boldsymbol{\mu}$. The joint distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_n$ will be throughout denoted by $\mathrm{P}^{(n)}_{\sigma,\mathbf{V};f_1}$ or equivalently by $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};f_1}$.

Below, we derive a LAN result that does not require distinct shape eigenvalues as in [7] but still requires classical assumptions on the radial density function f_1. Define the elliptical coordinates (again, without loss of generality we put $\boldsymbol{\mu} = \mathbf{0}$)

$$\mathbf{U}_i(\mathbf{V}) := \frac{\mathbf{V}^{-1/2}\mathbf{X}_i}{\|\mathbf{V}^{-1/2}\mathbf{X}_i\|} \quad \text{and} \quad d_i(\mathbf{V}) := \|\mathbf{V}^{-1/2}\mathbf{X}_i\|. \tag{1}$$

Under the assumption of ellipticity, the multivariate signs $\mathbf{U}_i(\mathbf{V})$, $i \in \{1, \ldots, n\}$ are i.i.d. uniform over the unit sphere in $\mathbb{R}^p$, and independent of the standardised elliptical distances $d_i(\mathbf{V})$. As in [7], we assume that f_1 belongs to the class of positive, absolutely continuous real functions $\mathcal{F}_1$ such that (i) the $d_i(\mathbf{V})$'s have density $\tilde{f}_1$ and distribution function $\tilde{F}_1$ with median one and (ii) denoting by $\dot{f}_1$ the a.e. derivative of f_1 and letting $\varphi_{f_1} := -\dot{f}_1/f_1$, the integrals

$$\mathcal{I}_p(f_1) := \int_0^1 \varphi^2_{f_1}(\tilde{F}_1^{-1}(u))\, du \quad \text{and} \quad \mathcal{J}_p(f_1) := \int_0^1 \varphi^2_{f_1}(\tilde{F}_1^{-1}(u))(\tilde{F}_1^{-1}(u))^2\, du \tag{2}$$

are finite. The latter constraints identify the scatter and shape matrices without requiring any moment assumptions and guarantee the finiteness of the Fisher information matrix. Note that under finite second-order moments, the scatter and shape matrices are proportional to the traditional covariance matrix $\boldsymbol{\Sigma}_{\mathrm{cov}}$ of the $\mathbf{X}_i$'s, see [7] for more details. As explained in the Introduction, the testing problem is invariant with respect to scale transformations so that the parameter of interest in this work are the eigenvalues $\mathrm{d}\mathring{\mathrm{vec}}\,(\boldsymbol{\Lambda})$ of the shape matrix. The parameter $\boldsymbol{\mu}$ is assumed to be $\mathbf{0}$ while the parameters σ and $\boldsymbol{\beta}$ are playing the role of nuisance parameters having to be estimated. The LAN result below involves local perturbations of $\mathrm{d}\mathring{\mathrm{vec}}\,(\boldsymbol{\Lambda}) \in \mathcal{H}_{0q}$ (such that $\lambda_q > \lambda_{q+1} = \cdots = \lambda_p$) of the form

$$\mathrm{d}\mathring{\mathrm{vec}}\,(\boldsymbol{\Lambda}^{(n)}) := \mathrm{d}\mathring{\mathrm{vec}}\,(\boldsymbol{\Lambda} + n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)})) =: \mathrm{d}\mathring{\mathrm{vec}}\,(\boldsymbol{\Lambda}) + n^{-1/2}\boldsymbol{\tau}^{(n)}, \tag{3}$$

where $\boldsymbol{\ell}^{(n)} =: (\ell_1^{(n)}, \ldots, \ell_p^{(n)})$ is a bounded sequence of $\mathbb{R}^p$ such that $\det(\boldsymbol{\Lambda}^{(n)}) = \det(\boldsymbol{\Lambda} + n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)})) = 1$ and $(\prod_{j=2}^p(\lambda_j + n^{-1/2}\ell_j^{(n)})^{-1}) = \lambda_1 + n^{-1/2}\ell_1^{(n)} \geq \lambda_2 + n^{-1/2}\ell_2^{(n)} \geq \cdots \geq \lambda_p + n^{-1/2}\ell_p^{(n)}$. Since, as $n \to \infty$,

$$0 \;=\; \det(\boldsymbol{\Lambda} + n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)})) - \det(\boldsymbol{\Lambda}) = n^{-1/2}\mathrm{tr}(\boldsymbol{\Lambda}^{-1}\mathrm{diag}(\boldsymbol{\ell}^{(n)})) + O(n^{-1}),$$

we readily obtain that $\boldsymbol{\ell}^{(n)}$ must be such that $\mathrm{tr}(\boldsymbol{\Lambda}^{-1}\mathrm{diag}(\boldsymbol{\ell}^{(n)})) = O(n^{-1/2})$ as $n \to \infty$. We have the following result.

Proposition 1. *Assume that $f_1 \in \mathcal{F}_1$, fix an arbitrary orthonormal matrix $\boldsymbol{\beta}$ and let $\boldsymbol{\ell}^{(n)}$ be a bounded sequence of $\mathbb{R}^p$ such that* (3) *holds with $\boldsymbol{\tau} := \lim_{n\to\infty}\boldsymbol{\tau}^{(n)} = \lim_{n\to\infty}\mathrm{d}\mathring{\mathrm{vec}}\,(\mathrm{diag}(\boldsymbol{\ell}^{(n)}))$. Let moreover $\mathbf{M}_p = \mathbf{M}_p(\mathrm{d}\mathring{\mathrm{vec}}\,(\boldsymbol{\Lambda}))$ be the $(p-1)\times p$ matrix such that* (i) $\mathbf{M}_p\mathrm{dvec}\,(\boldsymbol{\Lambda}^{-1}) = \mathbf{0}$ *and* (ii) $\mathbf{M}_p^\top\mathrm{d}\mathring{\mathrm{vec}}\,(\mathbf{L}) = \mathrm{dvec}(\mathbf{L})$ *for any matrix $\mathbf{L}$ such that* $\mathrm{tr}(\boldsymbol{\Lambda}^{-1}\mathbf{L}) = 0$. *Then, as $n \to \infty$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};f_1}$,*

$$\ln\Big(\frac{\mathrm{dP}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda}+n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)});f_1}}{\mathrm{dP}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};f_1}}\Big) = \boldsymbol{\tau}^\top\boldsymbol{\Delta}^{(n)}_{f_1}(\mathbf{V}) - \frac{1}{2}\boldsymbol{\tau}^\top\boldsymbol{\Gamma}_{f_1}(\mathbf{V})\boldsymbol{\tau} + o_{\mathrm{P}}(1), \tag{4}$$

where letting

$$\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) \;:=\; n^{-1}\sum_{i=1}^n \varphi_{f_1}\Big(\frac{d_i(\mathbf{V})}{\sigma}\Big)\frac{d_i(\mathbf{V})}{\sigma}\,\mathbf{U}_i(\mathbf{V})\mathbf{U}_i^\top(\mathbf{V}), \tag{5}$$

$\boldsymbol{\Delta}^{(n)}_{f_1}(\mathbf{V})$ *and* $\boldsymbol{\Gamma}_{f_1}(\mathbf{V})$ *in* (4) *are given by*

$$\boldsymbol{\Delta}^{(n)}_{f_1}(\mathbf{V}) := \frac{1}{2}\mathbf{M}_p n^{1/2}\mathrm{dvec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top\mathbf{S}^{(n)}_{f_1}(\mathbf{V})\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) \quad \text{and} \quad \boldsymbol{\Gamma}_{f_1}(\mathbf{V}) := \frac{\mathcal{J}_p(f_1)}{2p(p+2)}\mathbf{M}_p\boldsymbol{\Lambda}^{-2}\mathbf{M}_p^\top.$$

Proposition 1 entails that the considered sequence of experiments is locally and asymptotically normal (LAN). It allows to compute the asymptotic powers of tests against local alternatives using the classical third Le Cam Lemma and also to provide the shape of locally and asymptotically optimal tests. The quantity $\mathbf{\Delta}_{f_1}^{(n)}(\mathbf{V})$ in Proposition 1 is often called the central sequence. Following the classical Le Cam asymptotic theory, the tests we propose in this section will be based on $\mathbf{\Delta}_{f_1}^{(n)}(\mathbf{V})$. Locally in $\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda}) + n^{-1/2}\boldsymbol{\tau}^{(n)}$ with $\overset{\circ}{\text{dvec}}\,(\mathbf{\Lambda}) \in \mathcal{H}_{0q}$, the testing problem can be written as $\mathcal{H}_{0q} : \boldsymbol{\tau}^{(n)} \in \mathcal{M}(\mathbf{\Upsilon})$, where $\mathcal{M}(\mathbf{A})$ denotes the span of the the columns of $\mathbf{A}$ and letting $\mathbf{1}_k := (1, \ldots, 1)^\top \in \mathbb{R}^k$, $\mathbf{\Upsilon} := \text{diag}(\mathbf{I}_{q-1}, \mathbf{1}_{p-q})$. For such local linear restrictions, it directly follows from Proposition 1 that a locally and asymptotically most stringent test with specified nuisance (the notation Sp in the index is used to emphasise this) $\phi_{\text{Sp},f_1}^{(n)}$ rejects $\mathcal{H}_{0q}$ at the asymptotic level α when ($\boldsymbol{\beta}_{0q} := (\boldsymbol{\beta}_{q+1}, \ldots, \boldsymbol{\beta}_p)$)

$$\begin{aligned} S_{q,f_1}^{(n)}(\mathbf{V}) &= (\mathbf{\Delta}_{f_1}^{(n)}(\mathbf{V}))^\top(\mathbf{\Gamma}_{f_1}(\mathbf{V}))^{-1/2}(\mathbf{I}_{p-1} - (\mathbf{\Gamma}_{f_1}(\mathbf{V}))^{1/2}\mathbf{\Upsilon}(\mathbf{\Upsilon}^\top\mathbf{\Gamma}_{f_1}(\mathbf{V})\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}^\top(\mathbf{\Gamma}_{f_1}(\mathbf{V}))^{1/2})(\mathbf{\Gamma}_{f_1}(\mathbf{V}))^{-1/2}\mathbf{\Delta}_{f_1}^{(n)}(\mathbf{V}) \\ &= \frac{np(p+2)}{2\mathcal{J}_p(f_1)}(\text{tr}(\text{diag}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}_{f_1}^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q})) - (p-q)^{-1}\text{tr}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}_{f_1}^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q})) > \chi^2_{p-q-1;1-\alpha}, \end{aligned} \quad (6)$$

where $\mathbf{\Delta}_{f_1}^{(n)}(\mathbf{V})$ and $\mathbf{\Gamma}_{f_1}(\mathbf{V})$ are given in Proposition 1 and $\chi^2_{\nu;\tau}$ denotes the quantile of order $\tau \in (0,1)$ of the chi-square distribution with ν degrees of freedom. Note that obtaining the last line of (6) requires some easy computations and Lemma 4.1 in [7] . The test statistic (6) can be seen as the norm of the standardised central sequence $\mathbf{\Gamma}_{f_1}^{-1/2}(\mathbf{V})\mathbf{\Delta}_{f_1}^{(n)}(\mathbf{V})$ in the metric associated with the orthogonal complement of the space spanned by the columns of $\mathbf{\Gamma}_{f_1}^{1/2}(\mathbf{V})\mathbf{\Upsilon}$.

3. Optimal Gaussian test

The tests based on $S_{q,f_1}^{(n)}(\mathbf{V})$ in (6) typically are valid under standardised radial density f_1 only. They are clearly of theoretical interest since they settle the optimality bounds at given density f_1. Nevertheless, due to the central role of normal distributions in multivariate analysis, the Gaussian case ($f_1(r) = \phi_1(r) = \exp(-a_p r^2/2)$, where a_p is such that $f_1 \in \mathcal{F}_1$) deserves more investigations. In this section, we therefore provide the Gaussian version $S_{q,\phi_1}^{(n)}(\mathbf{V})$ of the optimal test statistics of Section 2. These tests require Gaussian densities, though; when Gaussian assumptions are questionable, the pseudo-Gaussian procedures discussed in Section 4 are preferable—all the more so that they do not imply any loss of efficiency at the Gaussian radial function. For $f_1 = \phi_1$, the test statistic $S_{q,f_1}^{(n)}(\mathbf{V})$ in (6) takes the form

$$S_{q,\phi_1}^{(n)}(\mathbf{V}) = \frac{n}{2}(\text{tr}(\text{diag}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}_{\phi_1}^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q})) - (p-q)^{-1}\text{tr}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}_{\phi_1}^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q})), \quad (7)$$

where (remember that $\mathbf{\Sigma} = \sigma^2\mathbf{V}$)

$$\mathbf{S}_{\phi_1}^{(n)}(\mathbf{V}) := \frac{a_p}{n}\sum_{i=1}^n \mathbf{\Sigma}^{-1/2}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{\Sigma}^{-1/2}.$$

The test statistic (7) clearly involves parameters that have to be estimated. The Gaussian MLEs, that is, the eigenvalues $\lambda_{1;\mathbf{S}}, \ldots \lambda_{p;\mathbf{S}}$ and eigenvectors $\boldsymbol{\beta}_{1;\mathbf{S}}, \ldots, \boldsymbol{\beta}_{p;\mathbf{S}}$ of the empirical covariance matrix

$$\mathbf{S}^{(n)} = \frac{1}{n}\sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i^\top = \sum_{j=1}^p \lambda_{j;\mathbf{S}}\,\boldsymbol{\beta}_{j;\mathbf{S}}\,\boldsymbol{\beta}_{j;\mathbf{S}}^\top$$

are natural candidates to use for the estimation of the various parameters involved in $S_{q,\phi_1}^{(n)}(\mathbf{V})$. Letting $\underline{\lambda}$ stands for the common value of $\lambda_{q+1} = \cdots = \lambda_p$ of the shape matrix $\mathbf{V}$ under $\mathcal{H}_{0q}$, we easily obtain that under $\mathcal{H}_{0q}$,

$$S_{q,\phi_1}^{(n)}(\mathbf{V}) = \frac{na_p^2}{2\sigma^4\underline{\lambda}^2}(\text{tr}(\text{diag}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}^{(n)}\boldsymbol{\beta}_{0q})) - (p-q)^{-1}\text{tr}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}^{(n)}\boldsymbol{\beta}_{0q})).$$

A natural consistent and constrained estimator of the ratio $a_p/\sigma^2\underline{\lambda}$ is simply the sample average $(p-q)^{-1}\sum_{j=q+1}^p \lambda_{j;\mathbf{S}}$ of the $p-q$ smallest eigenvalues while $\boldsymbol{\beta}_{0q}$ can naturally be replaced by $(\boldsymbol{\beta}_{q+1;\mathbf{S}}, \ldots, \boldsymbol{\beta}_{p;\mathbf{S}})$. Plugging these estimators into $S^{(n)}_{q,\phi_1}(\mathbf{V})$, we obtain the test statistic

$$S^{(n)}_{q,\phi_1} = \frac{n}{2((p-q)^{-1}\sum_{j=q+1}^{p}\lambda_{j;\mathbf{S}})^2}\Big(\sum_{j=q+1}^{p}\lambda^2_{j;\mathbf{S}} - (p-q)^{-1}\Big(\sum_{j=q+1}^{p}\lambda_{j;\mathbf{S}}\Big)^2\Big), \tag{8}$$

which is nothing but the well-known classical Gaussian test statistic studied for instance in [28]. The test statistic $S^{(n)}_{q,\phi_1}(\mathbf{V})$ explicitly depends on the parameter $\mathbf{V}$ (which is generally unknown in practice), while the test statistic $S^{(n)}_{q,\phi_1}$ obtained after replacing $\mathbf{V}$ in (7) by a natural estimator is a perfectly genuine test statistic. As shown in [1], the weak limit of $S^{(n)}_{q,\phi_1}$ under the null hypothesis is a chi-square random variable with $d(p,q) := (p-q+2)(p-q-1)/2$ degrees of freedom and therefore does not coincide with the weak limit of $S^{(n)}_{q,\phi_1}(\mathbf{V})$ (which is a chi-square random variable with $p-q-1$ degrees of freedom) under the same null hypothesis. This reflects the fact that the estimation of the nuisance parameters has an asymptotic cost in the problem we consider. This phenomenon did not appear in the PCA testing problems considered in [7], the structure of the problem being very different. The resulting unspecified nuisance (the notation Un in the index is used to emphasise this) test $\phi^{(n)}_{\mathrm{Un},\phi_1}$ rejects $\mathcal{H}_{0q}$ at the asymptotic level α when $S^{(n)}_{q,\phi_1} > \chi^2_{d(p,q);1-\alpha}$. Defining

$$T^{(n)}_{q,\phi_1}(\mathbf{V}) = \frac{n}{2}\big(\mathrm{tr}((\boldsymbol{\beta}_{0q}^\top \mathbf{S}^{(n)}_{\phi_1}(\mathbf{V})\boldsymbol{\beta}_{0q})^2) - (p-q)^{-1}\mathrm{tr}^2(\boldsymbol{\beta}_{0q}^\top \mathbf{S}^{(n)}_{\phi_1}(\mathbf{V})\boldsymbol{\beta}_{0q})\big), \tag{9}$$

we show in the next result that $S^{(n)}_{q,\phi_1}$ and $T^{(n)}_{q,\phi_1}(\mathbf{V})$ are asymptotically equivalent under the null hypothesis (and therefore also under contiguous alternatives). The asymptotic equivalence between $S^{(n)}_{q,\phi_1}$ ans $T^{(n)}_{q,\phi_1}(\mathbf{V})$ helps to illustrate the difference between the nuisance specified and unspecified test statistics $S^{(n)}_{q,\phi_1}(\mathbf{V})$ and $S^{(n)}_{q,\phi_1}$. The test statistic $S^{(n)}_{q,\phi_1}(\mathbf{V})$ is a function of the diagonal elements of $\boldsymbol{\beta}_{0q}^\top \mathbf{S}^{(n)}_{\phi_1}(\mathbf{V})\boldsymbol{\beta}_{0q}$ only (meaning that it takes into account the fact that $\boldsymbol{\beta}$ diagonalises $\mathbf{V}$) while $T^{(n)}_{q,\phi_1}(\mathbf{V})$ is a function of the entire matrix $\boldsymbol{\beta}_{0q}^\top \mathbf{S}^{(n)}_{\phi_1}(\mathbf{V})\boldsymbol{\beta}_{0q}$ (meaning intuitively that it does not take into account the fact that $\boldsymbol{\beta}$ diagonalises $\mathbf{V}$).

Proposition 2. *Under* $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};\phi_1}$ *with* $\mathrm{d\mathring{v}ec}(\boldsymbol{\Lambda}) \in \mathcal{H}_{0q}$, $S^{(n)}_{q,\phi_1} - T^{(n)}_{q,\phi_1}(\mathbf{V})$ *is* $o_{\mathrm{P}}(1)$ *as* $n \to \infty$.

Proposition 2 shows that the non-specification of the nuisance parameter has an asymptotic cost. Indeed, plugging in estimates of the nuisance parameters in the optimal test statistic (7) yields a test statistic that is asymptotically equivalent to $T^{(n)}_{q,\phi_1}(\mathbf{V})$. Clearly $T^{(n)}_{q,\phi_1}(\mathbf{V})$ can not have the same weak limit as $S^{(n)}_{q,\phi_1}(\mathbf{V})$. Proposition 2 moreover helps to quantify the cost of the non-specification since in the following corollary we obtain (using Proposition 2 and a standard application of the third Le Cam Lemma) the local asymptotic powers of the tests $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ and $\phi^{(n)}_{\mathrm{Un},\phi_1}$ respectively based on $S^{(n)}_{q,\phi_1}(\mathbf{V})$ and on $S^{(n)}_{q,\phi_1}$.

Proposition 3. *Fix an arbitrary orthonormal matrix* $\boldsymbol{\beta}$ *and let* $\boldsymbol{\ell}^{(n)}$ *be a bounded sequence of* $\mathbb{R}^p$ *such that* (3) *holds with* $\lim_{n\to\infty}\boldsymbol{\ell}^{(n)} = \boldsymbol{\ell} = (\ell_1,\ldots,\ell_p)$. *Then, under* $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda}+n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)});\phi_1}$ *with* $\mathrm{d\mathring{v}ec}(\boldsymbol{\Lambda}) \in \mathcal{H}_{0q}$,

(i) $\lim_{n\to\infty}\mathrm{P}[S^{(n)}_{q,\phi_1}(\mathbf{V}) > \chi^2_{p-q-1;1-\alpha}] = \mathrm{P}[Y > \chi^2_{p-q-1;1-\alpha}]$, *where* Y *is a chi-square random variable with* $p-q-1$ *degrees of freedom and noncentrality parameter*

$$s(\boldsymbol{\ell}) := \frac{1}{2\underline{\lambda}^2}\Big(\sum_{j=q+1}^{p}\ell_j^2 - (p-q)^{-1}\Big(\sum_{j=q+1}^{p}\ell_j\Big)^2\Big). \tag{10}$$

(ii) $\lim_{n\to\infty}\mathrm{P}[S^{(n)}_{q,\phi_1} > \chi^2_{d(p,q);1-\alpha}] = \mathrm{P}[Y > \chi^2_{d(p,q);1-\alpha}]$, *where* Y *is a chi-square random variable with* $d(p,q)$ *degrees of freedom and noncentrality parameter* $s(\boldsymbol{\ell})$.

Interestingly, the noncentrality parameters of the chi-square weak limits of both test statistics do coincide. The loss of asymptotic power generated by the non-specification of the nuisance translates only into a modification of the number of degrees of freedom. Since the degrees of freedom of the limiting chi-square distributions of $S^{(n)}_{q,\phi_1}(\mathbf{V})$ and $S^{(n)}_{q,\phi_1}$ are different, the derivation of the Asymptotic Relative Efficiencies (AREs) of the two considered tests is not straightforward. Expressions for AREs when the asymptotic distributions of the compared test statistics are of two different types have been obtained in [18] and in [21]. The same techniques can be used here to derive closed forms for the AREs of $\phi^{(n)}_{\mathrm{Un},\phi_1}$ with respect to $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ as shown in the following result.

Proposition 4. *Letting $F^{\chi^2}_{\nu,s}$ stand for the cumulative distribution function of a noncentral chi-square random variable with ν degrees of freedom and noncentrality parameter s. Let $s^{\alpha,\xi}_{\nu}$ be such that*

$$1 - \xi = F^{\chi^2}_{\nu, s^{\alpha,\xi}_{\nu}}((F^{\chi^2}_{\nu,0})^{-1}(1-\alpha)).$$

Under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\,\boldsymbol{\Lambda}_{\mathbf{V}}+n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)});\phi_1}$ with $\mathrm{d\mathring{v}ec}\,(\boldsymbol{\Lambda}_{\mathbf{V}}) \in \mathcal{H}_{0q}$, the AREs of $\phi^{(n)}_{\mathrm{Un},\phi_1}$ with respect to $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ are given by

$$\mathrm{ARE}_{p,\phi_1}(\phi^{(n)}_{\mathrm{Un},\phi_1}/\phi^{(n)}_{\mathrm{Sp},\phi_1}) = \frac{s^{\alpha,\xi}_{p-q-1}}{s^{\alpha,\xi}_{d(p,q)}},$$

where α is the (common) asymptotic size of the tests and ξ their (common) asymptotic power.

We should note at this point that the AREs do not depend on the perturbation $n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)})$ anymore, but only on p and q for fixed α and ξ. For this reason, we display in Fig. 1 the AREs of Proposition 4 as functions of p and q for fixed $\alpha = 0.05$ and $\xi \in \{0.1, 0.5, 0.9\}$. Inspection of Fig. 1 reveals that when $(p-q)$ increases, $\mathrm{ARE}_{p,\phi_1}(\phi^{(n)}_{\mathrm{Un},\phi_1}/\phi^{(n)}_{\mathrm{Sp},\phi_1})$ decreases, meaning that the cost of the non-specification of the nuisance increases with $p-q$ as expected. Fig. 1 also tends to indicate that for a given asymptotic level α, $\mathrm{ARE}_{p,\phi_1}(\phi^{(n)}_{\mathrm{Un},\phi_1}/\phi^{(n)}_{\mathrm{Sp},\phi_1})$ does not vary too much with the asymptotic power ξ.

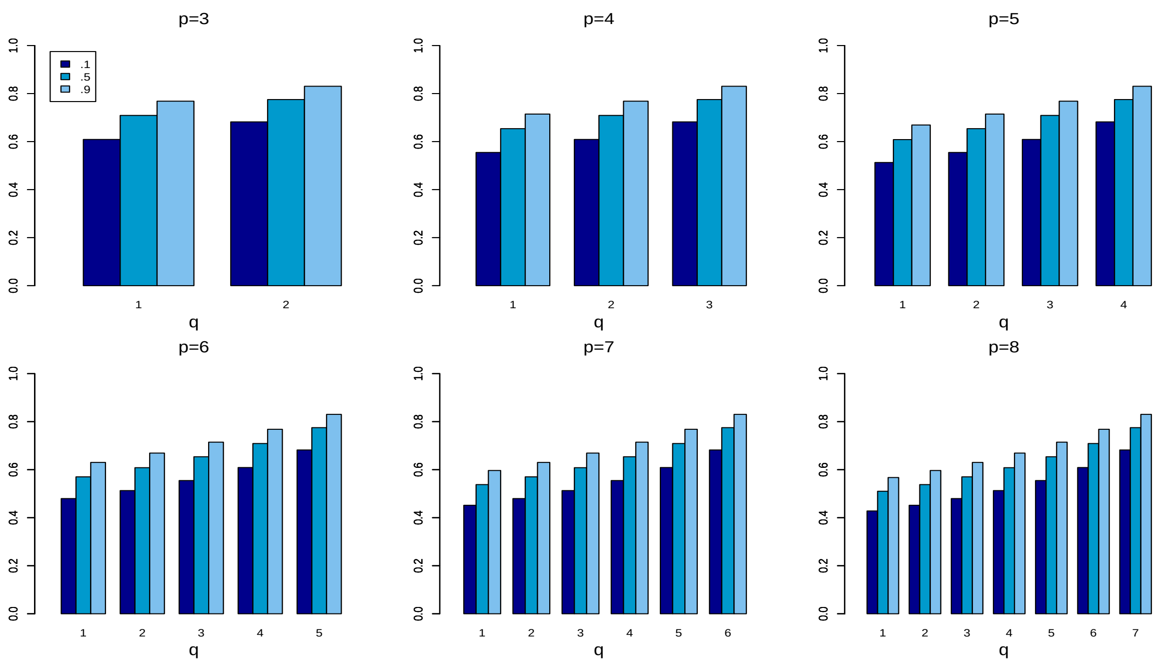

Fig. 1: Values of $\mathrm{ARE}_{p,\phi_1}(\phi^{(n)}_{\mathrm{Un},\phi_1}/\phi^{(n)}_{\mathrm{Sp},\phi_1})$ for various values of p and q. The different nuances of blue correspond to different values for the asymptotic power $\xi \in \{.1, .5, .9\}$. The nominal asymptotic level is $\alpha = 0.05$

To conclude this Section, we performed a Monte-Carlo simulation study to corroborate the results obtained in the Section. We generated $M = 40,000$ independent samples of i.i.d. observations

$$\mathbf{X}^{(\tau)}_1, \ldots, \mathbf{X}^{(\tau)}_n,$$

for $n = 20,000$ and $\tau = 0, 1/2, 1, \dots, 10$. The $\mathbf{X}_i^{(\tau)}$'s are i.i.d. with a common ($p =$)5-dimensional Gaussian distribution with mean zero and covariance matrix

$$\mathbf{\Sigma}(\tau) = \mathrm{diag}(3, 2, 2, 1, 1 - \frac{\tau}{\sqrt{n}}).$$

The value $\tau = 0$ provides a data generating process that belongs to the null hypothesis $\mathcal{H}_{03} : \lambda_3 > \lambda_4 = \lambda_5$ while for $\tau \in \{1/2, 1, \dots, 10\}$, the corresponding distributions are increasingly under the alternative. For each scenario, we performed the tests $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ and $\phi^{(n)}_{\mathrm{Un},\phi_1}$ with $p = 5$ and $q = 3$ at the nominal level $\alpha = .05$. In Fig. 2, we provide the empirical powers of $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ and $\phi^{(n)}_{\mathrm{Un},\phi_1}$ respectively as functions of τ. We also plot the limiting powers of $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ and $\phi^{(n)}_{\mathrm{Un},\phi_1}$ obtained in Proposition 3. Inspection of Fig. 2 clearly reveals that the conclusions drawn in this Section hold.

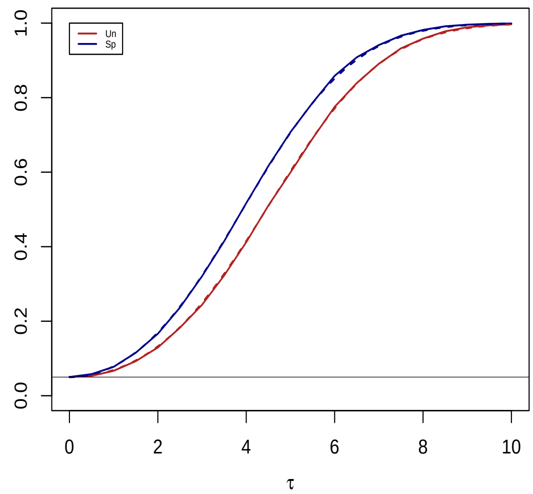

Fig. 2: The plain blue and red curves are the empirical powers of $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ and $\phi^{(n)}_{\mathrm{Un},\phi_1}$ respectively as functions of τ. The dotted blue and red curves are the corresponding limiting powers of $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ and $\phi^{(n)}_{\mathrm{Un},\phi_1}$ respectively obtained in Proposition 3. The tests are performed at the nominal asymptotic level $\alpha = .05$

4. Robust signed-rank-based tests

The parametric tests proposed in the previous Sections are valid under specified radial densities f_1 only, and therefore are of limited practical value. The importance of the Gaussian tests of Section 3 comes from the fact that they belong to usual practice, but Gaussian assumptions are quite unrealistic in most applications. The pseudo-Gaussian procedures of [29] are more appealing, as they only require finite fourth-order moments. The latter procedures are based on a kurtosis correction of the Gaussian test statistic $S^{(n)}_{q,\phi_1}$ in (8) we summarise here. Assume that the observations at hand are elliptically distributed with a radial density function f_1 that belongs to the collection $\mathcal{F}_4 \subset \mathcal{F}_1$ of radial densities such that the corresponding distribution has finite fourth order moments. This guarantees finiteness of the elliptical kurtosis coefficient

$$\kappa_p(f_1) := \frac{p\mathrm{E}[d_i^4(\mathbf{V})]}{(p+2)(\mathrm{E}[d_i^2(\mathbf{V})])^2} - 1.$$

It is well-known that the test statistic $S^{(n)}_{q,\phi_1}$ in (8) is asymptotically $\chi^2_{d(p,q)}$ under the null hypothesis if and only if $\kappa_p(f_1)$ takes the same value $\kappa_p(\phi_1) = 0$ as in the Gaussian case. Consequently, there is no guarantee that the corresponding test meets the asymptotic nominal level constraint under ellipticity, and it therefore makes little sense to use such a test in the elliptical case. This is why [29] considers the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ that rejects the null hypothesis at the nominal level α when

$$S^{(n)\dagger}_{q,\phi_1} := \frac{S^{(n)}_{q,\phi_1}}{1+\hat{\kappa}^{(n)}_p} > \chi^2_{d(p,q),1-\alpha}, \tag{11}$$

where

$$\hat{\kappa}^{(n)}_p := \frac{p(\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_i'(\mathbf{S}^{(n)})^{-1}\mathbf{X}_i)^2)}{(p+2)(\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_i'(\mathbf{S}^{(n)})^{-1}\mathbf{X}_i))^2} - 1$$

is a natural estimator of the kurtosis coefficient $\kappa_p(f_1)$. In the Gaussian case $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ is asymptotically equivalent to its original version $\phi^{(n)}_{\mathrm{Un},\phi_1}$. It may therefore be considered as a pseudo-Gaussian version of $\phi_{\mathrm{Un},\phi_1}$, since it extends its validity to the class of elliptical distributions with finite fourth-order moments without sacrificing its local and asymptotic power properties.

Still, moments of order four may be infinite and, being based on empirical covariance matrices, these procedures remain poorly robust. Signed-rank-based tests for this problem as for many others, enjoy an extremely attractive combination of robustness and efficiency properties. Following the general constructions of signed-rank tests used in [6] and [7], our test statistics will be obtained by replacing $\mathbf{S}^{(n)}_{f_1}(\mathbf{V})$ in (6) by a multivariate signed-rank measurable version of it. More precisely, our tests are based on signed-rank scatter matrices of the form

$$\mathbf{S}^{(n)}_K(\mathbf{V}) := \frac{1}{n}\sum_{i=1}^n K\Big(\frac{R^{(n)}_i(\mathbf{V})}{n+1}\Big)\mathbf{U}_i(\mathbf{V})\mathbf{U}_i(\mathbf{V})^\top, \tag{12}$$

where $K : r \in (0,1) \mapsto K(r)$ stands for some score function and $R^{(n)}_i(\mathbf{V})$ is the rank of $d_i(\mathbf{V})$ among $d_1(\mathbf{V}),\ldots,d_n(\mathbf{V})$. The various scores or score functions K appearing in our rank-based statistics will be assumed to satisfy a few regularity assumptions which, for convenience, are given here. The score function $K : (0,1) \to \mathbb{R}^+$ (i) is continuous and square-integrable, (ii) can be expressed as the difference of two monotone increasing functions, and (iii) satisfies $\int_0^1 K(u)du = p$. We define K_{f_1}, the score function associated with the radial density f_1 - which allows signed-rank tests to achieve f_1-parametric optimality - the following way

$$K_{f_1}(u) = \varphi_{f_1}(\tilde{F}_1^{-1}(u))(\tilde{F}_1^{-1}(u)).$$

For score functions K, K_1, K_2 satisfying these three points, let (throughout, U stands for a random variable uniformly distributed over $(0,1)$)

$$\mathcal{J}_p(K_1,K_2) := \mathrm{E}[K_1(U)K_2(U)],\ \ \mathcal{J}_p(K) := \mathcal{J}_p(K,K),\ \text{ and } \mathcal{J}_p(K,f_1) := \mathcal{J}_p(K,K_{f_1}); \tag{13}$$

note that with this notation, $\mathcal{J}_p(f_1) = \mathcal{J}_p(K_{f_1},K_{f_1})$. The power score functions $K_a(u) := p(a+1)u^a$ $(a > 0)$, with $\mathcal{J}_p(K_a) = p^2(a+1)^2/(2a+1)$, provide some traditional score functions satisfying the assumptions above: the Laplace, Wilcoxon and Spearman scores are obtained for $a = 0$, $a = 1$, and $a = 2$, respectively (the Laplace scores are also known as sign test scores). As for the score functions of the form K_{f_1}, an important particular case is that of van der Waerden or normal scores, obtained for $f_1 = \phi_1$, with ϕ_1 the Gaussian radial function. Then, denoting by Ψ_p the chi-square distribution function with p degrees of freedom, we have that

$$K_{\phi_1}(u) = \Psi_p^{-1}(u) \quad \text{and} \quad \mathcal{J}_p(\phi_1) = p(p+2).$$

Similarly, Student densities $f_1 = f^t_{1,\nu}$ (ν degrees of freedom) yield the scores

$$K_{f^t_{1,\nu}}(u) = \frac{p(p+\nu)G^{-1}_{p,\nu}(u)}{\nu + pG^{-1}_{p,\nu}(u)} \quad \text{and} \quad \mathcal{J}_p(f^t_{1,\nu}) = \frac{p(p+2)(p+\nu)}{p+\nu+2},$$

where $G_{p,\nu}$ stands for the Fisher-Snedecor distribution function with p and ν degrees of freedom. The choice of the score function will determine the type of radial densities under which the resulting tests will tend to outperform competitors. We provide some guidelines regarding this choice in the sequel. As in the Gaussian case discussed in the previous Section, some nuisance parameters will have to be estimated. In order to keep the good robustness properties of multivariate signed-rank based procedures, the classical choice is to provide estimators based on $\hat{\mathbf{V}}_{\rm Tyler}$, the shape estimator of [26], normalised so that it has determinant one. Letting $\hat{\mathbf{\Lambda}}_{\rm Tyler} = {\rm diag}(\hat{\lambda}_1, \ldots, \hat{\lambda}_p)$ and $\hat{\boldsymbol{\beta}}_{\rm Tyler} = (\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_p)$ be respectively the eigenvalues and the eigenvectors of $\hat{\mathbf{V}}_{\rm Tyler}$, a natural constrained estimator of $\mathbf{V}$ is $\hat{\mathbf{V}} = \tilde{\mathbf{V}}/{\rm det}^{1/p}(\tilde{\mathbf{V}})$, where, letting $\hat{\underline{\lambda}} := (p-q)^{-1} \sum_{j=q+1}^{p} \hat{\lambda}_j$ (a natural estimator of the common value $\underline{\lambda}$ of the $p-q$ smallest eigenvalues of the shape matrix under the null hypothesis) ,

$$\tilde{\mathbf{V}} := \sum_{j=1}^{q} \hat{\lambda}_j \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j^\top + \hat{\underline{\lambda}} \sum_{j=q+1}^{p} \hat{\boldsymbol{\beta}}_j \hat{\boldsymbol{\beta}}_j^\top .$$

Obviously, $\boldsymbol{\beta}_{0q}$ is replaced by $\hat{\boldsymbol{\beta}}_{0q} := (\hat{\boldsymbol{\beta}}_{q+1}, \ldots, \hat{\boldsymbol{\beta}}_p)$. As in [7] and [8], we will tacitly assume that $\hat{\mathbf{V}}$ is locally and asymptotically discrete in the sense that it only takes a bounded number of distinct values in balls with $O(n^{-1/2})$ radius centered at $\mathbf{V}$. Such discretization, however, is a purely technical requirement, with no practical consequences, and is only required in asymptotic statements related to signed-rank based statistics; see for instance [6]. Indeed, in practice the sample size n is fixed and we can always assume that the values taken by $\hat{\mathbf{V}}$ coincide with the values taken by some discretised version of $\hat{\mathbf{V}}$, constrained such that it takes only a possibly very large but finite number of different values in the balls of radius $O(n^{-1/2})$. Such a discretization has been explicitly conducted and its lack of consequences shown in [10], for instance. The resulting test $\phi^{(n)}_{{\rm Un},K}$ rejects the null hypothesis at the asymptotic level α when

$$S^{(n)}_{q,K} \quad = \quad \frac{np(p+2)}{2\mathcal{J}_p(K)} ({\rm tr}((\hat{\boldsymbol{\beta}}_{0q}^\top \mathbf{S}^{(n)}_K(\hat{\mathbf{V}}) \hat{\boldsymbol{\beta}}_{0q})^2) - (p-q)^{-1} {\rm tr}^2(\hat{\boldsymbol{\beta}}_{0q}^\top \mathbf{S}^{(n)}_K(\hat{\mathbf{V}}) \hat{\boldsymbol{\beta}}_{0q})) > \chi^2_{d(p,q);1-\alpha}. \tag{14}$$

The asymptotic properties of the test $\phi^{(n)}_{{\rm Un},K}$ are summarised in the following result.

Proposition 5. *Fix an arbitrary orthonormal matrix* $\boldsymbol{\beta}$ *and let* $\boldsymbol{\ell}^{(n)}$ *be a bounded sequence of* $\mathbb{R}^p$ *such that* (3) *holds with* $\lim_{n\to\infty} \boldsymbol{\ell}^{(n)} = \boldsymbol{\ell} = (\ell_1, \ldots, \ell_p)$*. Then, as* $n \to \infty$*,*

(i) $S^{(n)}_{q,K}$ *is asymptotically chi-square with* $d(p,q)$ *degrees of freedom under* $\bigcup_{g_1 \in \mathcal{F}_1} {\rm P}^{(n)}_{\sigma, \boldsymbol{\beta}, \mathbf{\Lambda}; g_1}$, $\overset{\circ}{\rm dvec}\,(\mathbf{\Lambda}) \in \mathcal{H}_{0q}$ *and*

(ii) $S^{(n)}_{q,K}$ *is asymptotically chi-square with* $d(p,q)$ *degrees of freedom and noncentrality parameter*

$$\frac{\mathcal{J}_p^2(K, g_1)}{p(p+2)\mathcal{J}_p(K)}\, s(\boldsymbol{\ell}) \tag{15}$$

under ${\rm P}^{(n)}_{\sigma, \boldsymbol{\beta}, \mathbf{\Lambda}+n^{-1/2}{\rm diag}(\boldsymbol{\ell}^{(n)}); g_1}$, $\overset{\circ}{\rm dvec}\,(\mathbf{\Lambda}) \in \mathcal{H}_{0q}$*,* $g_1 \in \mathcal{F}_1$*, with* $s(\boldsymbol{\ell})$ *defined in* (10).

Now, it follows along the same lines as in Propositions 3 and 5 that under ${\rm P}^{(n)}_{\sigma, \boldsymbol{\beta}, \mathbf{\Lambda}+n^{-1/2}{\rm diag}(\boldsymbol{\ell}^{(n)}); g_1}$, $g_1 \in \mathcal{F}_4$, $S^{(n)\dagger}_{q,\phi_1}$ in (11) converges weakly to a chi-square random variable with $d(p,q)$ degrees of freedom and noncentrality parameter

$$(1+\kappa_p(g_1))^{-1} s(\boldsymbol{\ell}). \tag{16}$$

It directly follows from (15) and (16) that the asymptotic relative efficiency (under ${\rm P}^{(n)}_{\sigma, \boldsymbol{\beta}, \mathbf{\Lambda}+n^{-1/2}{\rm diag}(\boldsymbol{\ell}^{(n)}); g_1}$, $g_1 \in \mathcal{F}_4$) of the signed rank-based test $\phi^{(n)}_{{\rm Un},K}$ with respect to $\phi^{\dagger}_{{\rm Un},\phi_1}$ is given by

$${\rm ARE}_{p,g_1}(\phi^{(n)}_{{\rm Un},K} / \phi^{\dagger}_{{\rm Un},\phi_1}) = \frac{(1+\kappa_p(g_1))\mathcal{J}_p^2(K, g_1)}{p(p+2)\mathcal{J}_p(K)}.$$

Numerical values of these AREs, for various values of the dimension p and selected radial densities g_1 (Student, Gaussian, and power-exponential), are provided for the van der Waerden test with $K = K_{\phi_1}$, the Laplace, the Wilcoxon and the Spearman tests (the score functions were defined below (13)) in Table 1. These values coincide with the "AREs for shape" obtained in [6]. In particular, as shown in [19], we have that

$$\inf_{g_1 \in \mathcal{F}_1} {\rm ARE}_{p,g_1}(\phi^{(n)}_{{\rm Un},K_{\phi_1}} / \phi^{\dagger}_{{\rm Un},\phi_1}) = 1; \tag{17}$$

Table 1: AREs of the van der Waerden (vdW), Laplace (L), Wilcoxon (W), and Spearman (SP) rank-based test $\phi^{(n)}_{\mathrm{Un},K}$ with respect to the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$, under p-dimensional ($p = 2, 3, 4, 6, 10$ and $p \to \infty$) Student (with 5, 8, and 12 degrees of freedom), normal, and power-exponential densities (with parameter $\eta \in \{2, 3, 5\}$), for $p \in \{2, 3, 4, 6, 10\}$, and $p \to \infty$.

		underlying density						
K	p	t_5	t_8	t_{12}	$\mathcal{N}$	e_2	e_3	e_5
vdW	2	2.204	1.215	1.078	1.000	1.129	1.308	1.637
	3	2.270	1.233	1.086	1.000	1.108	1.259	1.536
	4	2.326	1.249	1.093	1.000	1.093	1.223	1.462
	6	2.413	1.275	1.106	1.000	1.072	1.174	1.363
	10	2.531	1.312	1.126	1.000	1.050	1.121	1.254
	∞	3.000	1.500	1.250	1.000	1.000	1.000	1.000
L	2	1.500	0.750	0.625	0.500	0.392	0.365	0.347
	3	1.800	0.900	0.750	0.600	0.493	0.464	0.444
	4	2.000	1.000	0.833	0.667	0.565	0.537	0.517
	6	2.250	1.125	0.938	0.750	0.662	0.636	0.617
	10	2.500	1.250	1.041	0.833	0.766	0.746	0.730
	∞	3.000	1.500	1.250	1.000	1.000	1.000	1.000
W	2	2.301	1.230	1.067	0.934	0.965	1.042	1.168
	3	2.277	1.225	1.070	0.957	1.033	1.141	1.317
	4	2.432	1.273	1.094	0.945	0.955	1.006	1.095
	6	2.451	1.283	1.105	0.969	1.008	1.075	1.188
	10	2.426	1.264	1.088	0.970	1.032	1.106	1.233
	∞	2.250	1.125	0.938	0.750	0.750	0.750	0.750
SP	2	2.301	1.230	1.067	0.934	0.965	1.042	1.168
	3	2.277	1.225	1.070	0.957	1.033	1.141	1.317
	4	2.225	1.200	1.051	0.956	1.057	1.179	1.383
	6	2.128	1.146	1.007	0.936	1.057	1.189	1.414
	10	2.001	1.068	0.936	0.891	1.017	1.144	1.365
	∞	1.667	0.833	0.694	0.556	0.556	0.556	0.556

that is, the van der Waerden version of our rank tests asymptotically dominates (uniformly) the pseudo-Gaussian test of [29], with the local and asymptotic powers of the two tests coinciding only under Gaussian assumption.

We now discuss the choice of the score function K. If the data looks Gaussian we recommend to choose the score function K_{ϕ_1}; following (17), the van der Waerden test asymptotically dominates the pseudo-Gaussian test. The power score functions K_a also yield to tests with nice properties. Indeed, inspection of Table 1 reveals that choosing K_a with small a will lead to tests that tend to perform better than the competitors under heavy tails while for large values of a, the resulting tests will tend to outperform the competitors under distributions with lighter tails. The sign tests (with score function K_0) in particular should be used when the data looks to be strongly heavy-tailed.

Finally note that in the (very uncommon) case in which nuisance parameters are specified, a rank-based equivalent $\phi^{(n)}_{\mathrm{Sp},K}$ to $\phi^{(n)}_{\mathrm{Sp},f_1}$ in (6) rejects the null hypothesis at the asymptotic nominal level α when

$$S^{(n)}_{q,K}(\mathbf{V}) = \frac{np(p+2)}{2\mathcal{J}_p(K)}(\mathrm{tr}(\mathrm{diag}^2(\boldsymbol{\beta}_{0q}^\top \mathbf{S}^{(n)}_K(\mathbf{V})\boldsymbol{\beta}_{0q})) - (p-q)^{-1}\mathrm{tr}^2(\mathrm{diag}(\boldsymbol{\beta}_{0q}^\top \mathbf{S}^{(n)}_K(\mathbf{V})\boldsymbol{\beta}_{0q}))) > \chi^2_{p-q-1;1-\alpha}.$$

It can be shown that $\phi^{(n)}_{\mathrm{Sp},K_{\phi_1}}$ also uniformly dominates its pseudo-Gaussian competitor based on $\phi^{(n)}_{\mathrm{Sp},\phi_1}$ (using the same kind of construction as above).

5. Simulation study

We conclude the paper with a simulation study. We generated $M = 1,000$ independent samples of i.i.d. observations

$$\mathbf{X}_1^{(m,\tau)}, \ldots, \mathbf{X}_n^{(m,\tau)},$$

$\tau \in \{0, 1, 2, 4, 6, 8, 10\}$ and $m \in \{1, 2, \ldots, 6\}$. The $\mathbf{X}_i^{(1,\tau)}$'s are i.i.d. with a common (p =)5-dimensional centered Gaussian distribution with scatter matrix

$$\boldsymbol{\Sigma}_1(\tau) = \mathrm{diag}(3, 2, 2, 1, 1 - \frac{\tau}{\sqrt{n}}),$$

the $\mathbf{X}_i^{(2,\tau)}$'s are i.i.d. with a common (p =)5-dimensional centered Student distribution with 5 degrees of freedom and scatter matrix $\boldsymbol{\Sigma}_1(\tau)$ while the $\mathbf{X}_i^{(3,\tau)}$'s are i.i.d. with a common (p =)5-dimensional centered Student distribution with 1 degree of freedom and scatter matrix $\boldsymbol{\Sigma}_1(\tau)$. Therefore for $m \in \{1, 2, 3\}$, the value $\tau = 0$ provides data generating processes that belong to the null hypothesis $\mathcal{H}_{03} : \lambda_3 > \lambda_4 = \lambda_5$ while for $\tau \in \{1, 2, 4, 6, 8, 10\}$, the corresponding distributions are increasingly under the alternative. The $\mathbf{X}_i^{(4,\tau)}$'s are i.i.d. with a common (p =)5-dimensional centered Gaussian distribution with scatter matrix

$$\boldsymbol{\Sigma}_2(\tau) = \mathrm{diag}(3, 1, 1, 1, 1 - \frac{\tau}{\sqrt{n}}),$$

the $\mathbf{X}_i^{(5,\tau)}$'s are i.i.d. with a common (p =)5-dimensional centered Student distribution with 5 degrees of freedom and scatter matrix $\boldsymbol{\Sigma}_2(\tau)$ while the $\mathbf{X}_i^{(6,\tau)}$'s are i.i.d. with a common (p =)5-dimensional centered Student distribution with 1 degree of freedom and scatter matrix $\boldsymbol{\Sigma}_2(\tau)$. Therefore, for $m \in \{4, 5, 6\}$, the value $\tau = 0$ provides data generating processes that belong to the null hypothesis $\mathcal{H}_{01} : \lambda_1 > \lambda_2 = \cdots = \lambda_5$ while for $\tau \in \{1, 2, 4, 6, 8, 10\}$, the corresponding distributions are again increasingly under the alternative.

For $m \in \{1, 2, 3\}$, we estimated the location parameter using the affine equivariant median of [9] and centered the observations first, then we performed the following three tests for $\mathcal{H}_{03}$ at the nominal asymptotic level $\alpha = .05$: the pseudo-gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$, the van der Waerden-rank test $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$, and the sign test $\phi^{(n)}_{\mathrm{Un},K_0}$. We provide in Figures 3, 4, 5 and 6, the empirical rejection frequencies of the three tests as functions of τ for $n \in \{200, 500, 2000\}$ and $n = 5000$ respectively. For $m \in \{4, 5, 6\}$ we performed the same three tests (for $\mathcal{H}_{01}$ in this case) still at the same asymptotic nominal level $\alpha = .05$. We similarly provide in Figures 7, 8, 9 and 10, the empirical rejection frequencies of the three tests as functions of τ for $n \in \{200, 500, 2000\}$ and $n = 5000$ respectively.

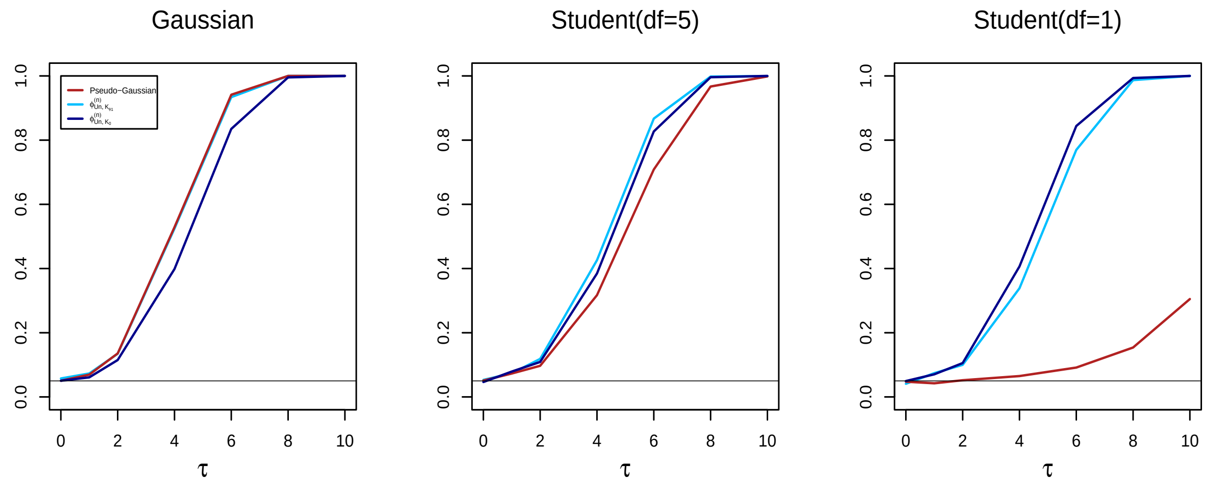

Fig. 3: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 200$ for $\mathcal{H}_{03}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

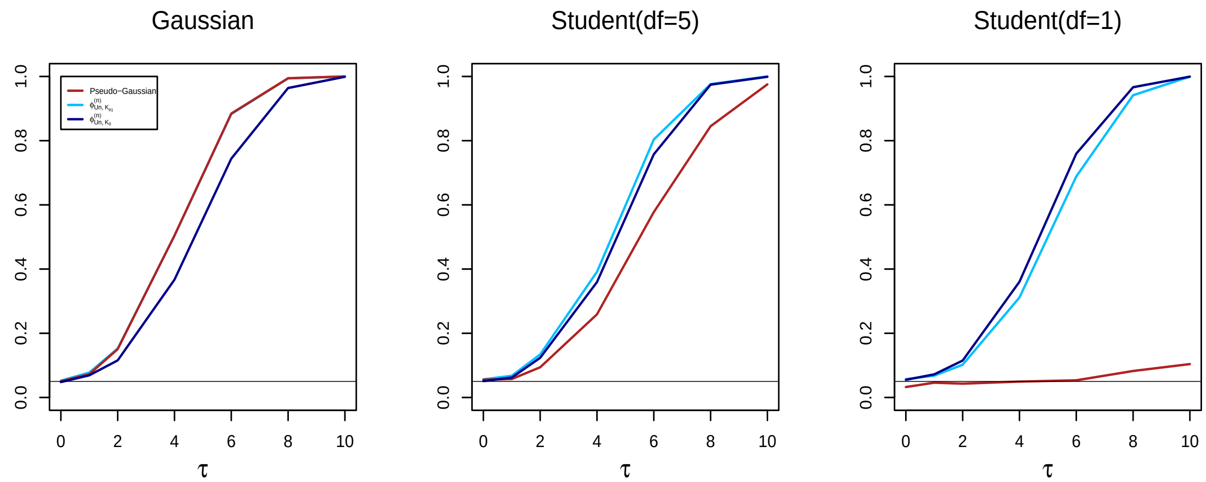

Fig. 4: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 500$ for $\mathcal{H}_{03}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

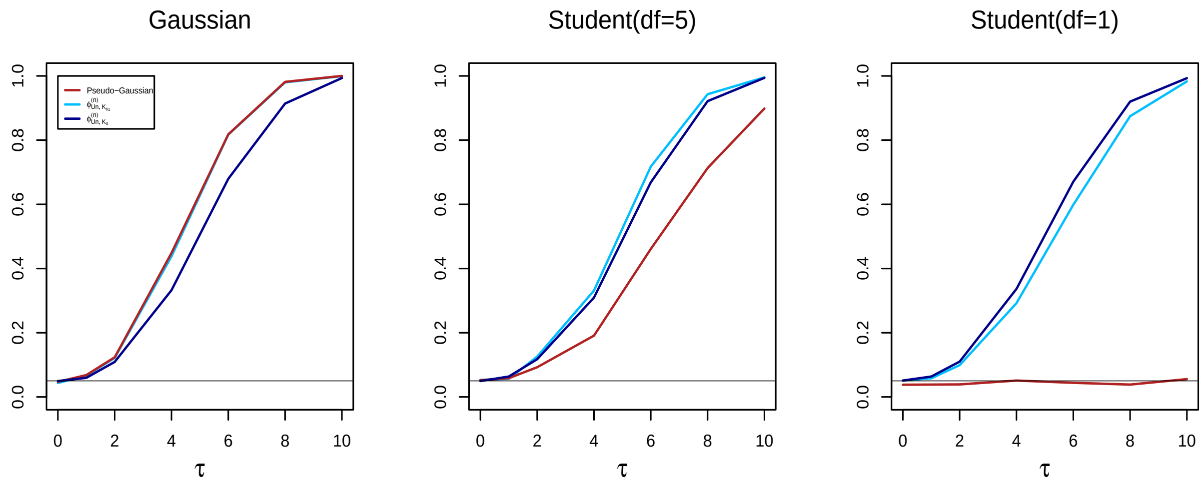

Fig. 5: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 2000$ for $\mathcal{H}_{03}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

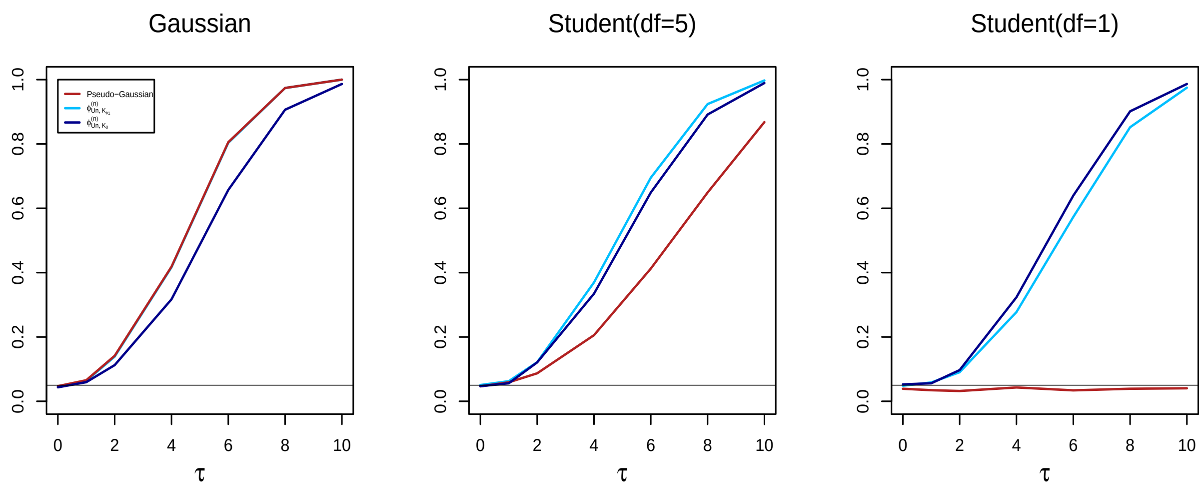

Fig. 6: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 5000$ for $\mathcal{H}_{03}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

Inspection of the various Figures clearly reveals that the signed-rank based tests we propose are robust to heavy tailed distributions while retaining high local powers. In particular, the van der Waerden test dominates the pseudo-gaussian competitor in every scenario as expected. Under very heavy tails, the pseudo-Gaussian procedure completely fails to be valid. Moreover, as expected, the location parameter $\boldsymbol{\mu}$ can indeed be estimated at no asymptotic cost.

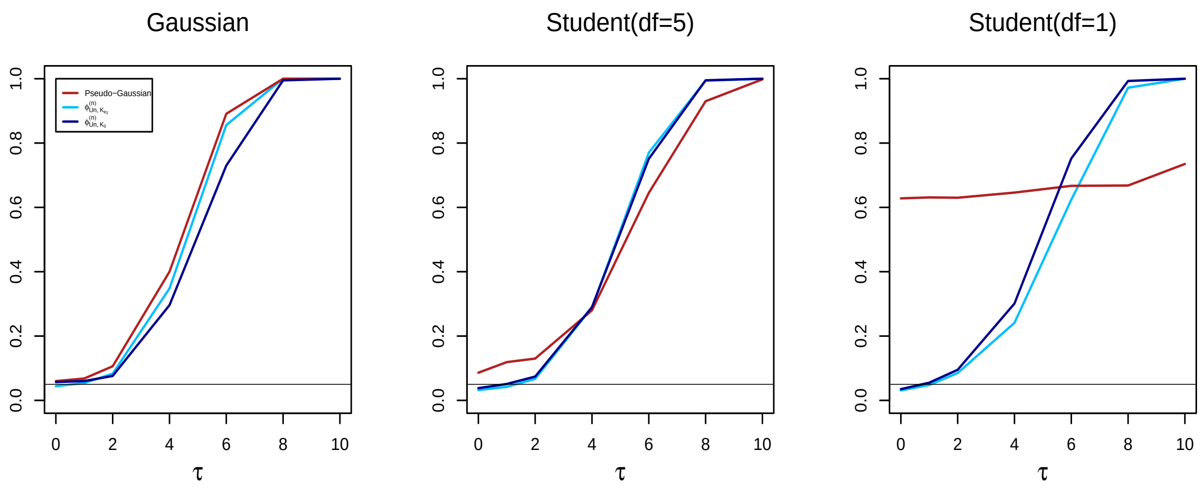

Fig. 7: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 200$ for $\mathcal{H}_{01}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

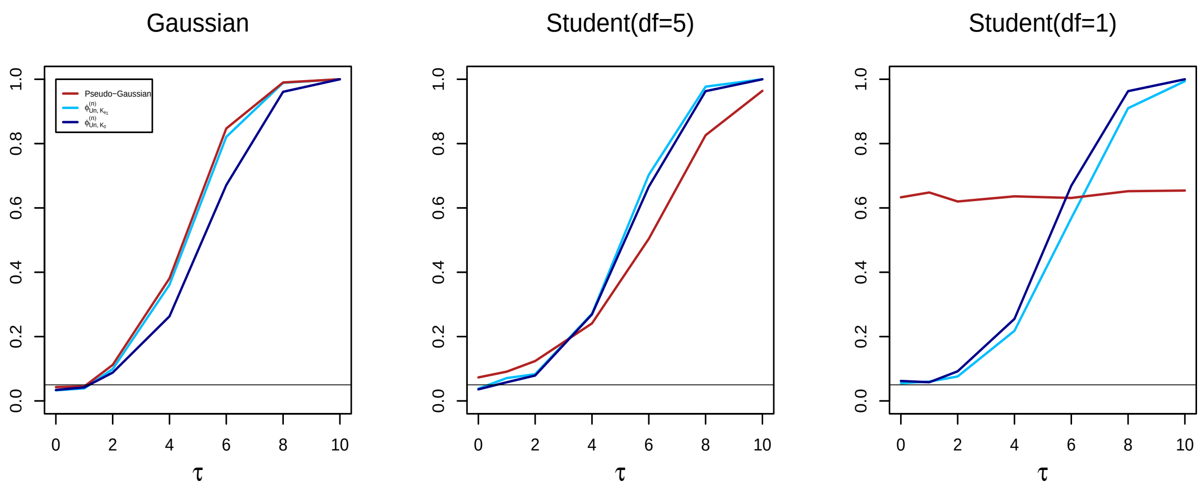

Fig. 8: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 500$ for $\mathcal{H}_{01}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

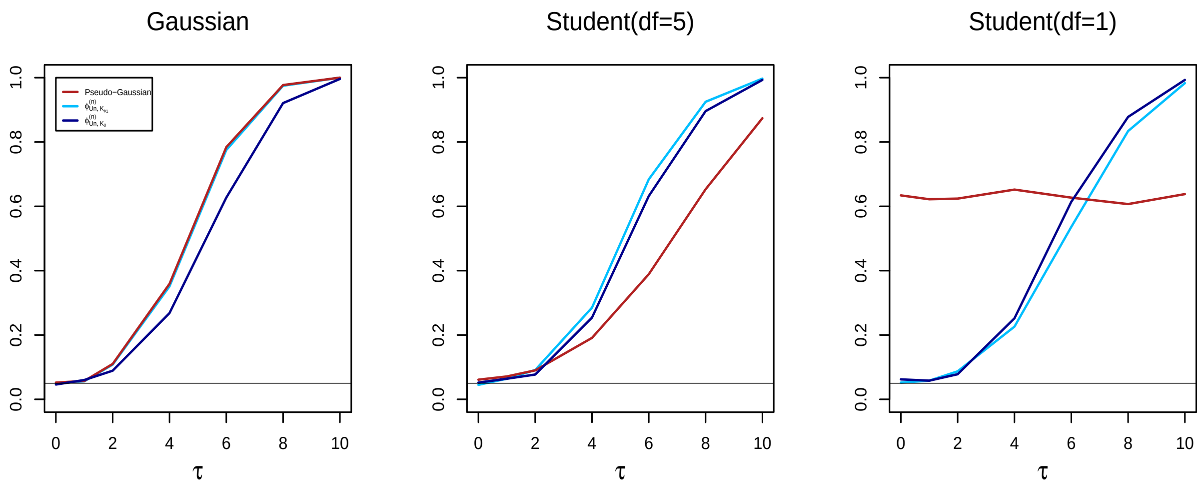

Fig. 9: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 2000$ for $\mathcal{H}_{01}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

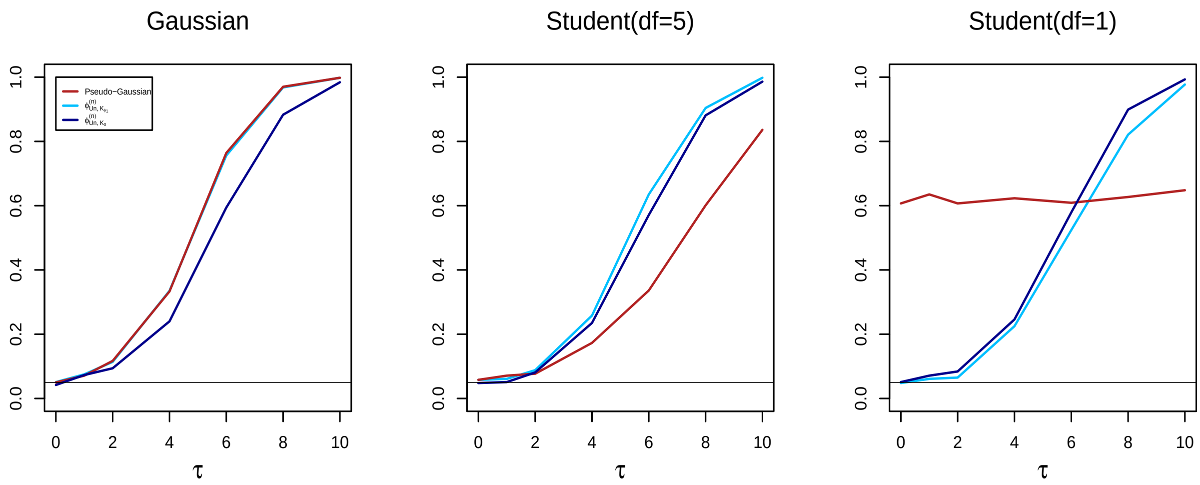

Fig. 10: Empirical powers of the pseudo-Gaussian test $\phi^{\dagger}_{\mathrm{Un},\phi_1}$ (in red), the signed-rank test (with Gaussian score function) $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ (in light blue) and the sign test (with a constant score function) $\phi^{(n)}_{\mathrm{Un},K_0}$ (in dark blue) performed with $n = 5000$ for $\mathcal{H}_{01}$ under various radial densities (Gaussian and Student). The three considered tests are defined in (11) and (14).

6. Conclusions

The two main contributions of the present paper are the following. First we showed that testing $\mathcal{H}_{0q} : \lambda_q > \lambda_{q+1} = \cdots = \lambda_p$ is quite different from the testing problems over the eigenvalues studied in [7]. The difference is that the estimation of the nuisance parameters have an asymptotic cost for this problem. We rigorously quantified this asymptotic cost in the Gaussian case and showed that it only depends on the dimensions p and q. Second, we proposed two new types of robust tests $\phi^{(n)}_{\mathrm{Sp},K}$ and $\phi^{(n)}_{\mathrm{Un},K}$ that are asymptotically valid under elliptical assumption for testing $\mathcal{H}_{0q}$. We showed that the tests $\phi^{(n)}_{\mathrm{Sp},K_{f_1}}$ are locally and asymptotically optimal under radial function f_1. On the other hand, we obtained that the van der Waerden rank test $\phi^{(n)}_{\mathrm{Un},K_{\phi_1}}$ still outperforms the existing pseudo-Gaussian competitor in any possible (elliptical) scenario. The van der Waerden test therefore shows a very attractive mixture of robustness and asymptotic efficiency properties. Although the eigenvectors associated with the $p-q$ smallest eigenvalues under the null hypothesis are not identified, the corresponding subspace is. We believe that it could be interesting to see how the corresponding projection onto the $(p-q)$-dimensional subspace could be locally perturbed to obtain the asymptotic behavior of log-likelihood ratios with perturbed eigenvalues and eigenspaces. We leave this for future work.

Acknowledgments

Thomas Verdebout's research is supported by an ARC grant of the Communauté Française de Belgique and a Projet de Recherche (PDR) from the Fonds National de la Recherche Scientifique (FNRS). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 811017.

Appendix

Proof of Proposition 1: First note that $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\,\boldsymbol{\Lambda}+n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)});f_1} \equiv \mathrm{P}^{(n)}_{\sigma,\mathbf{V}^{(n)};f_1}$, where

$$\mathbf{V}^{(n)} := \mathbf{V} + n^{-1/2}\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top = \boldsymbol{\beta}\boldsymbol{\Lambda}\boldsymbol{\beta}^\top + n^{-1/2}\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top.$$

Therefore, it directly follows from the LAN property (Proposition 2.1) in [6] on the shape parameter $\mathbf{V}$ that letting $\Lambda_n := \ln\Big(\frac{d\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\,\boldsymbol{\Lambda}+n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)});f_1}}{d\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\,\boldsymbol{\Lambda};f_1}}\Big)$,

$$\Lambda_n = (\overset{\circ}{\mathrm{vech}}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top))^\top\tilde{\boldsymbol{\Delta}}^{(n)}_{f_1}(\mathbf{V}) - \frac{1}{2}(\overset{\circ}{\mathrm{vech}}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top))^\top\tilde{\boldsymbol{\Gamma}}_{f_1}(\mathbf{V})\overset{\circ}{\mathrm{vech}}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top) + o_{\mathrm{P}}(1),$$

where $\mathrm{vech}\,(\mathbf{A}) = (A_{11}, \overset{\circ}{\mathrm{vech}}{}^\top(\mathbf{A}))^\top$ is the $p(p+1)/2$-dimensional vector obtained by stacking the upper triangular elements of a $p\times p$ matrix $\mathbf{A}$ and where letting $\mathbf{N}_p$ be the matrix such that (i) $\mathbf{N}_p\mathrm{vec}\,(\mathbf{V}^{-1}) = \mathbf{0}$ and (ii) $\mathbf{N}_p^\top\overset{\circ}{\mathrm{vech}}\,(\mathbf{W}) = \mathrm{vec}\,(\mathbf{W})$ for any symmetric matrix $\mathbf{W}$ such that $\mathrm{tr}(\mathbf{V}^{-1}\mathbf{W}) = 0$,

$$\tilde{\boldsymbol{\Delta}}^{(n)}_{f_1}(\mathbf{V}) := \frac{1}{2}\mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}))$$

and

$$\tilde{\boldsymbol{\Gamma}}^{(n)}_{f_1}(\mathbf{V}) := \frac{\mathcal{J}_p(f_1)}{4p(p+2)}\mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}(\mathbf{I}_{p^2} + \mathbf{K}_p)(\mathbf{V}^{-1/2})^{\otimes 2}\mathbf{N}_p^\top.$$

First note that the definition of $\boldsymbol{\ell}^{(n)}$ directly entails that $\mathrm{tr}(\mathbf{V}^{-1}\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top) = \mathrm{tr}(\boldsymbol{\Lambda}^{-1}\mathrm{diag}(\boldsymbol{\ell}^{(n)})) = o(1)$ as $n\to\infty$. Therefore we easily obtain that

$$\mathbf{N}_p^\top\overset{\circ}{\mathrm{vech}}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top) = \mathrm{vec}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top) + o(1) \tag{18}$$

and

$$\mathbf{M}_p^\top\overset{\circ}{\mathrm{dvec}}\,(\mathrm{diag}(\boldsymbol{\ell}^{(n)})) = \mathrm{dvec}\,(\mathrm{diag}(\boldsymbol{\ell}^{(n)})) + o(1) \tag{19}$$

as $n \to \infty$. Moreover since $\mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})] = \mathbf{I}_p$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};f_1}$, we have that

$$\mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}\mathrm{vec}\,(\mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})]) = \mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}\mathrm{vec}\,(\mathbf{I}_p) = \mathbf{N}_p\mathrm{vec}\,(\mathbf{V}^{-1}) = \mathbf{0}. \tag{20}$$

Using (20), we have that

$$\mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\mathbf{S}^{(n)}_{f_1}(\mathbf{V})) = \mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})]),$$

where $n^{1/2}\mathrm{vec}\,(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})])$ is $O_\mathrm{P}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};f_1}$. Using (18) and standard algebra properties, we therefore have that

$$\begin{aligned}(\mathrm{ve\mathring{c}h}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top))^\top\tilde{\boldsymbol{\Delta}}^{(n)}_{f_1}(\mathbf{V}) &= \frac{1}{2}(\mathrm{ve\mathring{c}h}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top))^\top\mathbf{N}_p(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}))\\ &= \frac{1}{2}\mathrm{vec}^\top(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top)(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})]) + o_\mathrm{P}(1)\\ &= \frac{1}{2}\mathrm{vec}^\top(\mathrm{diag}(\boldsymbol{\ell}^{(n)}))n^{1/2}\mathrm{vec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})])\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_\mathrm{P}(1),\end{aligned} \tag{21}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};f_1}$. Now the definitions of $\mathbf{H}_p$ and $\mathbf{M}_p$ together with (19) directly entail that (21) rewrites

$$\begin{aligned}(\mathrm{ve\mathring{c}h}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top))^\top\tilde{\boldsymbol{\Delta}}^{(n)}_{f_1}(\mathbf{V}) &= \frac{1}{2}\mathrm{dvec}^\top(\mathrm{diag}(\boldsymbol{\ell}^{(n)}))\mathbf{H}_p n^{1/2}\mathrm{vec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})])\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_\mathrm{P}(1)\\ &= \frac{1}{2}\mathrm{dvec}^\top(\mathrm{diag}(\boldsymbol{\ell}^{(n)}))n^{1/2}\mathrm{dvec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top(\mathbf{S}^{(n)}_{f_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{f_1}(\mathbf{V})])\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_\mathrm{P}(1)\\ &= \mathrm{d\mathring{v}ec}^\top(\mathrm{diag}(\boldsymbol{\ell}^{(n)}))\frac{1}{2}\mathbf{M}_p n^{1/2}\mathrm{dvec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top\mathbf{S}^{(n)}_{f_1}(\mathbf{V})\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_\mathrm{P}(1)\\ &= \boldsymbol{\tau}^\top\boldsymbol{\Delta}^{(n)}_{f_1}(\mathbf{V}) + o_\mathrm{P}(1)\end{aligned}$$

as $n \to \infty$. It follows along the same lines using (18) and (19) together with the fact that $\mathbf{H}_p\mathbf{K}_p = \mathbf{H}_p$ that

$$(\mathrm{ve\mathring{c}h}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top))^\top\tilde{\boldsymbol{\Gamma}}_{f_1}(\mathbf{V})\mathrm{ve\mathring{c}h}\,(\boldsymbol{\beta}\mathrm{diag}(\boldsymbol{\ell}^{(n)})\boldsymbol{\beta}^\top) = \boldsymbol{\tau}^\top\boldsymbol{\Gamma}_{f_1}(\mathbf{V})\boldsymbol{\tau} + o(1)$$

as $n \to \infty$. The result follows. □

Proof of Proposition 2: First define ($\boldsymbol{\beta}_\mathbf{S} := (\boldsymbol{\beta}_{1;\mathbf{S}}, \ldots, \boldsymbol{\beta}_{p;\mathbf{S}})$)

$$\mathbf{E} = \begin{pmatrix}\mathbf{E}_{11} & \mathbf{E}_{12}\\ \mathbf{E}_{21} & \mathbf{E}_{22}\end{pmatrix} := \boldsymbol{\beta}_\mathbf{S}^\top\,\boldsymbol{\beta},$$

where $\mathbf{E}_{11}$ and $\mathbf{E}_{22}$ are respectively the $q \times q$ upper diagonal and $(p-q)\times(p-q)$ lower diagonal blocks of $\mathbf{E}$. In this proof we also put

$$\boldsymbol{\Lambda}_\mathbf{S} := \mathrm{diag}(\lambda_{1;\mathbf{S}}, \ldots \lambda_{p;\mathbf{S}}) = \mathrm{diag}(\boldsymbol{\Lambda}_{1,\mathbf{S}}, \boldsymbol{\Lambda}_{2,\mathbf{S}}),$$

where $\boldsymbol{\Lambda}_{1,\mathbf{S}}$ is the $q\times q$ diagonal matrix with the q largest roots of $\mathbf{S}^{(n)}$ as diagonal elements and $\boldsymbol{\Lambda}_{2,\mathbf{S}}$ the $(p-q)\times(p-q)$ diagonal matrix with the $p-q$ smallest roots of $\mathbf{S}^{(n)}$. Note that directly follows from [1] that $n^{1/2}(\mathbf{E}_{22}\mathbf{E}_{22}^\top - \mathbf{I}_{p-q}) = o_\mathrm{P}(1)$ as $n \to \infty$ and that $\mathbf{E}_{21}$ and $\mathbf{E}_{12}$ are both $O_\mathrm{P}(n^{-1/2})$ as $n \to \infty$. Now, we have that

$$\boldsymbol{\beta}_{0q}^\top\mathbf{S}^{(n)}\boldsymbol{\beta}_{0q} = (\mathbf{E}_{12}^\top\boldsymbol{\Lambda}_{1,\mathbf{S}}\mathbf{E}_{12} + \mathbf{E}_{22}^\top\boldsymbol{\Lambda}_{2,\mathbf{S}}\mathbf{E}_{22})$$

so that since $\mathbf{E}_{12}^\top\mathbf{E}_{12} + \mathbf{E}_{22}^\top\mathbf{E}_{22} = \mathbf{I}_{p-q}$ ($\underline{\lambda}_{\rm Cov}$ denotes the common value of the $p-q$ smallest roots of the underlying covariance matrix under the null hypothesis),

$$\begin{aligned} n^{1/2}(\boldsymbol{\beta}_{0q}^\top\mathbf{S}^{(n)}\boldsymbol{\beta}_{0q} - \underline{\lambda}_{\rm Cov}\mathbf{I}_{p-q}) &= n^{1/2}(\mathbf{E}_{12}^\top(\boldsymbol{\Lambda}_{1,\mathbf{S}} - \underline{\lambda}_{\rm Cov}\mathbf{I}_q)\mathbf{E}_{12} + \mathbf{E}_{22}^\top(\boldsymbol{\Lambda}_{2,\mathbf{S}} - \underline{\lambda}_{\rm Cov}\mathbf{I}_{p-q})\mathbf{E}_{22}) \\ &= n^{1/2}\mathbf{E}_{22}^\top(\boldsymbol{\Lambda}_{2,\mathbf{S}} - \underline{\lambda}_{\rm Cov}\mathbf{I}_{p-q})\mathbf{E}_{22} + o_{\rm P}(1) \end{aligned}$$

as $n \to \infty$; the last equality follows easily from the fact that $n^{1/2}\mathbf{E}_{12}$ is $O_{\rm P}(1)$ as $n \to \infty$. Now since $n^{1/2}(\boldsymbol{\Lambda}_{2,\mathbf{S}} - \underline{\lambda}_{\rm Cov}\mathbf{I}_{p-q})$ is $O_{\rm P}(1)$ as $n \to \infty$ under the null hypothesis and $\mathbf{E}_{22}\mathbf{E}_{22}^\top = \mathbf{I}_{p-q} + o_{\rm P}(1)$ as $n \to \infty$, the Slutsky Lemma and standard computations entail that

$$\begin{aligned} T_{q,\phi_1}^{(n)}(\mathbf{V}) &= \frac{na_k^2}{2\sigma^4\underline{\lambda}^2}(\mathrm{tr}((\boldsymbol{\beta}_{0q}^\top\mathbf{S}^{(n)}\boldsymbol{\beta}_{0q})^2) - (p-q)^{-1}\mathrm{tr}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}^{(n)}\boldsymbol{\beta}_{0q})) \\ &= \frac{a_k^2}{2\sigma^4\underline{\lambda}^2}(\mathrm{tr}((n^{1/2}\mathbf{E}_{22}^\top(\boldsymbol{\Lambda}_{2,\mathbf{S}} - \underline{\lambda}_{\rm Cov}\mathbf{I}_{p-q})\mathbf{E}_{22})^2) - (p-q)^{-1}\mathrm{tr}^2(n^{1/2}\mathbf{E}_{22}^\top(\boldsymbol{\Lambda}_{2,\mathbf{S}} - \underline{\lambda}_{\rm Cov}\mathbf{I}_{p-q})\mathbf{E}_{22})) + o_{\rm P}(1) \\ &= \frac{a_k^2}{2\sigma^4\underline{\lambda}^2}(\mathrm{tr}((n^{1/2}\boldsymbol{\Lambda}_{2,\mathbf{S}})^2) - (p-q)^{-1}\mathrm{tr}^2(n^{1/2}\boldsymbol{\Lambda}_{2,\mathbf{S}})) + o_{\rm P}(1) = S_{q,\phi_1}^{(n)} + o_{\rm P}(1) \end{aligned}$$

as $n \to \infty$, where the last equality follows from the fact that $(p-q)^{-1}\sum_{j=q+1}^{p}\lambda_{j;\mathbf{S}}$ is a consistent estimator of $a_k/\sigma^2\underline{\lambda}$. □

Proof of Proposition 3: Point (ii) has been obtained in [27] while point (i) follows directly from a direct application of the Third Le Cam Lemma based on Proposition 1. □

Proof of Proposition 4: Combining Theorem A.1. in [18] and Proposition 3, we get the result. □

Proof of Proposition 5: First note that

$$\begin{aligned} S_{q,K}^{(n)} &= \frac{np(p+2)}{2\mathcal{J}_p(K)}(\mathrm{tr}((\hat{\boldsymbol{\beta}}_{0q}^\top\mathbf{S}_K^{(n)}(\hat{\mathbf{V}})\hat{\boldsymbol{\beta}}_{0q})^2) - (p-q)^{-1}\mathrm{tr}^2(\hat{\boldsymbol{\beta}}_{0q}^\top\mathbf{S}_K^{(n)}(\hat{\mathbf{V}})\hat{\boldsymbol{\beta}}_{0q})) \\ &= \frac{p(p+2)}{2\mathcal{J}_p(K)}n^{1/2}\mathrm{vec}^\top(\hat{\boldsymbol{\beta}}_{0q}^\top\mathbf{S}_K^{(n)}(\hat{\mathbf{V}})\hat{\boldsymbol{\beta}}_{0q})(\mathbf{I}_{(p-q)^2} - \frac{1}{p-q}\mathbf{J}_{p-q})n^{1/2}\mathrm{vec}(\hat{\boldsymbol{\beta}}_{0q}^\top\mathbf{S}_K^{(n)}(\hat{\mathbf{V}})\hat{\boldsymbol{\beta}}_{0q}) \\ &= \frac{p(p+2)}{2\mathcal{J}_p(K)}\Big(\mathbf{T}_k^{(n)}(\hat{\mathbf{V}})\Big)^\top\boldsymbol{\Sigma}_{p,q}\mathbf{T}_k^{(n)}(\hat{\mathbf{V}}), \end{aligned}$$

where $\boldsymbol{\Sigma}_{p,q} := \mathbf{I}_{(p-q)^2} - \frac{1}{p-q}\mathbf{J}_{p-q}$ and

$$\mathbf{T}_K^{(n)}(\hat{\mathbf{V}}) := n^{-1/2}\sum_{i=1}^{n}K\Big(\frac{R_i^{(n)}(\hat{\mathbf{V}})}{n+1}\Big)\mathrm{vec}(\hat{\boldsymbol{\beta}}_{0q}^\top(\mathbf{U}_i(\hat{\mathbf{V}})\mathbf{U}_i(\hat{\mathbf{V}})^\top - p^{-1}\mathbf{I}_p)\hat{\boldsymbol{\beta}}_{0q});$$

the last equality follows from the fact that for any real a, $a\,\boldsymbol{\Sigma}_{p,q}\mathrm{vec}(\mathbf{I}_{p-q}) = \mathbf{0}$. In the rest of this proof, we use the notation

$$\mathbf{E} = \begin{pmatrix}\mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22}\end{pmatrix} := \hat{\boldsymbol{\beta}}^\top\boldsymbol{\beta},$$

where $\mathbf{E}_{11}$ and $\mathbf{E}_{22}$ are respectively the $q \times q$ upper diagonal and $(p-q)\times(p-q)$ lower diagonal blocks of $\mathbf{E}$. Note that it follows as in [1] that $n^{1/2}(\mathbf{E}_{22}\mathbf{E}_{22}^\top - \mathbf{I}_{p-q}) = o_{\rm P}(1)$ as $n \to \infty$ and that $\mathbf{E}_{21}$ and $\mathbf{E}_{12}$ are both $O_{\rm P}(n^{-1/2})$ as $n \to \infty$.

Now we have that

$$\begin{aligned}\mathbf{T}_K^{(n)}(\hat{\mathbf{V}}) &= n^{-1/2}\sum_{i=1}^n K\Big(\frac{R_i^{(n)}(\hat{\mathbf{V}})}{n+1}\Big)\mathrm{vec}(\hat{\boldsymbol{\beta}}_{0q}^\top(\mathbf{U}_i(\hat{\mathbf{V}})\mathbf{U}_i(\hat{\mathbf{V}})^\top - p^{-1}\mathbf{I}_p)\hat{\boldsymbol{\beta}}_{0q})\\ &= (\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)n^{-1/2}\sum_{i=1}^n K\Big(\frac{R_i^{(n)}(\hat{\mathbf{V}})}{n+1}\Big)\mathrm{vec}(\mathbf{U}_i(\hat{\mathbf{V}})\mathbf{U}_i(\hat{\mathbf{V}})^\top - p^{-1}\mathbf{I}_p)\end{aligned}$$

so that since from (A2) in [8],

$$(\mathbf{I}_{p^2} - \frac{1}{p}\mathbf{J}_p)n^{1/2}\,\mathrm{vec}\,(\mathbf{S}_K^{(n)}(\hat{\mathbf{V}}) - \mathbf{S}_K^{(n)}(\mathbf{V})) + \frac{\mathcal{J}_p(K,g_1)}{4p(p+2)}\Big[\mathbf{I}_{p^2} + \mathbf{K}_p - \frac{2}{p}\mathbf{J}_p\Big](\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}} - \mathbf{V}) = o_{\mathrm{P}}(1) \tag{22}$$

as $n\to\infty$, under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};g_1}$. Now since $\boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)\mathbf{J}_p = \mathbf{0}$, it follows from (22) and the Slutzky Lemma that

$$\begin{aligned}\boldsymbol{\Sigma}_{p,q}\mathbf{T}_K^{(n)}(\hat{\mathbf{V}}) &= \boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)n^{1/2}\mathrm{vec}(\mathbf{S}_K^{(n)}(\hat{\mathbf{V}}))\\ &= \boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)n^{1/2}\mathrm{vec}(\mathbf{S}_K^{(n)}(\mathbf{V})) - \frac{\mathcal{J}_p(K,g_1)}{4p(p+2)}\boldsymbol{\Sigma}_{p,q}\mathbf{W}_n + o_{\mathrm{P}}(1)\end{aligned} \tag{23}$$

as $n\to\infty$, where

$$\mathbf{W}_n = (\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)\Big[\mathbf{I}_{p^2} + \mathbf{K}_p - \frac{2}{p}\mathbf{J}_p\Big](\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}} - \mathbf{V}).$$

We have that

$$\begin{aligned}\boldsymbol{\Sigma}_{p,q}\mathbf{W}_n &= 2\boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)(\mathbf{V}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}} - \mathbf{V})\\ &= 2\boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)(\hat{\mathbf{V}}^{-1/2})^{\otimes 2}n^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}} - \mathbf{V}) + o_{\mathrm{P}}(1)\\ &= 2\hat{\underline{\lambda}}^{-1}\boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)n^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}} - \mathbf{V}) + o_{\mathrm{P}}(1)\\ &= -2\hat{\underline{\lambda}}^{-1}\boldsymbol{\Sigma}_{p,q}n^{1/2}\mathrm{vec}\,(\hat{\boldsymbol{\beta}}_{0q}^\top\mathbf{V}\hat{\boldsymbol{\beta}}_{0q}) + o_{\mathrm{P}}(1)\end{aligned}$$

so that putting $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1,\dots\lambda_p) = \mathrm{diag}(\boldsymbol{\Lambda}_1, \underline{\lambda}\mathbf{I}_{p-q})$, where $\boldsymbol{\Lambda}_1$ is the $q\times q$ diagonal matrix with the q largest roots of $\boldsymbol{\Lambda}$ as diagonal elements,

$$\begin{aligned}\boldsymbol{\Sigma}_{p,q}n^{1/2}\mathrm{vec}\,(\hat{\boldsymbol{\beta}}_{0q}^\top\mathbf{V}\hat{\boldsymbol{\beta}}_{0q}) &= \boldsymbol{\Sigma}_{p,q}n^{1/2}\mathrm{vec}\,(\mathbf{E}_{21}\boldsymbol{\Lambda}_1\mathbf{E}_{21}^\top + \underline{\lambda}\mathbf{E}_{22}\mathbf{E}_{22}^\top)\\ &= n^{1/2}\underline{\lambda}\boldsymbol{\Sigma}_{p,q}\mathrm{vec}\,(\mathbf{E}_{22}\mathbf{E}_{22}^\top) + o_{\mathrm{P}}(1)\\ &= n^{1/2}\underline{\lambda}\boldsymbol{\Sigma}_{p,q}\mathrm{vec}\,(\mathbf{E}_{22}\mathbf{E}_{22}^\top - \mathbf{I}_{p-q}) + o_{\mathrm{P}}(1) = o_{\mathrm{P}}(1)\end{aligned} \tag{24}$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};g_1}$. Combining (23) and (24), we obtain that

$$\boldsymbol{\Sigma}_{p,q}\mathbf{T}_K^{(n)}(\hat{\mathbf{V}}) = \boldsymbol{\Sigma}_{p,q}(\hat{\boldsymbol{\beta}}_{0q}^\top\otimes\hat{\boldsymbol{\beta}}_{0q}^\top)n^{1/2}\mathrm{vec}(\mathbf{S}_K^{(n)}(\mathbf{V})) + o_{\mathrm{P}}(1) \tag{25}$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};g_1}$. We also have that

$$\hat{\boldsymbol{\beta}}_{0q}^\top = \hat{\boldsymbol{\beta}}_{0q}^\top\boldsymbol{\beta}\boldsymbol{\beta}^\top = \mathbf{E}_{22}\boldsymbol{\beta}_{0q}^\top + o_{\mathrm{P}}(1)$$

as $n\to\infty$ under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};g_1}$ so that since $n^{1/2}(\mathbf{E}_{22}\mathbf{E}_{22}^\top - \mathbf{I}_{p-q}) = o_{\mathrm{P}}(1)$ as $n\to\infty$, (25) entails that

$$\begin{aligned}S_{q,K}^{(n)} &= \frac{p(p+2)}{2\mathcal{J}_p(K)}\Big(n^{1/2}\mathrm{vec}(\mathbf{E}_{22}\boldsymbol{\beta}_{0q}^\top\mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q}\mathbf{E}_{22}^\top)^\top\boldsymbol{\Sigma}_{p,q}n^{1/2}\mathrm{vec}(\mathbf{E}_{22}\boldsymbol{\beta}_{0q}^\top\mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q}\mathbf{E}_{22}^\top)\Big) + o_{\mathrm{P}}(1)\\ &= \frac{p(p+2)}{2\mathcal{J}_p(K)}n\Big(\mathrm{tr}((\mathbf{E}_{22}\boldsymbol{\beta}_{0q}^\top\mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q}\mathbf{E}_{22}^\top)^2) - \frac{1}{p-q}\mathrm{tr}^2(\mathbf{E}_{22}\boldsymbol{\beta}_{0q}^\top\mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q}\mathbf{E}_{22}^\top)\Big) + o_{\mathrm{P}}(1)\\ &= \frac{p(p+2)}{2\mathcal{J}_p(K)}n\Big(\mathrm{tr}((\boldsymbol{\beta}_{0q}^\top\mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q})^2) - \frac{1}{p-q}\mathrm{tr}^2(\boldsymbol{\beta}_{0q}^\top\mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q})\Big) + o_{\mathrm{P}}(1).\end{aligned}$$

Now it follows from [6] that, still as $n \to \infty$ and under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};g_1}$, $n^{1/2}\mathrm{vec}(\boldsymbol{\beta}_{0q}^\top \mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q} - \mathbf{I}_{p-q})$ converges weakly to a Gaussian vector with mean zero and covariance matrix

$$\frac{\mathcal{J}_p(K)}{p(p+2)}(\mathbf{I}_{(p-q)^2} + \mathbf{K}_{p-q} - \frac{2}{p}\mathbf{J}_{p-q}).$$

Since

$$\frac{p(p+2)}{2\mathcal{J}_p(K)}\boldsymbol{\Sigma}_{p,q}\frac{\mathcal{J}_p(K)}{p(p+2)}(\mathbf{I}_{(p-q)^2} + \mathbf{K}_{p-q} - \frac{2}{p}\mathbf{J}_{p-q}) = \frac{1}{2}(\mathbf{I}_{(p-q)^2} + \mathbf{K}_{p-q} - \frac{2}{p-q}\mathbf{J}_{p-q})$$

is idempotent and such that

$$\mathrm{tr}(\frac{1}{2}(\mathbf{I}_{(p-q)^2} + \mathbf{K}_{p-q} - \frac{2}{p-q}\mathbf{J}_{p-q})) = d(p,q),$$

it follows from Theorem 9.2.1 in [22] that $S^{(n)}_{q,K}$ converges weakly to a chi-square random variable with $d(p,q)$ degrees of freedom. Point (i) follows. We now move to Point (ii). First note that from Proposition 1 and (19), we have that

$$\begin{aligned}
\boldsymbol{\tau}_n^\top \boldsymbol{\Delta}^{(n)}_{g_1} &= \frac{1}{2}(\mathrm{d}\overset{\circ}{\mathrm{vec}}\,(\mathrm{diag}(\boldsymbol{\ell}^{(n)})))^\top \mathbf{M}_p n^{1/2}\mathrm{dvec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top \mathbf{S}^{(n)}_{g_1}(\mathbf{V})\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) \\
&= \boldsymbol{\ell}^{(n)}\frac{1}{2}n^{1/2}\mathrm{dvec}\,(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top(\mathbf{S}^{(n)}_{g_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{g_1}(\mathbf{V})])\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_{\mathrm{P}}(1) \\
&= \boldsymbol{\ell}^{(n)}\mathbf{H}_p\frac{1}{2}n^{1/2}\mathrm{vec}(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top(\mathbf{S}^{(n)}_{g_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{g_1}(\mathbf{V})])\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_{\mathrm{P}}(1) \\
&= \mathrm{vec}^\top(\mathrm{diag}(\boldsymbol{\ell}^{(n)}))\frac{1}{2}n^{1/2}\mathrm{vec}(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\beta}^\top(\mathbf{S}^{(n)}_{g_1}(\mathbf{V}) - \mathrm{E}[\mathbf{S}^{(n)}_{g_1}(\mathbf{V})])\boldsymbol{\beta}\boldsymbol{\Lambda}^{-1/2}) + o_{\mathrm{P}}(1)
\end{aligned}$$

as $n \to \infty$ and under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda};g_1}$. Applying Le Cam third Lemma and denoting by $(\mathbf{0}_{(p-q)\times q}, \mathbf{I}_{p-q})$ the $(p-q)\times p$ matrix obtained by glueing $\mathbf{0}_{(p-q)\times q}$ and $\mathbf{I}_{p-q}$ together, we obtain that $n^{1/2}\mathrm{vec}(\boldsymbol{\beta}_{0q}^\top \mathbf{S}_K^{(n)}(\mathbf{V})\boldsymbol{\beta}_{0q} - \mathbf{I}_{p-q})$ converges weakly to a Gaussian vector with mean

$$\begin{aligned}
\boldsymbol{\mu} &= \frac{\mathcal{J}_p(K,g_1)}{2p(p+2)}(\boldsymbol{\beta}_{0q}^\top \otimes \boldsymbol{\beta}_{0q}^\top)(\mathbf{I}_{p^2} + \mathbf{K}_p - \frac{2}{p}\mathbf{J}_p)(\boldsymbol{\beta}\otimes\boldsymbol{\beta})\mathrm{vec}(\mathrm{diag}(\boldsymbol{\Lambda}^{-1}\boldsymbol{\ell}^{(n)})) \\
&= \frac{\mathcal{J}_p(K,g_1)}{p(p+2)}((\mathbf{0}_{(p-q)\times q}, \mathbf{I}_{p-q})^{\otimes 2} - \frac{1}{p}\mathrm{vec}(\mathbf{I}_{p-q})\mathrm{vec}^\top(\mathbf{I}_p))\mathrm{vec}(\mathrm{diag}(\boldsymbol{\Lambda}^{-1}\boldsymbol{\ell}^{(n)})) \\
&= \frac{\mathcal{J}_p(K,g_1)}{p(p+2)}(\underline{\lambda}^{-1}\mathrm{vec}(\mathrm{diag}(\ell^{(n)}_{q+1},\ldots,\ell^{(n)}_p)) - \frac{\mathrm{tr}(\mathrm{diag}(\boldsymbol{\Lambda}^{-1}\boldsymbol{\ell}^{(n)}))}{p}\mathrm{vec}(\mathbf{I}_{p-q}))
\end{aligned}$$

and covariance matrix

$$\frac{\mathcal{J}_p(K)}{p(p+2)}(\mathbf{I}_{(p-q)^2} + \mathbf{K}_{p-q} - \frac{2}{p}\mathbf{J}_{p-q})$$

under $\mathrm{P}^{(n)}_{\sigma,\boldsymbol{\beta},\boldsymbol{\Lambda}+n^{-1/2}\mathrm{diag}(\boldsymbol{\ell}^{(n)});g_1}$, $\mathrm{d}\overset{\circ}{\mathrm{vec}}\,(\boldsymbol{\Lambda}) \in \mathcal{H}_{0q}$, $g_1 \in \mathcal{F}_1$. The result follows. □

References

[1] T. W. Anderson, Asymptotic theory for principal component analysis, Ann. Mathem. Statist. 34 (1963) 122–148.

[2] E. A. Cator, H. P. Lopuhaä, Asymptotic expansion of the minimum covariance determinant estimators, J. Multivariate Anal. 101 (2010) 2372–2388.

[3] C. Croux, G. Haesbroeck, Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies, Biometrika 87 (2000) 603–618.

[4] A. Dürre, D. E. Tyler, D. Vogel, On the eigenvalues of the spatial sign covariance matrix in more than two dimensions, Statist. Probab. Lett. 111 (2016) 80–85.

[5] G. Frahm, Asymptotic distributions of robust shape matrices and scales, J. Multivariate Anal. 100 (2009) 1329–1337.
[6] M. Hallin, D. Paindaveine, Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity, Ann. Statist. 34 (2006) 2707–2756.
[7] M. Hallin, D. Paindaveine, T. Verdebout, Optimal rank-based testing for principal components, Ann. Statist. 38 (2010) 3245–3299.
[8] M. Hallin, D. Paindaveine, T. Verdebout, Efficient R-estimation of principal and common principal components, J. Amer. Statist. Assoc. 109 (2014) 1071–1083.
[9] T. P. Hettmansperger, R. H. Randles, A practical affine equivariant multivariate median, Biometrika 89 (2002) 851–860.
[10] P. Ilmonen, D. Paindaveine, Semiparametrically efficient inference based on signed ranks in symmetric independent component models, Ann. Statist. 39 (2011) 2448–2476.
[11] D. N. Lawley, Tests of significance for the latent roots of covariance and correlation matrices, Biometrika 43 (1956) 128–136.
[12] O. Ledoit, M. Wolf, Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, Ann. Statist. 30 (2002) 1081–1102.
[13] Z. Li, J. Yao, Testing the sphericity of a covariance matrix when the dimension is much larger than the sample size, Electron. J. Statist. 10 (2016) 2973–3010.
[14] H. P. Lopuhaa, Asymptotics of reweighted estimators of multivariate location and scatter, Ann. Statist. (1999) 1638–1665.
[15] W. Luo, B. Li, Combining eigenvalues and variation of eigenvectors for order determination, Biometrika 103 (2016) 875–887.
[16] B. Nadler, Nonparametric detection of signals by information theoretic criteria: performance analysis and an improved estimator, IEEE Trans. Signal Process. 58 (2010) 2746–2756.
[17] K. Nordhausen, H. Oja, D. E. Tyler, Asymptotic and bootstrap tests for subspace dimension, J. Multivariate Anal. 188 (2022) 104830.
[18] J. Nyblom, T. Mäkeläinen, Comparisons of tests for the presence of random walk coefficients in a simple linear model, J. Amer. Statist. Assoc. 78 (1983) 856–864.
[19] D. Paindaveine, A Chernoff–Savage result for shape: on the non-admissibility of pseudo-Gaussian methods, J. Multivariate Anal. 97 (2006) 2206–2220.
[20] D. Paindaveine, A canonical definition of shape, Statist. Probab. Lett. 78 (2008) 2240–2247.
[21] D. Paindaveine, On multivariate runs tests for randomness, J. Amer. Statist. Assoc. 104 (2009) 1525–1538.
[22] C. R. Rao, S. K. Mitra, Generalized Inverses of Matrices and its Applications, Wiley, New York, 1971.
[23] M. Salibián-Barrera, S. Van Aelst, G. Willems, Principal components analysis based on multivariate mm estimators with fast and robust bootstrap, J. Amer. Statist. Assoc. 101 (2006) 1198–1211.
[24] J. R. Schott, A high-dimensional test for the equality of the smallest eigenvalues of a covariance matrix, J. Multivariate Anal. 97 (2006) 827–843.
[25] S. Taskinen, C. Croux, A. Kankainen, E. Ollila, H. Oja, Influence functions and efficiencies of the canonical correlation and vector estimates based on scatter and shape matrices, J. Multivariate Anal. 97 (2006) 359–384.
[26] D. Tyler, A distribution-free m-estimator of multivariate scatter, Ann. Statist. 15 (1987) 234–251.
[27] D. E. Tyler, The asymptotic distribution of principal component roots under local alternatives to multiple roots, Ann. Statist. (1983) 1232–1242.
[28] J. Virta, Testing for subsphericity when n and p are of different asymptotic order, Statist. Probab. Lett. 179 (2021) 109209.
[29] C. M. Waternaux, Principal components in the nonnormal case: the test of equality of q roots, J. Multivariate Anal. 14 (1984) 323–335.