SFB 823

Optimal rank-based tests for common principal components

Marc Hallin, Davy Paindaveine, Thomas Verdebout

Discussion Paper

Nr. 53/2011

Optimal Rank-based Tests for Common Principal Components

MARC HALLIN[1,2,*] DAVY PAINDAVEINE[1,3,†] and THOMAS VERDEBOUT[4,‡]

[1] *ECARES*
Université Libre de Bruxelles
Av. F.D. Roosevelt, 50, CP114/04,
B-1050 Bruxelles, Belgique.
E-mail: mhallin@ulb.ac.be*;* dpaindav@ulb.ac.be*.*

[2] *ORFE*
Princeton University
Sherrerd Hall
Princeton, NJ 08544, USA.

[3] *LSTA*
Université Pierre et Marie Curie - Paris 6
Place Jussieu, 4
F-75252 Paris Cedex 05, France.

[4] *EQUIPPE*
Université Lille Nord de France
Domaine Universitaire du Pont de Bois, BP 60149
F-59653 Villeneuve-d'Ascq Cedex, France.
E-mail: thomas.verdebout@univ-lille3.fr*.*

This paper provides optimal testing procedures for the m-sample null hypothesis of Common Principal Components (CPC) under possibly non Gaussian and heterogenous elliptical densities. We first establish, under very mild assumptions that do not require finite moments of order four, the local asymptotic normality (LAN) of the model. Based on that result, we show that the pseudo-Gaussian test proposed in Hallin et al. (2010a) is locally and asymptotically optimal under Gaussian densities. We also show how to compute its local powers and asymptotic relative efficiencies (AREs). A numerical evaluation of those AREs, however, reveals that, while remaining valid, this test is poorly efficient away from the Gaussian. Moreover, it still requires finite moments of order four. We therefore propose rank-based procedures that remain valid under any possibly heterogenous m-tuple of elliptical densities, irrespective of any moment assumptions—in elliptical families, indeed, principal components naturally can be based on the scatter matrices characterizing the density contours, hence do not require finite variances. Those rank-based tests are not only validity-robust in the sense that they survive arbitrary elliptical population

*Académie Royale de Belgique, CentER, Tilburg University, and ECORE. Research supported by the Sonderforschungsbereich "Statistical modelling of nonlinear dynamic processes" (SFB 823) of the Deutsche Forschungsgemeinschaft. Marc Hallin and Davy Paindaveine are also members of ECORE, the association between CORE and ECARES.

†Research supported by an Action de Recherche Concertée, Communauté française de Belgique.

‡Research supported by a BQR (Bonus Qualité Recherche) fund, Université Lille III, France.

densities: we show that they also are efficiency-robust, in the sense that their local powers do not deteriorate under non-Gaussian alternatives. In the homogeneous case, the normal-score version of our tests uniformly dominates, in the Pitman sense, the optimal pseudo-Gaussian test. Theoretical results are obtained via a nonstandard application of Le Cam's methodology in the context of *curved* LAN experiments. The finite-sample properties of the proposed tests are investigated through simulations.
Keywords: Common Principal Components, rank-based methods, local asymptotic normality, robustness.

1. Introduction.

Principal components—arguably, the oldest and most popular tool of multivariate analysis—were originally introduced by Pearson (1901), then rediscovered by Hotelling (1933), in a one-sample context. Multisample principal component problems only came much later, when Flury (1984) introduced the Common Principal Components (CPC) model. CPC models since then have been used in a number of applications, mainly in a biometric context (see e.g. Airoldi and Hoffmann (1984), Flury and Riedl (1988)). Under such a model, $m \geq 2$ populations of dimension k, with covariance matrices $\mathbf{\Sigma}_i^{\rm Cov}$, $i = 1, \dots, m$, are assumed to share, with possibly different eigenvalues, the same principal components: namely, these covariance matrices factorize into $\mathbf{\Sigma}_i^{\rm Cov} = \boldsymbol{\beta}\mathbf{\Lambda}_i^{\rm Cov}\boldsymbol{\beta}'$ for some m-tuple of positive diagonal matrices $\mathbf{\Lambda}_i^{\rm Cov}$, $i = 1, \dots, m$, and some orthogonal matrix $\boldsymbol{\beta}$—the matrix of *common eigenvectors*, which does not depend on i and characterizes the *common principal components*. CPC models later on have been generalized (Flury 1988) into *partial* CPC models, in which only a subset of $q < k$ principal components are common to the m populations. More recently, a broader class of models, which includes CPC and partial CPC, but also possible common eigenspaces, has been considered by Boik (2002).

Before considering a statistical analysis based on such model, however, it is natural to check whether the CPC assumption is compatible with the data under study. Flury (1984) therefore developed a Gaussian likelihood ratio test $\phi_{\mathcal{N}}^{(n)}$ for the null hypothesis $\mathcal{H}_0$ of common principal components. This test is based on the asymptotically chi-square null distribution of $-2\log\Lambda$ where, denoting by $\mathbf{S}_i^{(n)}$, $i = 1, \dots, m$ the empirical covariance matrices computed from m mutually independent samples of k-dimensional independent observations and by $\hat{\boldsymbol{\beta}}$ the (constrained) maximum likelihood estimator of $\boldsymbol{\beta}$,

$$\Lambda := \prod_{i=1}^{m} \left(\frac{\det(\hat{\boldsymbol{\beta}}'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}})}{\det(\mathrm{diag}(\hat{\boldsymbol{\beta}}'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}}))} \right)^{n_i/2} \tag{1.1}$$

(we write $\mathrm{diag}(\mathbf{A})$ for the diagonal matrix having the same diagonal elements as a squared matrix $\mathbf{A}$). Under $\mathcal{H}_0$, $\hat{\boldsymbol{\beta}}'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}}$ should be nearly diagonal, hence $\det(\hat{\boldsymbol{\beta}}'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}})$ and $\det(\mathrm{diag}(\hat{\boldsymbol{\beta}}'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}}))$ approximately equal, in which case Λ is close to one; under the alternative, Λ is closer to zero (hence, $-2\log\Lambda$ is large), leading to the rejection of the CPC hypothesis. The asymptotically chi-square distribution of $-2\log\Lambda$ follows from the classical asymptotic result of Wilks (1938), a result which, however, is valid under Gaussian assumptions only.

It is well known that Gaussian likelihood ratio tests (LRT) for hypotheses involving covariance matrices, are quite sensitive to violations of Gaussian assumptions and to the presence of outliers (on this latter point, see e.g. Boente and Orellana (2001)). The test based on (1.1) is no exception to that rule: $-2\log\Lambda$ under non-Gaussian densities is no longer asymptotically chi-square, but converges in distribution to a weighted sum of independent chi-square variables.

This type of asymptotic behavior is not uncommon, and there exists an extensive literature on how to preserve the chi-square asymptotics of LRT statistics: Muirhead and Waternaux (1980), Browne (1984), Satorra and Bentler (1988), Bentler and Dungeon (1996), provide adjusted LRTs for various problems. Shapiro and Browne (1987) give a necessary and sufficient condition under which such adjusted test statistics remain asymptotically chi-square.

Using this result by Shapiro and Browne, Boik (2002) constructs a test—$\phi^{(n)}_{\text{Boik}}$, say—for the null hypothesis of CPC based on a statistic which remains asymptotically chi-square under families of elliptical distributions with finite fourth-order moments and common kurtosis parameter. That is, denoting by $\kappa_k(g_i)$ the kurtosis in the ith elliptical distribution, $i = 1, \ldots, m$ (see Section 6 for a precise definition), the validity of this test requires the somewhat stringent assumption of *homokurticity* $\kappa_k(g_1) = \ldots = \kappa_k(g_m)$.

In a series of papers, Boente *et al.* (2001, 2002, and 2009) generalize Boik's test by substituting robust scatter matrices for the regular covariance matrices, which reduces the impact of possible outliers. In terms of validity robustness, however, the resulting tests do not improve much on Boik's, as they merely replace the assumption of homokurticity with an assumption of the form $\varsigma(g_1) = \ldots = \varsigma(g_m)$, where the parameter $\varsigma(g)$ depends on the chosen concept of scatter—a homogeneity assumption that is hardly more natural or realistic than Boik's homokurticity assumption.

Hallin et al. (2010a) amend this situation by introducing a pseudo-Gaussian test $\phi^{(n)}_{\text{HPV}}$ the validity of which, unlike that of $\phi^{(n)}_{\mathcal{N}}$ and $\phi^{(n)}_{\text{Boik}}$, resists heterokurtic violations of Gaussian assumptions. Under Gaussian densities, this test is asymptotically equivalent to Flury's LRT. However, $\phi^{(n)}_{\text{HPV}}$ still requires finite fourth-order moments, and follows as a robustified version of Flury's LRT $\phi^{(n)}_{\mathcal{N}}$, the exact optimality properties of which have not been investigated so far. These issues (certainly, the fourth-order moment assumption) are not dramatic, and most statisticians would feel comfortable using such tests. Unfortunately, it appears from the power analysis in Sections 8.1 and 8.2 below that $\phi^{(n)}_{\text{HPV}}$ exhibits disturbingly low power against non-Gaussian alternatives. In the two-population case with bivariate t_5 densities (homokurtic case with finite fourth-order moments), the asymptotic relative efficiency of $\phi^{(n)}_{\text{HPV}}$ with respect to the locally optimal procedure is as low as .4286, whereas the normal-score (van der Waerden) test we are proposing here achieves asymptotic relative efficiency .9446—more than twice as much. Under $t_{4.2}$ densities, the figures are .1202 and .9303, respectively! It seems, thus, that the robustification of $\phi^{(n)}_{\mathcal{N}}$ into $\phi^{(n)}_{\text{HPV}}$ is obtained at the expense of efficiency—which, most statisticians will agree, is quite a heavy price.

The objective of this paper is to remedy those validity and efficiency problems by proposing rank-based tests that outperform the available parametric ones on both counts. These rank tests enjoy enhanced validity properties; in particular, they allow for heterokurticity and do not require any moment assumptions at all. In the same time, they exhibit increased efficiency-robustness: see Tables 1 and 2. The asymptotic relative efficiencies of the van der Waerden version of our tests with respect to $\phi^{(n)}_{\mathrm{HPV}}$, for instance, are uniformly larger than one—a theoretical finding that is supported by the simulation results of Section 8.2.

Reaching that double objective requires overcoming several technical difficulties. First, the traditional covariance-based concept of common principal components has to be extended in order to cope with the possible absence of second-order moments. In elliptical families, the *scatter* and *shape* matrices that characterize the elliptical equidensity contours quite naturally qualify as moment-free substitutes for covariance matrices (with which they coincide, up to a scalar factor, in case second-order moments do exist). Second, based on a parametrization involving those scatter and shape matrices, we establish the local asymptotic normality (LAN) of the model in the vicinity of the CPC hypothesis. This provides us with a clear definition of optimality, and a way of computing local powers and asymptotic relative efficiencies—with, however, the additional difficulty that the *limiting local experiments* associated with the scatter- or shape-matrix parametrization are not full-rank Gaussian shift experiments but *curved* ones. Such curved LAN structures were previously studied in Hallin *et al.* (2010b). As a by-product, we establish the exact optimality properties of Flury's $\phi^{(n)}_{\mathcal{N}}$ and Hallin *et al.*'s $\phi^{(n)}_{\mathrm{HPV}}$. Finally, following the method used in Hallin *et al.* (2010b) for the one-sample case, we construct rank-based versions of the optimal test statistics associated with various radial densities, and compute the corresponding local powers and asymptotic relative efficiencies.

2. Outline of the paper.

The outline of the paper is as follows. Section 3 states the assumptions to be used in the sequel. The parametrization we propose for testing the CPC hypothesis in an elliptical context is described in Section 4. Section 5 provides the uniform local and asymptotic (ULAN) property, on which the construction of optimal tests will be based. Section 6 derives Gaussian and pseudo-Gaussian tests for the CPC hypothesis, and Section 7 introduces optimal rank-based tests for the same. In Section 8, the performances of the proposed tests are investigated both in terms of asymptotic relative efficiencies (Section 8.1) and simulations (Section 8.2). Finally, the Appendix collects technical proofs.

3. Main assumptions.

For the sake of convenience, we are collecting here the main assumptions to be used in the sequel.

3.1. Elliptical symmetry

Denote by $(\mathbf{X}_{i1}, \ldots, \mathbf{X}_{in_i})$, $i = 1, \ldots, m$ a collection of m mutually independent samples of i.i.d. k-dimensional random vectors with *elliptically symmetric densities*. More precisely, the n_i observations $\mathbf{X}_{ij}$, $j = 1, \ldots, n_i$ in sample i are independent, with density

$$\underline{f_i}(\mathbf{x}) := c_{k,f_i} \, (\det(\mathbf{\Sigma}_i))^{-1/2} \, f_i\Big(((\mathbf{x}-\boldsymbol{\theta}_i)'\mathbf{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\theta}_i))^{1/2}\Big), \tag{3.1}$$

for some k-dimensional *location* parameter $\boldsymbol{\theta}_i$, some symmetric and positive definite *scatter* matrix $\mathbf{\Sigma}_i$ and some *radial density* function $f_i : \mathbb{R}_0^+ \mapsto \mathbb{R}^+$; c_{k,f_i} is a normalization constant such that $\int_{\mathbb{R}^k} \underline{f_i}(\mathbf{x}) \mathrm{d}\mathbf{x} = 1$. Note that, despite the terminology, the radial density f_i is not a probability density (over the positive real line), since it does not integrate to one; but $\tilde{f}_i := \mu_{k-1;f_i}^{-1} r^{k-1} f_i$ (for the sake of simplicity, we write $\tilde{f}_i$ instead of $\tilde{f}_{ik}$), where $\mu_{\ell;f} := \int_0^\infty r^\ell f(r)\,\mathrm{d}r$, is. Define

$$\mathcal{F} := \Big\{ f : f(r) > 0 \text{ a.e. and } \mu_{k-1;f} < \infty \Big\} \text{ and } \mathcal{F}_1 := \Big\{ f \in \mathcal{F} : \frac{1}{\mu_{k-1;f}} \int_0^1 r^{k-1} f(r)\,\mathrm{d}r = \frac{1}{2} \Big\};$$

$\mathcal{F}_1$ is a class of *standardized* radial densities, in the sense that, for any radial density $f \in \mathcal{F}_1$, the probability density $\tilde{f}(r) := \mu_{k-1;f}^{-1} r^{k-1} f(r)$ is a properly standardized probability density. By "standardized", here, we mean that the corresponding median is one; the median, for a nonvanishing density over $\mathbb{R}_0^+$, indeed, is a scale parameter which does not require any moment conditions.

Summarizing this, we throughout assume that the following assumption holds true.

Assumption (A). The observations $\mathbf{X}_{ij}$, $j = 1, \ldots, n_i$, $i = 1, \ldots, m$ are mutually independent, with probability densities $\underline{f_i}$ given in (3.1), for some m-tuple of (possibly distinct) radial densities $\mathrm{f} := (f_1, \ldots, f_m) \in (\mathcal{F}_a)^m$, where $\mathcal{F}_a \subset \mathcal{F}_1$ is defined below.

Under Assumption (A), the elliptical distances $d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i) := \|\mathbf{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)\|$, $j = 1, \ldots, n_i$, $i = 1, \ldots, m$, have probability density $\tilde{f}_i$, with median one, which identifies the *scatter* matrices $\mathbf{\Sigma}_i$, $i = 1, \ldots, m$ also in the absence of any moments. Under finite second-order moments, however, $\mathbf{\Sigma}_i$ is proportional to the covariance matrix $\mathbf{\Sigma}_i^{\mathrm{Cov}}$ of $\mathbf{X}_{ij}$. Assumption (A) allows for heterogeneity of the m elliptical densities, that is, we may have $f_i \neq f_{i'}$ for $i \neq i'$. Classical examples are the k-variate multinormal distributions, with standardized radial densities $f_i(r) = \phi(r) := \exp(-a_k r^2/2)$, the k-variate Student distributions, with standardized radial densities (for $\nu \in \mathbb{R}_0^+$ degrees of freedom) $f_i(r) = f_\nu^t(r) := (1 + a_{k,\nu} r^2/\nu)^{-(k+\nu)/2}$, and the k-variate power-exponential distributions, with standardized radial densities of the form $f_i(r) = f_\eta^e(r) := \exp(-b_{k,\eta} r^{2\eta})$, $\eta \in \mathbb{R}_0^+$; the positive constants a_k, $a_{k,\nu}$, and $b_{k,\eta}$ are such that $f_i \in \mathcal{F}_1$.

The equidensity contours associated with (3.1) are hyper-ellipsoids centered at $\boldsymbol{\theta}_i$, the shape and orientation of which are determined by the scatter matrix $\mathbf{\Sigma}_i$. The *multivariate signs* $\mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i) := \mathbf{\Sigma}_i^{-1/2}(\mathbf{X}_{ij} - \boldsymbol{\theta}_i)/d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i)$ and *standardized radial distances* $d_{ij}(\boldsymbol{\theta}_i, \mathbf{\Sigma}_i)$ just defined are $\mathbf{X}_{ij}$'s (within-group) elliptical coordinates associated with those ellipsoids. The observations then decompose into $\mathbf{X}_{ij} = \boldsymbol{\theta}_i + d_{ij}\mathbf{\Sigma}_i^{1/2}\mathbf{U}_{ij}$, where the $\mathbf{U}_{ij}$'s, $j = 1, \ldots, n_i$, $i = 1, \ldots, m$ are i.i.d. uniform over the unit sphere in $\mathbb{R}^k$, and the d_{ij}'s are i.i.d., independent of the $\mathbf{U}_{ij}$'s, with standardized probability density $\tilde{f}_i$ over $\mathbb{R}^+$ and distribution function $\tilde{F}_i$. In the sequel, the notation $\tilde{g}_i$ and $\tilde{G}_i$ will be used for the same functions computed from a standardized radial density $g_i (\in \mathcal{F}_1)$.

The derivation of locally and asymptotically optimal tests at a given m-tuple $\mathrm{f} = (f_1, \ldots, f_m)$ of radial densities will be based on the *uniform local and asymptotic normality* (ULAN) of the corresponding model. This ULAN property only holds under some mild regularity conditions on the f_i's. More precisely, ULAN (see Proposition 5.1 below) requires the f_i's to belong to the collection $\mathcal{F}_a$ of those radial densities $f \in \mathcal{F}_1$ which are absolutely continuous, with almost everywhere derivative $\dot{f}$ such that, letting $\varphi_f := -\dot{f}/f$ and denoting by $\tilde{F}$ the distribution function associated with $\tilde{f}$, the integrals

$$\mathcal{I}_k(f) := \int_0^1 \varphi_f^2(\tilde{F}^{-1}(u))\,\mathrm{d}u \quad \text{and} \quad \mathcal{J}_k(f) := \int_0^1 \varphi_f^2(\tilde{F}^{-1}(u))(\tilde{F}^{-1}(u))^2\,\mathrm{d}u$$

are finite. The quantities $\mathcal{I}_k(f_i)$ and $\mathcal{J}_k(f_i)$ play the roles of *radial Fisher information* for location and shape/scale, respectively, in sample i, $i = 1, \ldots, m$ (see Hallin and Paindaveine 2006).

3.2. Asymptotic behavior of sample sizes.

Actually, we throughout consider triangular arrays of observations, of the form

$$(\mathbf{X}_{11}^{(n)}, \ldots, \mathbf{X}_{1n_1^{(n)}}^{(n)}, \mathbf{X}_{21}^{(n)}, \ldots, \mathbf{X}_{2n_2^{(n)}}^{(n)}, \ldots, \mathbf{X}_{m1}^{(n)}, \ldots, \mathbf{X}_{mn_m^{(n)}}^{(n)}),$$

indexed by the total sample size $n := \sum_{i=1}^m n_i^{(n)}$, where the sequences $n_i^{(n)}$ satisfy the following assumption.

Assumption (B). For all $i = 1, \ldots, m$, $n_i = n_i^{(n)} \to \infty$ as $n \to \infty$.

This assumption is weaker than the assumption usually made in (univariate or multivariate) multisample problems, where it is required that $n_i^{(n)}/n$ be bounded away from 0 and 1 for all i as $n \to \infty$. Letting $r_i^{(n)} := n_i^{(n)}/n$, it is easy to check that Assumption (B) is actually equivalent to the Noether conditions

$$\max\left(\frac{1 - r_i^{(n)}}{r_i^{(n)}}, \frac{r_i^{(n)}}{1 - r_i^{(n)}}\right) = o(n) \quad \text{as } n \to \infty, \text{ for all } i.$$

However, in the derivation of asymptotic distributions under local alternatives, we will need—mainly, for notational comfort—the following classical reinforcement:

Assumption (B′). For all $i = 1, \ldots, m$, $r_i^{(n)} \to r_i \in (0, 1)$, as $n \to \infty$.

For notational simplicity, we henceforth omit superfluous ${}^{(n)}$ superscripts.

3.3. Score functions.

The signed-rank tests considered in Section 6 are based on m-tuples $\mathrm{K} = (K_1, \ldots, K_m)$ of *score functions*, which we assume to satisfy the following regularity conditions.

Assumption (C). For any $i = 1, \ldots, m$, the mapping (from $(0,1)$ to $\mathbb{R}$) $u \mapsto K_i(u)$
(C1) is continuous and square-integrable;
(C2) can be expressed as the difference of two monotone increasing functions, and
(C3) satisfies $\int_0^1 K_i(u)\,\mathrm{d}u = k$.

Assumption (C3) is a normalization constraint that is automatically satisfied by the score functions $K_i(u) = K_{f_i}(u) := \varphi_{f_i}(\tilde{F}_i^{-1}(u))\tilde{F}_i^{-1}(u)$ leading to local and asymptotic optimality at m-tuples of radial densities $\mathrm{f} = (f_1, \ldots, f_m)$ for which ULAN holds; see Section 5.

For score functions K, K_1, K_2 satisfying Assumption (C), let (throughout, U stands for a random variable uniformly distributed over $(0,1)$), $\mathcal{J}_k(K_1, K_2) := \mathrm{E}[K_1(U)K_2(U)]$. For simplicity, we write $\mathcal{J}_k(K)$ for $\mathcal{J}_k(K, K)$, $\mathcal{J}_k(K, f)$ for $\mathrm{E}[K(U)K_f(U)]$, etc.

The power score functions $K_a(u) := k(a+1)u^a$ $(a \geq 0)$ provide some traditional score functions satisfying Assumption (C), with $\mathcal{J}_k(K_a) = k^2(a+1)^2/(2a+1)$: the Laplace, Wilcoxon and Spearman scores are obtained for $a = 0$, 1, and 2, respectively. As for score functions of the form K_{f_i}, an important particular case is that of van der Waerden or normal scores, obtained for $f_i = \phi$. Then, denoting by Ψ_k the chi-square distribution function with k degrees of freedom, $K_\phi(u) = \Psi_k^{-1}(u)$, and $\mathcal{J}_k(\phi) = k(k+2)$. Similarly, writing $G_{k,\nu}$ for the Fisher-Snedecor distribution function with k and ν degrees of freedom, Student densities $f_i = f_\nu^t$ yield

$$K_{f_\nu^t}(u) = \frac{k(k+\nu)G_{k,\nu}^{-1}(u)}{\nu + kG_{k,\nu}^{-1}(u)} \quad \text{and} \quad \mathcal{J}_k(f_\nu^t) = \frac{k(k+2)(k+\nu)}{k+\nu+2}.$$

4. Parametrization of m-sample elliptical models.

A natural notation for the joint distribution of the n-tuple $(\mathbf{X}'_{11}, \ldots, \mathbf{X}'_{mn_m})'$ under Assumption (A), parameter values $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m$, $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_m$, and the m-tuple f of radial densities, is $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_m;\boldsymbol{\Sigma}_1,\ldots,\boldsymbol{\Sigma}_m;\mathrm{f}}$. Such parametrization, however, is not well adapted to the present context, due to the fact that eigenvectors and eigenvalues are complicated functions of the scatter matrices. As in Hallin *et al.* (2010b), a parametrization based on eigenvectors and eigenvalues, which we now describe, will prove much more adequate.

4.1. Scatter, scale, and shape.

Since the eigenvectors $\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(m)}$ of $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_m$ are scale-free functions of the $\boldsymbol{\Sigma}_i$'s, it is appropriate to first decompose each $\boldsymbol{\Sigma}_i$ into a product $\boldsymbol{\Sigma}_i = \sigma_i^2\mathbf{V}_i$, where σ_i is a scalar *global scale* parameter and $\mathbf{V}_i$ a *shape* matrix (see Hallin and Paindaveine (2006) for details) for sample i. Paindaveine (2008) has shown the advantages of doing so by defining σ_i^2 as $(\det \mathbf{\Sigma}_i)^{1/k}$. This definition, which we adopt here, implies that the eigenvalues $\lambda_{ij}^{\mathbf{V}}$ of the shape matrices $\mathbf{V}_i$ are such that $\prod_{j=1}^k \lambda_{ij}^{\mathbf{V}} = 1$ for all $i = 1, \ldots, m$; clearly, $\mathbf{V}_i$ and $\mathbf{\Sigma}_i$ share the same eigenvectors.

4.2. Shape eigenvalues and eigenvectors.

Shape matrices in turn factorize into $\mathbf{V}_i = \boldsymbol{\beta}^{(i)\prime} \mathbf{\Lambda}_i^{\mathbf{V}} \boldsymbol{\beta}^{(i)}$, with $\mathbf{\Lambda}_i^{\mathbf{V}} := \mathrm{diag}(\lambda_{i1}^{\mathbf{V}}, \ldots, \lambda_{ik}^{\mathbf{V}})$ (throughout $\mathrm{diag}(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_m)$ stands for the block-diagonal matrix with diagonal blocks $\mathbf{B}_1$, $\mathbf{B}_2, \ldots, \mathbf{B}_m$). Even in case the $\lambda_{ij}^{\mathbf{V}}$'s are all distinct, this factorization, due to possible permutations of eigenvalues and the columns of $\boldsymbol{\beta}^{(i)}$, is not unique, and it is usually imposed, without any loss of generality, that the diagonal elements of $\mathbf{\Lambda}_i^{\mathbf{V}}$ are ranked in decreasing order of magnitude, which provides each eigenvalue $\lambda_{ij}^{\mathbf{V}}$ and the corresponding eigenvector $\boldsymbol{\beta}_j^{(i)}$ with a well-defined label j.

That way of labeling eigenvalues and eigenvectors is used in the statement of ULAN in Section 5 below. The same labeling however is no longer adequate when describing the null hypothesis $\mathcal{H}_0$ of CPC. The existence of a common $\boldsymbol{\beta}$ indeed induces a matching between the eigenvalues of the various populations, and it would be natural to label them so that $\lambda_{i1}^{\mathbf{V}}$, for $i = 1, \ldots, m$ be associated with $\boldsymbol{\beta}$'s first column $\boldsymbol{\beta}_1$, $\lambda_{i2}^{\mathbf{V}}$, $i = 1, \ldots, m$ with $\boldsymbol{\beta}_2$, etc. Under such labeling, $\mathcal{H}_0$ would take the simple form $\boldsymbol{\beta}^{(1)} = \ldots = \boldsymbol{\beta}^{(m)}$ instead of "there exist $(m-1)$ permutation matrices $\mathbf{M}_2^{\Pi}, \ldots, \mathbf{M}_m^{\Pi}$ such that $\boldsymbol{\beta}^{(1)} = \mathbf{M}_2^{\Pi} \boldsymbol{\beta}^{(2)} = \ldots = \mathbf{M}_m^{\Pi} \boldsymbol{\beta}^{(m)} =: \boldsymbol{\beta}$". This $\boldsymbol{\beta}$-induced labeling, however, only exists under $\mathcal{H}_0$, and, being $\boldsymbol{\beta}$-dependent, only holds over a neighborhood of $\boldsymbol{\beta}$; hence, it is local. We therefore adopt the "traditional" labeling in the statement of ULAN, and switch to the local $\boldsymbol{\beta}$-induced labels when optimal tests are to be derived (these tests, typically, will involve the ordering of eigenvalues induced by some adequate estimator $\hat{\boldsymbol{\beta}}$).

Establishing ULAN for a parametrization involving eigenvector matrices $\boldsymbol{\beta}^{(i)}$ and eigenvalues $\mathbf{\Lambda}_i^{\mathbf{V}}$ requires a differentiable correspondence between the $\mathbf{V}_i$'s and the corresponding $(\boldsymbol{\beta}^{(i)}, \mathbf{\Lambda}_i^{\mathbf{V}})$'s. Therefore, we need the following assumption.

Assumption (D). For all $i = 1, \ldots, m$, the scatter $\mathbf{\Sigma}_i$ (equivalently, the shape $\mathbf{V}_i$) has k distinct eigenvalues: $\lambda_{i1}^{\mathbf{\Sigma}} > \ldots > \lambda_{ik}^{\mathbf{\Sigma}}$.

While ULAN indeed requires Assumption (D), the tests we will propose, as we will show, remain (asymptotically) valid under the weaker

Assumption (D′). For any $1 \leq j \neq j' \leq k$, there exists $i \in \{1, \ldots, m\}$ such that $\lambda_{ij}^{\mathbf{\Sigma}} \neq \lambda_{ij'}^{\mathbf{\Sigma}}$.

Now, under the null and Assumption (D′), the matrix $\boldsymbol{\beta} := (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k)$ of common eigenvectors is identified up to an arbitrary permutation of its columns (we still forget about the irrelevant sign changes of the $\boldsymbol{\beta}_j$'s). However, it is easy to fix an ordering, hence to make the $\boldsymbol{\beta}_j$'s—hence also the corresponding $\lambda_{ij}^{\mathbf{\Sigma}}$'s—(individually) identifiable.

For instance, one can require that $\lambda^{\mathbf{\Sigma}}_{11} \geq \lambda^{\mathbf{\Sigma}}_{12} \geq \ldots \geq \lambda^{\mathbf{\Sigma}}_{1k} (> 0)$, and that, for any sequence of the form $\lambda^{\mathbf{\Sigma}}_{1j} = \lambda^{\mathbf{\Sigma}}_{1,j+1} = \ldots = \lambda^{\mathbf{\Sigma}}_{1,j+\ell}$, one has $\lambda^{\mathbf{\Sigma}}_{2j} \geq \lambda^{\mathbf{\Sigma}}_{2,j+1} \geq \ldots \geq \lambda^{\mathbf{\Sigma}}_{2,j+\ell}$. Recursively, if further ties occur among those $\lambda^{\mathbf{\Sigma}}_{2,j}$'s, the ranking can be based on the way the $\lambda^{\mathbf{\Sigma}}_{3,j}$'s are ordered, etc. Clearly, Assumption (D′) ensures that this correctly defines a unique ordering of the common principal directions and corresponding eigenvalues. Note that the largest eigenspace common to $\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_m$ (equivalently, to $\mathbf{V}_1, \ldots, \mathbf{V}_m$) then has dimension less than or equal to one.

4.3. Parameter space.

The parametrization we are adopting in the sequel is similar to that considered in the one-sample case by Hallin et al. (2010b); it is based on the $L := mk(k+2)$-dimensional vector (we denote by $\mathrm{dvec}\,(\mathbf{A}) =: (\mathbf{A}_{11}, (\overset{\circ}{\mathrm{dvec}}\,(\mathbf{A}))')'$ the vector obtained by stacking the diagonal elements of a squared matrix $\mathbf{A}$)

$$
\begin{aligned}
\boldsymbol{\vartheta} &:= (\boldsymbol{\vartheta}'_{I}, \boldsymbol{\vartheta}'_{II}, \boldsymbol{\vartheta}'_{III}, \boldsymbol{\vartheta}'_{IV})' \\
&:= (\boldsymbol{\theta}'_1, \ldots, \boldsymbol{\theta}'_m, \sigma^2_1, \ldots, \sigma^2_m, (\overset{\circ}{\mathrm{dvec}}\,\mathbf{\Lambda}^{\mathbf{V}}_1)', \ldots, (\overset{\circ}{\mathrm{dvec}}\,\mathbf{\Lambda}^{\mathbf{V}}_m)', (\mathrm{vec}\,\boldsymbol{\beta}^{(1)})', \ldots, (\mathrm{vec}\,\boldsymbol{\beta}^{(m)})')',
\end{aligned}
$$

where $\boldsymbol{\theta}_i$ and σ^2_i are the location and scale parameters, $\mathbf{\Lambda}^{\mathbf{V}}_i := \mathrm{diag}(\lambda^{\mathbf{V}}_{i1}, \ldots, \lambda^{\mathbf{V}}_{ik})$ and $\boldsymbol{\beta}^{(i)}$ the shape eigenvalue and eigenvector matrices, respectively, in population i, $i = 1, \ldots, m$; the reason why $\lambda^{\mathbf{V}}_{i1}$ is omitted in the parametrization is that, $\mathbf{V}_i$ being a shape matrix, $\lambda^{\mathbf{V}}_{i1} = 1/\prod_{j=2}^{m} \lambda^{\mathbf{V}}_{ij}$. The parameter space is thus $\mathbf{\Theta} := \mathbb{R}^{mk} \times (\mathbb{R}^+_0)^m \times (\mathcal{C}^{k-1})^m \times (\mathrm{vec}\,(\mathcal{SO}_k))^m$, where $\mathcal{C}^{k-1}$ is the open cone of $(\mathbb{R}^+_0)^{k-1}$ with strictly ordered (from largest to smallest) coordinates, and $\mathcal{SO}_k$ stands for the class of $k \times k$ real orthogonal matrices with determinant one. Note that Assumption (D) is explicitly incorporated in the definition of $\mathbf{\Theta}$.

We denote by $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f}$ or $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_I, \boldsymbol{\vartheta}_{II}, \boldsymbol{\vartheta}_{III}, \boldsymbol{\vartheta}_{IV}; \mathrm{f}}$ the joint distribution of the n observations under parameter value $\boldsymbol{\vartheta}$ and standardized radial densities $\mathrm{f} = (f_1, \ldots, f_m)$.

5. Uniform local asymptotic normality (ULAN).

As mentioned in Section 1, we plan to construct tests that are optimal at correctly specified densities, in the sense of Le Cam's asymptotic theory of statistical experiments. In this section, we state the ULAN result (with respect to $\boldsymbol{\vartheta} \in \mathbf{\Theta}$, for fixed radial densities $\mathrm{f} = (f_1, \ldots, f_m)$) on which optimality will be based. Denote by

$$
\begin{aligned}
\boldsymbol{\vartheta}^{(n)} := \; & (\boldsymbol{\vartheta}^{(n)\prime}_{I}, \boldsymbol{\vartheta}^{(n)\prime}_{II}, \boldsymbol{\vartheta}^{(n)\prime}_{III}, \boldsymbol{\vartheta}^{(n)\prime}_{IV})' := (\boldsymbol{\theta}^{(n)\prime}_1, \ldots, \boldsymbol{\theta}^{(n)\prime}_m, \\
& \sigma^{2(n)}_1, \ldots, \sigma^{2(n)}_m, (\overset{\circ}{\mathrm{dvec}}\,\mathbf{\Lambda}^{\mathbf{V}(n)}_1)', \ldots, (\overset{\circ}{\mathrm{dvec}}\,\mathbf{\Lambda}^{\mathbf{V}(n)}_m)', (\mathrm{vec}\,\boldsymbol{\beta}^{(1),(n)})', \ldots, (\mathrm{vec}\,\boldsymbol{\beta}^{(m),(n)})')'
\end{aligned}
$$

a local sequence such that $\boldsymbol{\vartheta}^{(n)} \in \mathbf{\Theta}$ and $\boldsymbol{\vartheta}^{(n)} - \boldsymbol{\vartheta} = O(n^{-1/2})$. Letting

$$
\mathbf{r}^{(n)} := \mathrm{diag}((r^{(n)}_1)^{-1/2}, \ldots, (r^{(n)}_m)^{-1/2})
$$

(see Section 3.2), define

$$\boldsymbol{\varsigma}^{(n)} := \mathrm{diag}\big(\boldsymbol{\varsigma}_{I}^{(n)}, \boldsymbol{\varsigma}_{II}^{(n)}, \boldsymbol{\varsigma}_{III}^{(n)}, \boldsymbol{\varsigma}_{IV}^{(n)}\big) := \mathrm{diag}\big(\mathbf{r}^{(n)} \otimes \mathbf{I}_k, \mathbf{r}^{(n)}, \mathbf{r}^{(n)} \otimes \mathbf{I}_{k-1}, \mathbf{r}^{(n)} \otimes \mathbf{I}_{k^2}\big) \quad (5.1)$$

and consider further sequences of the form $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$, where

$$\begin{aligned}\boldsymbol{\tau}^{(n)} &= (\boldsymbol{\tau}_{I}^{(n)\prime}, \boldsymbol{\tau}_{II}^{(n)\prime}, \boldsymbol{\tau}_{III}^{(n)\prime}, \boldsymbol{\tau}_{IV}^{(n)\prime})' \\ &= (\mathbf{t}_1^{(n)\prime}, \ldots, \mathbf{t}_m^{(n)\prime}, s_1^{(n)}, \ldots, s_m^{(n)}, \mathbf{l}_1^{(n)\prime}, \ldots, \mathbf{l}_m^{(n)\prime}, (\mathrm{vec}\,\mathbf{b}^{(1),(n)})', \ldots, (\mathrm{vec}\,\mathbf{b}^{(m),(n)})')'\end{aligned}$$

is such that $\sup_n \boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)} < \infty$ and $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)} \in \boldsymbol{\Theta}$. Under Assumption (B′), we also write $\boldsymbol{\varsigma}$ for $\lim_{n\to\infty} \boldsymbol{\varsigma}^{(n)}$.

Strong restrictions are imposed on $\boldsymbol{\tau}^{(n)} = (\boldsymbol{\tau}_{I}^{(n)\prime}, \boldsymbol{\tau}_{II}^{(n)\prime}, \boldsymbol{\tau}_{III}^{(n)\prime}, \boldsymbol{\tau}_{IV}^{(n)\prime})'$ in order for the perturbed parameter values $\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$ to belong to $\boldsymbol{\Theta}$. In particular, the perturbed orthogonal matrices should remain orthogonal; we refer to Hallin et al. (2010b) for details.

The statement of ULAN in Proposition 5.1 below still requires some additional notation. Write $\mathbf{V}^{\otimes 2}$ for the Kronecker product $\mathbf{V} \otimes \mathbf{V}$. Denoting by $\mathbf{e}_\ell$ the ℓth vector of the canonical basis of $\mathbb{R}^k$, let $\mathbf{K}_k := \sum_{i,j=1}^{k} (\mathbf{e}_i\mathbf{e}_j') \otimes (\mathbf{e}_j\mathbf{e}_i')$ be the classical $(k^2 \times k^2)$ *commutation matrix*. Define $\mathbf{H}_k$ as the $k \times k^2$ matrix such that $\mathbf{H}_k\mathrm{vec}\,(\mathbf{A}) = \mathrm{dvec}\,(\mathbf{A})$ for any $k \times k$ matrix $\mathbf{A}$. For any $k \times k$ diagonal matrix $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$, write $\mathbf{M}_k^{\boldsymbol{\Lambda}}$ for the $(k-1) \times k$ matrix $\big(-\lambda_1(\lambda_2^{-1}, \ldots, \lambda_k^{-1})' \,\vdots\, \mathbf{I}_{k-1}\big)$ and $\mathbf{L}_k^{\boldsymbol{\beta}^{(i)}, \boldsymbol{\Lambda}_i^{\mathbf{V}}}$ for $(\mathbf{L}_{k;12}^{\boldsymbol{\beta}^{(i)}, \boldsymbol{\Lambda}_i^{\mathbf{V}}}\ \mathbf{L}_{k;13}^{\boldsymbol{\beta}^{(i)}, \boldsymbol{\Lambda}_i^{\mathbf{V}}} \ldots \mathbf{L}_{k;(k-1)k}^{\boldsymbol{\beta}^{(i)}, \boldsymbol{\Lambda}_i^{\mathbf{V}}})'$, with $\mathbf{L}_{k;jh}^{\boldsymbol{\beta}^{(i)}, \boldsymbol{\Lambda}_i^{\mathbf{V}}} := (\lambda_{ih}^{\mathbf{V}} - \lambda_{ij}^{\mathbf{V}})(\boldsymbol{\beta}_h^{(i)} \otimes \boldsymbol{\beta}_j^{(i)})$. Finally, let $\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}} := (\mathbf{G}_{k;12}^{\boldsymbol{\beta}^{(i)}}\ \mathbf{G}_{k;13}^{\boldsymbol{\beta}^{(i)}} \ldots \mathbf{G}_{k;(k-1)k}^{\boldsymbol{\beta}^{(i)}})$, with $\mathbf{G}_{k;jh}^{\boldsymbol{\beta}^{(i)}} := \mathbf{e}_j \otimes \boldsymbol{\beta}_h^{(i)} - \mathbf{e}_h \otimes \boldsymbol{\beta}_j^{(i)}$, and $\boldsymbol{\nu}^{(i)} := \mathrm{diag}(\nu_{12}^{(i)}, \nu_{13}^{(i)}, \ldots, \nu_{(k-1)k}^{(i)})$ with $\nu_{jh}^{(i)} := \lambda_{ij}^{\mathbf{V}}\lambda_{ih}^{\mathbf{V}}/(\lambda_{ij}^{\mathbf{V}} - \lambda_{ih}^{\mathbf{V}})^2$. We then have the following ULAN result.

Proposition 5.1. *Let Assumptions (A) (with* $\mathrm{f} = (f_1, \ldots, f_m) \in (\mathcal{F}_a)^m$*), (B) and (D) hold. Then, the family* $\mathcal{P}_{\mathrm{f}}^{(n)} := \big\{\mathrm{P}_{\boldsymbol{\vartheta};\mathrm{f}}^{(n)} \,|\, \boldsymbol{\vartheta} \in \boldsymbol{\Theta}\big\}$ *is ULAN, with central sequence*

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}} = \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{(n)} := \Big(\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{I(n)\prime},\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{II(n)\prime},\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{III(n)\prime},\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{IV(n)\prime}\Big)',$$

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{I} = \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{I,1} \\ \vdots \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_m}^{I,m} \end{pmatrix},\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{II} = \begin{pmatrix} \Delta_{\boldsymbol{\vartheta};f_1}^{II,1} \\ \vdots \\ \Delta_{\boldsymbol{\vartheta};f_m}^{II,m} \end{pmatrix},\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{III} = \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{III,1} \\ \vdots \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_m}^{III,m} \end{pmatrix},\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};\mathrm{f}}^{IV} = \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{IV,1} \\ \vdots \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_m}^{IV,m} \end{pmatrix},$$

where (with $d_{ij} = d_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$ *and* $\mathbf{U}_{ij} = \mathbf{U}_{ij}(\boldsymbol{\theta}_i, \mathbf{V}_i)$*)*

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_i}^{I,i} := \frac{1}{\sqrt{n_i}\sigma_i}\sum_{j=1}^{n_i} \varphi_{f_i}\left(\frac{d_{ij}}{\sigma_i}\right)\mathbf{V}_i^{-1/2}\mathbf{U}_{ij}, \qquad \Delta_{\boldsymbol{\vartheta};f_i}^{II,i} := \frac{1}{2\sqrt{n_i}\sigma_i^2}\sum_{j=1}^{n_i}\left(\varphi_{f_i}\left(\frac{d_{ij}}{\sigma_i}\right)\frac{d_{ij}}{\sigma_i} - k\right),$$

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_i}^{III,i} := \frac{1}{2\sqrt{n_i}}\mathbf{M}_k^{\boldsymbol{\Lambda}_i^{\mathbf{V}}}\mathbf{H}_k\left((\boldsymbol{\Lambda}_i^{\mathbf{V}})^{-1/2}\boldsymbol{\beta}^{(i)\prime}\right)^{\otimes 2}\sum_{j=1}^{n_i}\mathrm{vec}\left(\varphi_{f_i}\left(\frac{d_{ij}}{\sigma_i}\right)\frac{d_{ij}}{\sigma_i}\,\mathbf{U}_{ij}\mathbf{U}_{ij}'\right),$$

$$\boldsymbol{\Delta}^{IV,i}_{\boldsymbol{\vartheta};f_i} := \frac{1}{2\sqrt{n_i}} \mathbf{G}_k^{\boldsymbol{\beta}^{(i)}} \mathbf{L}_k^{\boldsymbol{\beta}^{(i)},\boldsymbol{\Lambda}_i^{\mathbf{V}}} \left(\mathbf{V}_i^{\otimes 2}\right)^{-1/2} \sum_{j=1}^{n_i} \mathrm{vec}\left(\varphi_{f_i}\left(\frac{d_{ij}}{\sigma_i}\right)\frac{d_{ij}}{\sigma_i}\mathbf{U}_{ij}\mathbf{U}'_{ij}\right),$$

$i = 1, \ldots, m$, and with block-diagonal information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\mathrm{f}} := \mathrm{diag}(\boldsymbol{\Gamma}^{I}_{\boldsymbol{\vartheta};\mathrm{f}}, \boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};\mathrm{f}}, \boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};\mathrm{f}}, \boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};\mathrm{f}}), \tag{5.2}$$

where $\boldsymbol{\Gamma}^{I}_{\boldsymbol{\vartheta};\mathrm{f}} = \mathrm{diag}(\boldsymbol{\Gamma}^{I,1}_{\boldsymbol{\vartheta};f_1}, \ldots, \boldsymbol{\Gamma}^{I,m}_{\boldsymbol{\vartheta};f_m})$, $\boldsymbol{\Gamma}^{II}_{\boldsymbol{\vartheta};\mathrm{f}} = \mathrm{diag}(\Gamma^{II,1}_{\boldsymbol{\vartheta};f_1}, \ldots, \Gamma^{II,m}_{\boldsymbol{\vartheta};f_m})$, $\boldsymbol{\Gamma}^{III}_{\boldsymbol{\vartheta};\mathrm{f}} = \mathrm{diag}(\boldsymbol{\Gamma}^{III,1}_{\boldsymbol{\vartheta};f_1}, \ldots, \boldsymbol{\Gamma}^{III,m}_{\boldsymbol{\vartheta};f_m})$, *and* $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};\mathrm{f}} = \mathrm{diag}(\boldsymbol{\Gamma}^{IV,1}_{\boldsymbol{\vartheta};f_1}, \ldots, \boldsymbol{\Gamma}^{IV,m}_{\boldsymbol{\vartheta};f_m})$, *with*

$$\boldsymbol{\Gamma}^{I,i}_{\boldsymbol{\vartheta};f_i} := \frac{\mathcal{I}_k(f_i)}{k\sigma_i^2}\mathbf{V}_i^{-1}, \quad \Gamma^{II,i}_{\boldsymbol{\vartheta};f_i} := \frac{\mathcal{J}_k(f_i) - k^2}{4\sigma_i^4},$$

$$\boldsymbol{\Gamma}^{III,i}_{\boldsymbol{\vartheta};f_i} := \frac{\mathcal{J}_k(f_i)}{4k(k+2)} \mathbf{M}_k^{\boldsymbol{\Lambda}_i^{\mathbf{V}}} \mathbf{H}_k((\boldsymbol{\Lambda}_i^{\mathbf{V}})^{-1})^{\otimes 2}\left[\mathbf{I}_{k^2} + \mathbf{K}_k\right]\mathbf{H}'_k(\mathbf{M}_k^{\boldsymbol{\Lambda}_i^{\mathbf{V}}})',$$

and

$$\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta};f_i} := \frac{\mathcal{J}_k(f_i)}{4k(k+2)} \mathbf{G}_k^{\boldsymbol{\beta}^{(i)}}(\boldsymbol{\nu}^{(i)})^{-1}\left(\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}}\right)'.$$

More precisely, for any $\boldsymbol{\vartheta}^{(n)} = \boldsymbol{\vartheta} + O(n^{-1/2})$ *and any bounded sequence* $\boldsymbol{\tau}^{(n)}$*, we have, under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};\mathrm{f}}$*,*

$$\begin{aligned}\Lambda^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}/\boldsymbol{\vartheta}^{(n)};\mathrm{f}} &:= \log\left(\mathrm{dP}^{(n)}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\mathrm{f}}/\mathrm{dP}^{(n)}_{\boldsymbol{\vartheta}^{(n)};\mathrm{f}}\right) \\ &= (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta}^{(n)};\mathrm{f}} - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\mathrm{f}}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1)\end{aligned}$$

and $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}^{(n)};\mathrm{f}} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};\mathrm{f}})$*, as* $n \to \infty$*.*

Proposition 5.1, which is the multi-sample version of Theorem 2.1 in Hallin et al. (2010b), is the key result for constructing optimal inference procedures for eigenvectors and eigenvalues in multisample elliptical families. However, the standard methods for defining locally and asymptotically optimal tests under ULAN, which are based on the fact that local experiments converge to Gaussian shift experiments, do not apply here. Indeed, the parameter space $\boldsymbol{\Theta}$ is a nonlinear manifold of $\mathbb{R}^L$ (since $(\mathrm{vec}\,(\mathcal{SO}_k))^m$ is a nonlinear manifold of $\mathbb{R}^{mk^2}$). Just as in the one-sample situation, local limiting experiments therefore are *curved* Gaussian experiments. The problem of constructing optimal tests for diferentiable hypotheses in curved experiments has been considered in Hallin et al. (2010b), where general results are provided, which we apply in the present situation.

Consider a parameter value $\boldsymbol{\vartheta}_0$ satisfying $\mathcal{H}_0$ for some common eigenvector matrix $\boldsymbol{\beta}$. As explained in Section 4.2, $\boldsymbol{\beta}'\mathbf{V}_i\boldsymbol{\beta} =: \boldsymbol{\Lambda}_i^{\mathbf{V};\boldsymbol{\beta}}$, in general, is a reordered version of $\boldsymbol{\Lambda}_i^{\mathbf{V}}$, since the eigenvalues in $\boldsymbol{\Lambda}_i^{\mathbf{V}}$ are ranked in decreasing order of magnitude but not necessarily so in the locally $\boldsymbol{\beta}$-reordered (we also call it $\boldsymbol{\vartheta}_0$-reordered) $\boldsymbol{\Lambda}_i^{\mathbf{V};\boldsymbol{\beta}}$. At $\boldsymbol{\vartheta}_0$, the locally reordered $\mathbf{\Lambda}_i^{\mathbf{V};\boldsymbol{\beta}}$ is a much more natural parameter than the original $\mathbf{\Lambda}_i^{\mathbf{V}}$. In that local reparametrization, the null hypothesis $\mathcal{H}_0$ of CPC actually consists of the intersection of the nonlinear manifold $\mathbf{\Theta}$ and the linear one $\mathcal{C} := \mathbb{R}^{mk} \times (\mathbb{R}_0^+)^m \times (\mathcal{C}_{k-1})^m \times \mathcal{M}(\mathbf{1}_m \otimes \mathbf{I}_{k^2})$, where $\mathcal{M}(\mathbf{A})$ stands for the vector space spanned by the columns of $\mathbf{A}$.

Proposition 3.2 in Hallin et al. (2010b) on locally and asympotically optimal tests for differentiable hypotheses in curved ULAN experiments imply that, in the present context, a most stringent (at $\boldsymbol{\vartheta}_0 = (\boldsymbol{\vartheta}'_I, \boldsymbol{\vartheta}'_{II}, \boldsymbol{\vartheta}'_{III}, \mathbf{1}'_m \otimes (\mathrm{vec}\,(\boldsymbol{\beta}))')'$) test for $\mathcal{H}_0$ can be based on the quadratic form provided by the "classical" most stringent test for the (linear) null hypothesis consisting of the intersection of $\mathcal{C}$ and the tangent space to $\mathbf{\Theta}$ at $\boldsymbol{\vartheta}_0$. That intersection, still in the vicinity of $\boldsymbol{\vartheta}_0$, reduces to

$$\left\{ \begin{pmatrix} \boldsymbol{\vartheta}_I + n^{-1/2}\boldsymbol{\varsigma}_I^{(n)}\boldsymbol{\tau}_I^{(n)} \\ \boldsymbol{\vartheta}_{II} + n^{-1/2}\boldsymbol{\varsigma}_{II}^{(n)}\boldsymbol{\tau}_{II}^{(n)} \\ \boldsymbol{\vartheta}_{III} + n^{-1/2}\boldsymbol{\varsigma}_{III}^{(n)}\boldsymbol{\tau}_{III}^{(n)} \\ \mathrm{vec}\,(\boldsymbol{\beta} + n^{-1/2}(r_1^{(n)})^{-1/2}\mathbf{b}^{(1),(n)}) \\ \vdots \\ \mathrm{vec}\,(\boldsymbol{\beta} + n^{-1/2}(r_m^{(n)})^{-1/2}\mathbf{b}^{(m),(n)}) \end{pmatrix} \begin{array}{l} \text{such that} \quad \boldsymbol{\beta}'\mathbf{b}^{(i),(n)} + (\mathbf{b}^{(i),(n)})'\boldsymbol{\beta} = \mathbf{0},\ i = 1, \ldots, m \\ \qquad \text{and } (r_1^{(n)})^{-1/2}\mathbf{b}^{(1),(n)} = \ldots = (r_m^{(n)})^{-1/2}\mathbf{b}^{(m),(n)} \end{array} \right\}.$$

Solving this system leads to

$$\boldsymbol{\varsigma}_{IV}^{(n)}\boldsymbol{\tau}_{IV}^{(n)} = \big((r_1^{(n)})^{-1/2}(\mathrm{vec}\,\mathbf{b}^{(1),(n)})', \ldots, (r_m^{(n)})^{-1/2}(\mathrm{vec}\,\mathbf{b}^{(m),(n)})'\big)' \in \mathcal{M}(\mathbf{\Psi}),$$

with

$$\mathbf{\Psi} := \mathbf{1}_m \otimes \begin{pmatrix} \mathbf{I}_k - \boldsymbol{\beta}_1\boldsymbol{\beta}'_1 & -\boldsymbol{\beta}_2\boldsymbol{\beta}'_1 & \ldots & -\boldsymbol{\beta}_k\boldsymbol{\beta}'_1 \\ -\boldsymbol{\beta}_1\boldsymbol{\beta}'_2 & \mathbf{I}_k - \boldsymbol{\beta}_2\boldsymbol{\beta}'_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\boldsymbol{\beta}_k\boldsymbol{\beta}'_{k-1} \\ -\boldsymbol{\beta}_1\boldsymbol{\beta}'_k & \ldots & -\boldsymbol{\beta}_{k-1}\boldsymbol{\beta}'_k & \mathbf{I}_k - \boldsymbol{\beta}_k\boldsymbol{\beta}'_k \end{pmatrix},$$

where $\boldsymbol{\beta}_\ell$ denotes $\boldsymbol{\beta}$'s ℓth column. Hence, the null hypothesis of CPC, locally at $\boldsymbol{\vartheta}_0$, takes the form $\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)} \in \mathcal{M}(\mathbf{\Upsilon})$, where

$$\mathbf{\Upsilon} := \mathrm{diag}\big(\mathbf{\Upsilon}^{I}, \mathbf{\Upsilon}^{II}, \mathbf{\Upsilon}^{III}, \mathbf{\Upsilon}^{IV}\big) := \mathrm{diag}\big(\mathbf{I}_{mk}, \mathbf{I}_m, \mathbf{I}_{m(k-1)}, \mathbf{\Psi}\big).$$

It then follows from Hallin et al. (2010b, Section 4.1) that, for given f, a locally and asymptotically most stringent test $\phi_{\mathrm{f}}^{(n)}$, say, rejects $\mathcal{H}_0$ for large values of $Q_{\hat{\boldsymbol{\vartheta}},\mathrm{f}}^{(n)}$, where (throughout, we denote by $\mathbf{A}^-$ the Moore-Penrose inverse of $\mathbf{A}$)

$$\begin{aligned} Q_{\boldsymbol{\vartheta},\mathrm{f}}^{(n)} &:= (\mathbf{\Delta}_{\boldsymbol{\vartheta},\mathrm{f}})' \Big(\mathbf{\Gamma}_{\boldsymbol{\vartheta},\mathrm{f}}^{-} - (\boldsymbol{\varsigma}^{(n)})^{-1}\mathbf{\Upsilon}\big[\mathbf{\Upsilon}'(\boldsymbol{\varsigma}^{(n)})^{-1}\mathbf{\Gamma}_{\boldsymbol{\vartheta},\mathrm{f}}(\boldsymbol{\varsigma}^{(n)})^{-1}\mathbf{\Upsilon}\big]^{-}\mathbf{\Upsilon}'(\boldsymbol{\varsigma}^{(n)})^{-1}\Big)\mathbf{\Delta}_{\boldsymbol{\vartheta},\mathrm{f}} \qquad (5.3) \\ &= (\mathbf{\Delta}_{\boldsymbol{\vartheta},\mathrm{f}}^{IV})' \Big((\mathbf{\Gamma}_{\boldsymbol{\vartheta},\mathrm{f}}^{IV})^{-} - (\boldsymbol{\varsigma}_{IV}^{(n)})^{-1}\mathbf{\Upsilon}^{IV}\big[(\mathbf{\Upsilon}^{IV})'(\boldsymbol{\varsigma}_{IV}^{(n)})^{-1}\mathbf{\Gamma}_{\boldsymbol{\vartheta},\mathrm{f}}^{IV}(\boldsymbol{\varsigma}_{IV}^{(n)})^{-1}\mathbf{\Upsilon}^{IV}\big]^{-}(\mathbf{\Upsilon}^{IV})'(\boldsymbol{\varsigma}_{IV}^{(n)})^{-1}\Big)\mathbf{\Delta}_{\boldsymbol{\vartheta},\mathrm{f}}^{IV}, \end{aligned}$$

and $\hat{\boldsymbol{\vartheta}} := \hat{\boldsymbol{\vartheta}}^{(n)}$ denotes a sequence of estimators satisfying the following Assumption (E) with $\mathcal{K}$ reducing to {f}.

Assumption (E). We say that a sequence of estimators $\boldsymbol{\vartheta}^{(n)}$ of $\boldsymbol{\vartheta}$, $n \in \mathbb{N}$, satisfies Assumption (E) for some given collection $\mathcal{K}$ of m-tuples of standardized radial densities if, as $n \to \infty$ as in Assumption (B), $\boldsymbol{\vartheta}^{(n)}$ is

(E1) *constrained:* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}\big[\boldsymbol{\vartheta}^{(n)} \in \mathcal{H}_0\big] = 1$ for all n, $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$, and $\mathrm{g} \in \mathcal{K}$;

(E2) $n^{1/2}(\boldsymbol{\varsigma}^{(n)})^{-1}$*-consistent:* for all $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$, $n^{1/2}\big(\boldsymbol{\varsigma}^{(n)}\big)^{-1}(\boldsymbol{\vartheta}^{(n)} - \boldsymbol{\vartheta}_0) = O_{\mathrm{P}}(1)$, as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, for all $\mathrm{g} \in \mathcal{K}$;

(E3) *locally asymptotically discrete:* for all $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ and all $c > 0$, there exists $M = M(c) > 0$ such that the number of possible values of $\boldsymbol{\vartheta}^{(n)}$ in balls of the form $\{\mathbf{t} \in \mathbb{R}^L : n^{1/2}\|\big(\boldsymbol{\varsigma}^{(n)}\big)^{-1}(\mathbf{t} - \boldsymbol{\vartheta}_0)\| \leq c\}$ is bounded by M, uniformly in n.

Assumption (E3) is a theoretical assumption that has no impact in practice (see pages 125 or 188 of Le Cam and Yang (2000) for a discussion). Any estimator satisfying (E1) and (E2) can be turned into an estimator also satisfying (E3) by discretization (see, e.g., Hallin et al. (2006)), a fact we will no further emphasize in the notation by tacitly assuming, in the statement of asymptotic results, that any $\boldsymbol{\vartheta}^{(n)}$, when necessary, has been adequately discretized.

The sequences of tests $\phi^{(n)}_{\mathrm{f}}$ associated with the m-tuple f achieve local asymptotic optimality at f. Moreover, they are of a purely parametric nature since, in general, they are valid at f only—that is, they achieve the correct nominal asymptotic level under correctly specified f only, even when based on an estimator $\boldsymbol{\vartheta}^{(n)}$ satisfying Assumption (E) under a broad collection $\mathcal{K}$ of densities. An exception is the Gaussian test $\phi^{(n)}_{\mathcal{N}}$ associated with an m-tuple of Gaussian radial densities which, with a Gaussian MLE $\boldsymbol{\vartheta}^{(n)}$, remains valid under any m-tuple $\mathrm{g} = (g_1, \ldots, g_m)$ such that g_i has Gaussian kurtosis (that is, in the notation of Section 6 below, $\kappa_k(g_i) = 0$) for all $i = 1, \ldots, m$ (this, of course, requires finite fourth-order moments). Clearly, this is somewhat unsatisfactory in practice, and there is a need to define alternative optimal tests, that remain valid under much broader conditions. The next two sections are devoted to the construction of such tests.

6. Gaussian and pseudo-Gaussian tests.

In this section, we construct a pseudo-Gaussian version $\phi^{(n)\dagger}_{\mathcal{N}}$ of the Gaussian test $\phi^{(n)}_{\mathcal{N}}$, that is, a test that shares the optimality properties of $\phi^{(n)}_{\mathcal{N}}$ in the multinormal case, while remaining valid under a much broader class of densities—namely, the class of all (possibly heterokurtic) m-tuples of elliptic densities with finite fourth-order moments. Our construction is based on a general method proposed by Hallin and Paindaveine (2008a), which exploits the ULAN structure of the experiment. Finally, we show that this pseudo-Gaussian test $\phi^{(n)\dagger}_{\mathcal{N}}$ asymptotically coincides with the test $\phi^{(n)}_{\mathrm{HPV}}$ proposed, on heuristic grounds, in Hallin *et al.* (2010a), the optimality properties of which thus follow.

Let $(\mathcal{F}^4_1)^m$ denote the collection of m-tuples of standardized radial densities yielding finite fourth-order moments in each population:

$$(\mathcal{F}_1^4)^m := \Big\{ \mathrm{g} = (g_1, \ldots, g_m) \in (\mathcal{F}_1)^m : E_k(g_i) := \int_0^1 (\tilde{G}_{ik}^{-1}(u))^4 \mathrm{d}u < \infty,\ i = 1, \ldots, m \Big\},$$

where $r \mapsto \tilde{G}_{ik}(r) := (\mu_{k-1;g_i})^{-1} \int_0^r s^{k-1} g_i(s)\, \mathrm{d}s$ stands for the distribution function, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}$, of the $d_{ij}(\boldsymbol{\vartheta})$'s, $j = 1, \ldots, n_i$. Then, writing $D_k(g_i) := \int_0^1 (\tilde{G}_{ik}^{-1}(u))^2 \mathrm{d}u$,

$$\kappa_k(g_i) := \frac{k}{k+2} \times \frac{E_k(g_i)}{D_k^2(g_i)} - 1,$$

for any $\mathrm{g} \in (\mathcal{F}_1^4)^m$, is a measure of kurtosis in the ith elliptic population under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}$; see, e.g., page 54 of Anderson (2003). If g_i is Gaussian, $E_k(g_i) = k(k+2)/a_k^2$ and $D_k(g_i) = k/a_k$, so that $\kappa_k(g_i) = 0$.

Since the optimal Gaussian test $\phi_{\mathcal{N}}^{(n)}$ of Section 5 is based on a quadratic form in the eigenvector part $\boldsymbol{\Delta}^{\scriptscriptstyle IV}_{\boldsymbol{\vartheta};\phi}$ of the Gaussian central sequence, defining a pseudo-Gaussian version of $\phi_{\mathcal{N}}^{(n)}$ clearly requires controlling the asymptotic behavior of $\boldsymbol{\Delta}^{\scriptscriptstyle IV}_{\boldsymbol{\vartheta};\phi}$ also away from the Gaussian case. This is made possible by the following result.

Lemma 6.1. *Assume that (A), (B), and (D′) hold. Fix any $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ (the eigenvalues are the $\boldsymbol{\vartheta}_0$-ordered eigenvalues) and $\mathrm{g} \in (\mathcal{F}_1^4)^m$. Then,*

(i) under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, $\boldsymbol{\Delta}^{\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi}$ is asymptotically normal, with mean zero and covariance matrix $\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi} := \mathrm{diag}\big(\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV,1}_{\boldsymbol{\vartheta}_0;\phi}, \ldots, \boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV,m}_{\boldsymbol{\vartheta}_0;\phi}\big)$, where $\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV,i}_{\boldsymbol{\vartheta}_0;\phi} := \frac{a_k^2 E_k(g_i)}{4k(k+2)}\, \mathbf{G}_k^{\boldsymbol{\beta}} (\boldsymbol{\nu}^{(i)})^{-1} (\mathbf{G}_k^{\boldsymbol{\beta}})'$;

(ii) reinforcing (D′) into (D) and defining $\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}} := \mathrm{diag}\big(\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV,1}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}}, \ldots, \boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV,m}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}}\big)$, with $\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV,i}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}} := \frac{a_k D_k(g_i)}{4k}\, \mathbf{G}_k^{\boldsymbol{\beta}} (\boldsymbol{\nu}^{(i)})^{-1} (\mathbf{G}_k^{\boldsymbol{\beta}})'$, we have that

$$\boldsymbol{\Delta}^{\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0 + n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\phi} - \boldsymbol{\Delta}^{\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi} + \boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}} \boldsymbol{\tau}^{(n)}_{\scriptscriptstyle IV}$$

is $o_{\mathrm{P}}(1)$ as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$;

(iii) still with (D′) reinforced into (D), $\boldsymbol{\Delta}^{\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi} - \boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}} \boldsymbol{\tau}^{(n)}_{\scriptscriptstyle IV}$ is asymptotically normal, with mean zero and covariance matrix $\boldsymbol{\Gamma}^{\mathrm{g},\scriptscriptstyle IV}_{\boldsymbol{\vartheta}_0;\phi}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0 + n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\mathrm{g}}$ for any $\mathrm{g} \in (\mathcal{F}_a^4)^m := (\mathcal{F}_1^4)^m \cap (\mathcal{F}_a)^m$.

Point (i) of this Lemma directly follows from the multivariate central limit theorem. Note that Assumption (D′) is sufficient for asymptotic normality since the common value $\boldsymbol{\beta}$ of the eigenvector matrix is well identified under Assumption (D′). However, points (ii) and (iii) require ULAN and therefore Assumption (D); they directly follow from Lemma 4.2 in Hallin *et* al. (2010b); the proof is therefore omitted.

Transposed to the present context, and temporarily assuming that the actual $\mathrm{g} \in (\mathcal{F}_1^4)^m$ is known, the pseudo-Gaussian test of Hallin and Paindaveine (2008a) is rejecting the null hypothesis of CPC for large values of

$$Q^{\mathcal{N}(n)}_{\boldsymbol{\vartheta}_0,\mathrm{g}} := \big(\boldsymbol{\Delta}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;\phi}\big)'(\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta}_0;\phi})^{\perp}\boldsymbol{\Delta}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;\phi}, \tag{6.1}$$

with

$$(\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi})^{\perp} := (\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi})^{-} - (\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi})^{-}\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi,\mathrm{g}}(\varsigma^{(n)}_{\mathit{IV}})^{-1}\boldsymbol{\Upsilon}^{\mathit{IV}}$$
$$\Big[\big(\boldsymbol{\Upsilon}^{\mathit{IV}}\big)'(\varsigma^{(n)}_{\mathit{IV}})^{-1}\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi,\mathrm{g}}(\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi})^{-}\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi,\mathrm{g}}(\varsigma^{(n)}_{\mathit{IV}})^{-1}\boldsymbol{\Upsilon}^{\mathit{IV}}\Big]^{-}\big(\boldsymbol{\Upsilon}^{\mathit{IV}}\big)'(\varsigma^{(n)}_{\mathit{IV}})^{-1}\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi,\mathrm{g}}(\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi})^{-},$$

where $\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi}$ and $\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV}}_{\boldsymbol{\vartheta};\phi,\mathrm{g}}$ are defined in Lemma 6.1, and eigenvalues have been $\boldsymbol{\vartheta}_0$-reordered as explained in Sections 3 and 4. Now, using the fact that

$$\big(\boldsymbol{\Gamma}^{\mathrm{g},\mathit{IV},i}_{\boldsymbol{\vartheta};\phi}\big)^{-} = \frac{k(k+2)}{a_k^2 E_k(g_i)}\mathbf{G}^{\boldsymbol{\beta}^{(i)}}_k \boldsymbol{\nu}^{(i)}(\mathbf{G}^{\boldsymbol{\beta}^{(i)}}_k)',$$

the quadratic form $Q^{\mathcal{N}(n)}_{\boldsymbol{\vartheta}_0,\mathrm{g}}$ after some algebra rewrites

$$Q^{\mathcal{N}(n)}_{\boldsymbol{\vartheta}_0,\mathrm{g}} = \sum_{i=1}^{m}\sum_{1\le j<j'\le k}\frac{n_i}{1+\kappa_k(g_i)}\,(\boldsymbol{\beta}'_j\mathbf{S}^{(n)}_{\phi,i}\boldsymbol{\beta}_{j'})^2 \tag{6.2}$$
$$-\sum_{i,i'=1}^{m}\sum_{1\le j<j'\le k}\frac{n_i n_{i'}}{n}\,\frac{1}{(1+\kappa_k(g_i))(1+\kappa_k(g_{i'}))}\,\frac{\nu_{jj'}(\mathrm{g})}{(\nu^{(i)}_{jj'}\nu^{(i')}_{jj'})^{1/2}}\,(\boldsymbol{\beta}'_j\mathbf{S}^{(n)}_{\phi,i}\boldsymbol{\beta}_{j'})\,(\boldsymbol{\beta}'_j\mathbf{S}^{(n)}_{\phi,i'}\boldsymbol{\beta}_{j'}),$$

where

$$\mathbf{S}^{(n)}_{\phi,i} := \frac{k}{\sigma^2 D_k(g_i)}\boldsymbol{\beta}(\boldsymbol{\Lambda}^{\mathbf{V};\boldsymbol{\beta}}_i)^{-1/2}\boldsymbol{\beta}'\Big[\frac{1}{n_i}\sum_{j=1}^{n_i}(\mathbf{X}_{ij}-\boldsymbol{\theta}_i)(\mathbf{X}_{ij}-\boldsymbol{\theta}_i)'\Big]\boldsymbol{\beta}(\boldsymbol{\Lambda}^{\mathbf{V};\boldsymbol{\beta}}_i)^{-1/2}\boldsymbol{\beta}'$$

and

$$\mathrm{diag}(\nu_{12}(\mathrm{g}),\ldots,\nu_{(k-1)k}(\mathrm{g})) := \left(\sum_{i=1}^{m}\frac{r_i^{(n)}}{1+\kappa_k(g_i)}\,(\boldsymbol{\nu}^{(i)})^{-1}\right)^{-1} =: \boldsymbol{\nu}(\mathrm{g}).$$

In order to obtain a genuine test statistic (that is, a random variable that does not depend anymore on $\boldsymbol{\vartheta}_0$ nor g) which nevertheless, under any $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$ (with $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ and $\mathrm{g} \in (\mathcal{F}_1^4)^m$) and contiguous alternatives, is asymptotically equivalent to $Q^{\mathcal{N}(n)}_{\boldsymbol{\vartheta}_0,\mathrm{g}}$, it is sufficient to

(a) replace $\boldsymbol{\vartheta}_0$ in (6.2) with some estimator $\boldsymbol{\vartheta}^{(n)}$ satisfying Assumption (E) for the class $\mathcal{K} = (\mathcal{F}_1^4)^m$, and
(b) replace the coefficients $D_k(g_i)$ and the kurtoses $\kappa_k(g_i)$ with consistent (still under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$, $\mathrm{g} \in (\mathcal{F}_1^4)^m$) estimators $\hat{D}_i^{(n)}$ and $\hat{\kappa}_i^{(n)}$, respectively.

In this pseudo-Gaussian context, a natural estimator for $\boldsymbol{\vartheta}_0$ is

$$\boldsymbol{\vartheta}^{(n)}_{\mathcal{N}} := \Big(\bar{\mathbf{X}}'_1,\ldots,\bar{\mathbf{X}}'_m,\hat{\sigma}^2_1,\ldots,\hat{\sigma}^2_m, \tag{6.3}$$
$$(\mathrm{d}\overset{\circ}{\mathrm{vec}}\,\hat{\boldsymbol{\Lambda}}_1)'/\textstyle\prod_{j=1}^k(\hat{\lambda}_{1j})^{1/k},\ldots,(\mathrm{d}\overset{\circ}{\mathrm{vec}}\,\hat{\boldsymbol{\Lambda}}_m)'/\prod_{j=1}^k(\hat{\lambda}_{mj})^{1/k},\mathbf{1}'_m\otimes(\mathrm{vec}\,\hat{\boldsymbol{\beta}})'\Big)',$$

where $\bar{\mathbf{X}}_i := n_i^{-1}\sum_{j=1}^{n_i}\mathbf{X}_{ij}$, $\hat{\mathbf{\Lambda}}_i = \mathrm{diag}(\hat{\lambda}_{i1},\ldots,\hat{\lambda}_{ik})$, $i=1,\ldots,m$, and $\hat{\boldsymbol{\beta}}$ are the maximum likelihood estimators of the corresponding parameters in the CPC model (see Flury 1986), and $\hat{\sigma}_i^2$ denotes the empirical median of the $d_{ij}^2(\bar{\mathbf{X}}_i, \hat{\boldsymbol{\beta}}\hat{\mathbf{\Lambda}}_i\hat{\boldsymbol{\beta}}'/\prod_{j=1}^k(\hat{\lambda}_{ij})^{1/k})$'s, $j=1,\ldots,n_i$. Note that the estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{\Lambda}}_i$, resulting from the Flury and Gautschi (1986) algorithm, do not provide consistent estimators of $\boldsymbol{\beta}$ and the $\boldsymbol{\vartheta}_0$-reordered eigenvalues matrices $\mathbf{\Lambda}_i^{\mathbf{V};\boldsymbol{\beta}}$, respectively, because of the possibly different ordering of eigenvalues (and also because the determinant of $\hat{\mathbf{\Lambda}}_i$ is not equal to one in general). However, Flury (1986) shows that they are root-n consistent for $\mathbf{N}^{\Pi}\boldsymbol{\beta} = \mathbf{N}^{\Pi}\boldsymbol{\beta}^{(1)} = \mathbf{N}^{\Pi}\mathbf{M}_2^{\Pi_2}\boldsymbol{\beta}^{(2)} = \ldots = \mathbf{N}^{\Pi}\mathbf{M}_m^{\Pi_m}\boldsymbol{\beta}^{(m)}$ and the corresponding reordered version of the $\mathbf{\Lambda}_i^{\mathbf{V};\boldsymbol{\beta}}$'s for some global permutation matrix $\mathbf{N}^{\Pi}$. Now, since both the null hypothesis and the test statistic $Q_{\boldsymbol{\vartheta}_0,g}^{\mathcal{N}(n)}$ are invariant with respect to such global permutations, no reordering of the eigenvalues is needed here. Note that $D_k(g_i)$ is consistently estimated by $k\prod_{j=1}^k(\hat{\lambda}_{ij})^{1/k}/\hat{\sigma}_i^2$, $i=1,\ldots,m$. Finally, an obvious choice for $\hat{\kappa}_i^{(n)}$ is then

$$\hat{\kappa}_i^{(n)} := \frac{k}{k+2}\times\frac{n_i^{-1}\sum_{j=1}^{n_i} d_{ij}^4(\bar{\mathbf{X}}_i, \hat{\boldsymbol{\beta}}\hat{\mathbf{\Lambda}}_i\hat{\boldsymbol{\beta}}')}{\left(n_i^{-1}\sum_{j=1}^{n_i} d_{ij}^2(\bar{\mathbf{X}}_i, \hat{\boldsymbol{\beta}}\hat{\mathbf{\Lambda}}_i\hat{\boldsymbol{\beta}}')\right)^2} - 1.$$

Letting $\mathbf{S}_i^{(n)} := {n_i}^{-1}\sum_{j=1}^{n_i}(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)(\mathbf{X}_{ij}-\bar{\mathbf{X}}_i)'$, this leads to the test statistic

$$Q_{\mathcal{N}}^{(n)\dagger} := \sum_{i=1}^m \sum_{1\leq j<j'\leq k} \frac{n_i}{1+\hat{\kappa}_i^{(n)}}\,(\hat{\lambda}_{ij}\hat{\lambda}_{ij'})^{-1}\,(\hat{\boldsymbol{\beta}}_j'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}}_{j'})^2 \tag{6.4}$$
$$-\sum_{i,i'=1}^m \sum_{1\leq j<j'\leq k} \frac{n_i n_{i'}}{n}\,\frac{(\hat{\lambda}_{ij}\hat{\lambda}_{ij'})^{-1/2}(\hat{\lambda}_{i'j}\hat{\lambda}_{i'j'})^{-1/2}}{(1+\hat{\kappa}_i^{(n)})(1+\hat{\kappa}_{i'}^{(n)})}\,\frac{\hat{\nu}_{jj'}}{(\hat{\nu}_{jj'}^{(i)}\,\hat{\nu}_{jj'}^{(i')})^{1/2}}\,(\hat{\boldsymbol{\beta}}_j'\mathbf{S}_i^{(n)}\hat{\boldsymbol{\beta}}_{j'})\,(\hat{\boldsymbol{\beta}}_j'\mathbf{S}_{i'}^{(n)}\hat{\boldsymbol{\beta}}_{j'}),$$

where we write $\hat{\nu}_{jj'}^{(i)}$ and $\hat{\nu}_{jj'}$, respectively, for the $\nu_{jj'}^{(i)}$ and $\nu_{jj'}(\mathrm{g})$ values computed from the $\hat{\lambda}_{ij}$ and $\hat{\kappa}_i^{(n)}$ estimators. The resulting pseudo-Gaussian test $\phi_{\mathcal{N}}^{(n)\dagger}$ rejects the null hypothesis of CPC, at asymptotic level α, as soon as $Q_{\mathcal{N}}^{(n)\dagger}$ exceeds the α-upper quantile of the chi-square distribution with $(m-1)k(k-1)/2$ degrees of freedom.

To investigate the asymptotic behavior of this pseudo-Gaussian test under local alternatives, we consider perturbations $\boldsymbol{\vartheta}_0 + n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$ such that, letting Assumption (B′) hold and putting $\boldsymbol{\varsigma\tau} := \lim_{n\to\infty}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$, with

$$\boldsymbol{\varsigma}_{IV}^{(n)}\boldsymbol{\tau}_{IV}^{(n)} = ((r_1^{(n)})^{-1/2}(\mathrm{vec}\,\mathbf{b}^{(1),(n)})', \ldots, (r_m^{(n)})^{-1/2}(\mathrm{vec}\,\mathbf{b}^{(m),(n)})')'$$

and

$$\boldsymbol{\varsigma}_{IV}\boldsymbol{\tau}_{IV} = (r_1^{-1/2}(\mathrm{vec}\,\mathbf{b}^{(1)})', \ldots, r_m^{-1/2}(\mathrm{vec}\,\mathbf{b}^{(m)})')',$$

we still have, for all $i=1,\ldots,m$, $\boldsymbol{\beta}'\mathbf{b}^{(i)} + (\mathbf{b}^{(i)})'\boldsymbol{\beta} = \mathbf{0}$ (where $\boldsymbol{\beta}$ is the common value, under $\boldsymbol{\vartheta}_0$, of the m eigenvector matrices). Assume furthermore that the corresponding perturbed value of $\boldsymbol{\vartheta}_0$ does not belong to $\mathcal{H}_0$ anymore (does not belong to the linear manifold $\mathcal{C}$). Letting, for any such $\boldsymbol{\vartheta}_0$ and any $\mathrm{g} \in (\mathcal{F}_1^4)^m$,

$$\mathbf{C}^{\mathcal{N}}_{\boldsymbol{\vartheta}_0;\mathrm{g}} := \mathrm{diag}\bigg(\frac{1}{1+\kappa_k(g_1)}\,(\boldsymbol{\nu}^{(1)})^{-1}, \ldots, \frac{1}{1+\kappa_k(g_m)}\,(\boldsymbol{\nu}^{(m)})^{-1} \bigg)$$

and

$$\mathbf{D}^{\mathcal{N}(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}} := \mathbf{C}^{\mathcal{N}}_{\boldsymbol{\vartheta}_0;\mathrm{g}} - \mathbf{C}^{\mathcal{N}}_{\boldsymbol{\vartheta}_0;\mathrm{g}} \big[((\mathbf{r}^{(n)})^{-1}\mathbf{1}_m\mathbf{1}_m'(\mathbf{r}^{(n)})^{-1}) \otimes \boldsymbol{\nu}(\mathrm{g}) \big]\, \mathbf{C}^{\mathcal{N}}_{\boldsymbol{\vartheta}_0;\mathrm{g}},$$

standard algebra yields

$$\begin{aligned} l^{\mathcal{N}}_{\boldsymbol{\vartheta}_0,\boldsymbol{\tau};\mathrm{g}} &:= \lim_{n\to\infty} \Big\{ (\boldsymbol{\tau}^{(n)}_{\mathit{IV}})'(\mathbf{I}_m \otimes \mathbf{G}^{\boldsymbol{\beta}}_k)\, \mathbf{D}^{\mathcal{N}(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}} (\mathbf{I}_m \otimes \mathbf{G}^{\boldsymbol{\beta}}_k)'(\boldsymbol{\tau}^{(n)}_{\mathit{IV}}) \Big\} \\ &= \sum_{i,i'=1}^{m} (\mathrm{vec}\,\mathbf{b}^{(i)})'\mathbf{G}^{\boldsymbol{\beta}}_k \Big[\delta_{ii'}\mathbf{T}^{\mathcal{N}(i,i')}_{\mathrm{g}} - (r_i\, r_{i'})^{1/2}\, \mathbf{T}^{\mathcal{N}(i,i')}_{\mathrm{g}}\, \boldsymbol{\nu}(\mathrm{g}) \mathbf{T}^{\mathcal{N}(i,i')}_{\mathrm{g}} \Big] \big(\mathbf{G}^{\boldsymbol{\beta}}_k\big)'(\mathrm{vec}\,\mathbf{b}^{(i')}), \end{aligned} \tag{6.5}$$

where $\mathbf{T}^{\mathcal{N}(i,i')}_{\mathrm{g}} := \big((1+\kappa_k(g_i))(1+\kappa_k(g_{i'}))\big)^{-1/2}\, (\boldsymbol{\nu}^{(i)})^{-1/2}(\boldsymbol{\nu}^{(i')})^{-1/2}$; $\mathbf{r}^{(n)}$ was defined on Page 10. The following result then summarizes the asymptotic properties of $Q^{(n)\dagger}_{\mathcal{N}}$ and $\phi^{(n)\dagger}_{\mathcal{N}}$.

Proposition 6.1. *Assume that (A), (B), and (D′) hold. Then,*

(i) $Q^{(n)\dagger}_{\mathcal{N}}$ *is asymptotically chi-square with* $(m-1)k(k-1)/2$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0}\bigcup_{\mathrm{g}\in(\mathcal{F}_1^4)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}\}$*, and (provided that (D′) is reinforced into (D) and (B) into (B′)) asymptotically noncentral chi-square, still with* $(m-1)k(k-1)/2$ *degrees of freedom, but with noncentrality parameter* $l^{\mathcal{N}}_{\boldsymbol{\vartheta},\boldsymbol{\tau};\mathrm{g}}$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\mathrm{g}}$, $\boldsymbol{\vartheta}\in\mathcal{H}_0$, $\boldsymbol{\varsigma\tau} := \lim_{n\to\infty}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$ *as described above, and* $\mathrm{g}\in(\mathcal{F}_a^4)^m$*;*

(ii) $\phi^{(n)\dagger}_{\mathcal{N}}$ *has asymptotic level* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0}\bigcup_{\mathrm{g}\in(\mathcal{F}_1^4)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}\}$*;*

(iii) letting (D′) be reinforced into (D), $\phi^{(n)\dagger}_{\mathcal{N}}$ *is locally and asymptotically most stringent, at asymptotic level* α*, for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0}\bigcup_{\mathrm{g}\in(\mathcal{F}_1^4)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{H}_0}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\phi}\}$.

One can easily check that $\phi^{(n)\dagger}_{\mathcal{N}}$ actually coincides with the test $\phi^{(n)}_{\mathrm{HPV}}$ proposed in Hallin et al. (2010a); theorem 6.1 therefore clarifies the asymptotic optimality properties of the latter.

7. Optimal rank-based tests.

7.1. A rank-based central sequence for eigenvectors.

Even though the pseudo-Gaussian test $\phi^{(n)\dagger}_{\mathcal{N}}$ of the previous section is valid under a broad class of densities, it still requires finite fourth-order moments, and may be poorly robust, since it is based on empirical covariance matrices. In this section, we show how ranks (actually, a multivariate generalization of signed ranks) allow us to improve on the performances of pseudo-Gaussian tests both in terms of validity and efficiency.

A general result by Hallin and Werker (2003) implies that, in adaptive semiparametric models for which fixed-f submodels are ULAN and fixed-$\boldsymbol{\vartheta}$ submodels are generated by a group $\mathcal{G}_{\boldsymbol{\vartheta}}$ of transformations (acting on the observation space), invariant versions of central sequences exist under very general assumptions. In the present case, the ULAN structure of fixed-f submodels is established in Section 5. As for the fixed-$\boldsymbol{\vartheta}$ submodels, consider the group $\mathcal{G}^{\boldsymbol{\vartheta}},_\circ$ of *continuous monotone radial transformations* $\mathcal{g}_h$ of the form

$$\begin{aligned}\mathbf{X} &\mapsto \mathcal{g}_h(\mathbf{X}_{11},\ldots,\mathbf{X}_{mn_m})\\ &:= (\boldsymbol{\theta}_1 + h_1(d_{11}(\boldsymbol{\theta}_1,\boldsymbol{\beta}^{(1)}\boldsymbol{\Lambda}_1^{\mathbf{V}}\boldsymbol{\beta}^{(1)\prime}))\boldsymbol{\beta}^{(1)}(\boldsymbol{\Lambda}_1^{\mathbf{V}})^{1/2}\boldsymbol{\beta}^{(1)\prime}\mathbf{U}_{11}(\boldsymbol{\theta}_1,\boldsymbol{\beta}^{(1)}\boldsymbol{\Lambda}_1^{\mathbf{V}}\boldsymbol{\beta}^{(1)\prime}),\ldots,\\ &\qquad \boldsymbol{\theta}_m + h_m(d_{mn_m}(\boldsymbol{\theta}_m,\boldsymbol{\beta}^{(m)}\boldsymbol{\Lambda}_m^{\mathbf{V}}\boldsymbol{\beta}^{(m)\prime}))\boldsymbol{\beta}^{(m)}(\boldsymbol{\Lambda}_m^{\mathbf{V}})^{1/2}\boldsymbol{\beta}^{(m)\prime}\mathbf{U}_{mn_m}(\boldsymbol{\theta}_m,\boldsymbol{\beta}^{(m)}\boldsymbol{\Lambda}_m^{\mathbf{V}}\boldsymbol{\beta}^{(m)\prime})),\end{aligned}$$

where for all $i = 1,\ldots,m$, $h_i : \mathbb{R}^+ \to \mathbb{R}^+$ is continuous, monotone increasing, and such that $h_i(0) = 0$ and $\lim_{r\to\infty} h_i(r) = \infty$. Letting $\underline{\sigma}^2 := (\sigma_1^2,\ldots,\sigma_m^2)'$, this group is a generating group for the submodel $\bigcup_{\underline{\sigma}^2}\bigcup_{\mathrm{f}}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^I,\underline{\sigma}^2,\boldsymbol{\vartheta}^{III},\boldsymbol{\vartheta}^{IV};\mathrm{f}}\}$ (a nonparametric family). The invariance principle suggests basing inference on statistics that are measurable with respect to the corresponding maximal invariant, namely the vectors $(\mathbf{U}_{11},\ldots,\mathbf{U}_{mn_m})$ (a multivariate generalization of signs) along with the vector $(R_{11},\ldots,R_{mn_m})$ of ranks, where $\mathbf{U}_{ij} = \mathbf{U}_{ij}(\boldsymbol{\theta}_i,\boldsymbol{\beta}^{(i)}\boldsymbol{\Lambda}_i^{\mathbf{V}}\boldsymbol{\beta}^{(i)\prime})$, and $R_{ij} = R_{ij}(\boldsymbol{\theta}_i,\boldsymbol{\beta}^{(i)}\boldsymbol{\Lambda}_i^{\mathbf{V}}\boldsymbol{\beta}^{(i)\prime})$ denotes the rank of $d_{ij}(\boldsymbol{\theta}_i,\boldsymbol{\beta}^{(i)}\boldsymbol{\Lambda}_i^{\mathbf{V}}\boldsymbol{\beta}^{(i)\prime})$ among $d_{i1}(\boldsymbol{\theta}_i,\boldsymbol{\beta}^{(i)}\boldsymbol{\Lambda}_i^{\mathbf{V}}\boldsymbol{\beta}^{(i)\prime}),\ldots,d_{in_i}(\boldsymbol{\theta}_i,\boldsymbol{\beta}^{(i)}\boldsymbol{\Lambda}_i^{\mathbf{V}}\boldsymbol{\beta}^{(i)\prime})$. Such invariant statistics of course are distribution-free under $\bigcup_{\underline{\sigma}^2}\bigcup_{\mathrm{f}}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^I,\underline{\sigma}^2,\boldsymbol{\vartheta}^{III},\boldsymbol{\vartheta}^{IV};\mathrm{f}}\}$.

The existence of central sequences that are measurable with respect to the multivariate signs $\mathbf{U}_{ij}$ and the ranks R_{ij} (recall that central sequences are always defined up to $o_{\mathrm{P}}(1)$ quantities) is established by the asymptotic representation result of Lemma 7.1(i) below.

Denoting by K an m-tuple of score functions satisfying Assumption (C), consider the random vectors

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K} := ((\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,1}_{\boldsymbol{\vartheta};K_1})',\ldots,(\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,m}_{\boldsymbol{\vartheta};K_m})')',$$

with

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta};K_i} := \frac{1}{2\sqrt{n_i}}\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}}\mathbf{L}_k^{\boldsymbol{\beta}^{(i)},\boldsymbol{\Lambda}_i^{\mathbf{V}}}\left(\mathbf{V}_i^{\otimes 2}\right)^{-1/2}\sum_{j=1}^{n_i}K_i\left(\frac{R_{ij}}{n_i+1}\right)\mathrm{vec}\left(\mathbf{U}_{ij}\mathbf{U}'_{ij}\right). \tag{7.1}$$

In order to describe the asymptotic behavior of $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K}$, similarly define

$$\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta};K;g} := ((\boldsymbol{\Delta}^{IV,1}_{\boldsymbol{\vartheta};K_1;g_1})',\ldots,(\boldsymbol{\Delta}^{IV,m}_{\boldsymbol{\vartheta};K_m;g_m})')',$$

with

$$\boldsymbol{\Delta}^{IV,i}_{\boldsymbol{\vartheta};K_i;g_i} := \frac{1}{2\sqrt{n_i}}\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}}\mathbf{L}_k^{\boldsymbol{\beta}^{(i)},\boldsymbol{\Lambda}_i^{\mathbf{V}}}\left(\mathbf{V}_i^{\otimes 2}\right)^{-1/2}\sum_{j=1}^{n_i}K_i\left(\tilde{G}_{ik}\left(\frac{d_{ij}}{\sigma}\right)\right)\mathrm{vec}\big(\mathbf{U}_{ij}\mathbf{U}'_{ij}\big).$$

We then have the following result.

Lemma 7.1. *Assume that (A), (B), (C), and (D′) hold. Fix any $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ (the eigenvalues are the $\boldsymbol{\vartheta}_0$-reordered eigenvalues) and $\mathrm{g} \in (\mathcal{F}_1)^m$. Then,*

- *(i)* $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta}_0;K} = \boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;K;\mathrm{g}} + o_{L^2}(1)$, *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, *as* $n \to \infty$;
- *(ii) under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, $\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;K;\mathrm{g}}$ *is asymptotically normal with mean zero and covariance matrix* $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K} := \mathrm{diag}(\boldsymbol{\Gamma}^{IV,1}_{\boldsymbol{\vartheta}_0;K_1}, \ldots, \boldsymbol{\Gamma}^{IV,m}_{\boldsymbol{\vartheta}_0;K_m})$, *with* $\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}_0;K_i} := \frac{\mathcal{J}_k(K_i)}{4k(k+2)} \mathbf{G}^{\boldsymbol{\beta}}_k (\boldsymbol{\nu}^{(i)})^{-1} (\mathbf{G}^{\boldsymbol{\beta}}_k)'$;
- *(iii) reinforcing (D′) into (D) and defining* $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} := \mathrm{diag}(\boldsymbol{\Gamma}^{IV,1}_{\boldsymbol{\vartheta}_0;K_1,g_1}, \ldots, \boldsymbol{\Gamma}^{IV,m}_{\boldsymbol{\vartheta}_0;K_m,g_m})$, *with* $\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}_0;K_i,g_i} := \frac{\mathcal{J}_k(K_i,g_i)}{4k(k+2)} \mathbf{G}^{\boldsymbol{\beta}}_k (\boldsymbol{\nu}^{(i)})^{-1} (\mathbf{G}^{\boldsymbol{\beta}}_k)'$, *and assuming moreover that* $\mathrm{g} \in (\mathcal{F}_a)^m$, $\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta}_0;K;\mathrm{g}} - \boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} \boldsymbol{\tau}^{(n)}_{IV}$ *is asymptotically normal with mean zero and covariance matrix* $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K}$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\mathrm{g}}$.

An immediate corollary of the *asymptotic representation result* in Part (i) of this lemma is that $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};\mathrm{f}} := \underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K_\mathrm{f}}$, with $K_\mathrm{f} := (K_{f_1}, \ldots, K_{f_m})$, constitutes a signed-rank version of the eigenvector part $\boldsymbol{\Delta}^{IV}_{\boldsymbol{\vartheta};\mathrm{f}}$ of the f-central sequence; Parts (ii) and (iii) provide the asymptotic distribution of $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K_\mathrm{f}}$, under the null and local alternatives.

In order to construct a test statistic based on $\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta};K}$, we also need to know how it is affected (asymptotically, under the null hypothesis and contiguous alternatives) by the substitution, for $\boldsymbol{\vartheta}$, of an estimator $\boldsymbol{\vartheta}^{(n)}$ satisfying Assumption (E). This important step is taken care of by the *asymptotic linearity* result of Lemma 7.2. This lemma uses the local reordering of eigenvalues described in the previous sections.

Lemma 7.2. *Assume that (A), (B), (C), and (D′) hold, and let* $\boldsymbol{\vartheta}^{(n)}$ *be an estimator satisfying Assumption (E). Fix* $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ *(with common value* $\boldsymbol{\beta}$ *of the eigenvector matrices and the corresponding reordering of eigenvalues). Then, for all* $\mathrm{g} \in (\mathcal{F}_a)^m$,

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\hat{\boldsymbol{\vartheta}};K} - \underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta}_0;K} + \boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} (\boldsymbol{\varsigma}^{(n)}_{IV})^{-1} n^{1/2} \big[\mathbf{1}_m \otimes \mathrm{vec}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\big] = o_{\mathrm{P}}(1)$$

as $n \to \infty$, *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$.

See the appendix for the proof. Finally, the construction of the rank-based tests of Section 7.2 requires consistent estimation of the cross-information quantities $\mathcal{J}_k(K_i, g_i)$, $i = 1, \ldots, m$. The following method, which is inspired by a local maximum likelihood argument, heavily relies on the asymptotic linearity result of Lemma 7.2, and was first proposed, in a different context, by Hallin et al. (2006). Fix $i \in \{1, \ldots, m\}$ and $\mathrm{g} \in (\mathcal{F}_a)^m$, and let $\boldsymbol{\vartheta}^{(n)}$ satisfy Assumption (E). Denote by $\hat{\boldsymbol{\beta}}$ the estimator of the common eigenvector matrix in $\boldsymbol{\vartheta}^{(n)}$; that is, assume that $\boldsymbol{\vartheta}^{(n)} =: (\hat{\boldsymbol{\vartheta}}'_{I}, \hat{\boldsymbol{\vartheta}}'_{II}, \hat{\boldsymbol{\vartheta}}'_{III}, \mathbf{1}'_m \otimes (\mathrm{vec}\,\hat{\boldsymbol{\beta}})')'$. Define, for any $\rho \geq 0$,

$$\mathrm{vec}\,(\hat{\boldsymbol{\beta}}(\rho)) := \mathrm{vec}\,(\hat{\boldsymbol{\beta}}) + n_i^{-1/2} \rho\, k(k+2) \mathbf{G}^{\hat{\boldsymbol{\beta}}}_k \hat{\boldsymbol{\nu}}^{(i)} (\mathbf{G}^{\hat{\boldsymbol{\beta}}}_k)' \underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i}. \tag{7.2}$$

Consider the (almost surely) piecewise continuous quadratic form

$$\rho \mapsto h^{(n)}_i(\rho) := (\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)},K_i})' (\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i})^{-} \underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)}(\rho);K_i},$$

where $\boldsymbol{\vartheta}^{(n)}(\rho)$ is simply obtained from $\boldsymbol{\vartheta}^{(n)}$ by replacing $\hat{\boldsymbol{\beta}}$ with $\hat{\boldsymbol{\beta}}(\rho)$, that is, $\boldsymbol{\vartheta}^{(n)}(\rho) := (\hat{\boldsymbol{\vartheta}}'_{I}, \hat{\boldsymbol{\vartheta}}'_{II}, \hat{\boldsymbol{\vartheta}}'_{III}, \mathbf{1}'_m \otimes (\mathrm{vec}\,\hat{\boldsymbol{\beta}}(\rho))')'$. Then, Lemma 7.2, the consistency of $\boldsymbol{\vartheta}^{(n)}$, and the definition of $\hat{\boldsymbol{\beta}}(\rho)$ in (7.2) imply that

$$h_i^{(n)}(\rho) = (\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i})'(\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i})^{-}\Big[\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i} - \boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i,g_i} n_i^{1/2}\,\mathrm{vec}\,(\hat{\boldsymbol{\beta}}(\rho) - \hat{\boldsymbol{\beta}})\Big] + o_{\mathrm{P}}(1)$$
$$= (\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i})'(\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i})^{-}\Big[\mathbf{I}_{k^2} - \rho k(k+2)\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i,g_i}\mathbf{G}_k^{\hat{\boldsymbol{\beta}}}\hat{\boldsymbol{\nu}}^{(i)}(\mathbf{G}_k^{\hat{\boldsymbol{\beta}}})'\Big]\underset{\sim}{\boldsymbol{\Delta}}{}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i} + o_{\mathrm{P}}(1) \quad (7.3)$$

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$. Now, note that $k(k+2)\mathbf{G}_k^{\boldsymbol{\beta}}\boldsymbol{\nu}^{(i)}(\mathbf{G}_k^{\boldsymbol{\beta}})'$ is the Moore-Penrose generalized inverse of $\frac{1}{4k(k+2)}\mathbf{G}_k^{\boldsymbol{\beta}}(\boldsymbol{\nu}^{(i)})^{-1}(\mathbf{G}_k^{\boldsymbol{\beta}})'$. Hence, recalling that $\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}_0;K_i} := \frac{\mathcal{J}_k(K_i)}{4k(k+2)}\mathbf{G}_k^{\boldsymbol{\beta}}(\boldsymbol{\nu}^{(i)})^{-1}(\mathbf{G}_k^{\boldsymbol{\beta}})'$ and $\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} := \frac{\mathcal{J}_k(K_i,g_i)}{4k(k+2)}\mathbf{G}_k^{\boldsymbol{\beta}}(\boldsymbol{\nu}^{(i)})^{-1}(\mathbf{G}_k^{\boldsymbol{\beta}})'$, (7.3) can be rewritten as

$$h_i^{(n)}(\rho) = \big(1 - \mathcal{J}_k(K_i, g_i)\rho\big)\, h_i^{(n)}(0) + o_{\mathrm{P}}(1), \quad (7.4)$$

still as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$. Since $h_i^{(n)}(0) > 0$, an intuitively appealing estimator for $(\mathcal{J}_k(K_i, g_i))^{-1}$, in view of (7.4), is given by $\hat{\rho} := \inf\{\rho > 0 : h_i^{(n)}(\rho) < 0\}$. By proceeding along the same lines as in Hallin et al. (2006), it is easily shown that $\hat{\mathcal{J}}_k(K_i, g_i) := \hat{\rho}^{-1}$ is, after adequate discretization (which still has no impact in fixed-n_i practice), a consistent estimator of $\mathcal{J}_k(K_i, g_i)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$.

7.2. Optimal rank-based tests.

Motivated by the form of the pseudo-Gaussian statistic in (6.1), consider the signed-rank statistic

$$\underset{\sim}{Q}{}^{(n)}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} := \big(\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta}_0;K}\big)'\,(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}})^{\perp}\,\underset{\sim}{\boldsymbol{\Delta}}{}^{IV}_{\boldsymbol{\vartheta}_0;K}, \quad (7.5)$$

with

$$(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K,\mathrm{g}})^{\perp} := (\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K})^{-} - (\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K})^{-}\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)}_{IV})^{-1}\boldsymbol{\Upsilon}^{IV}$$
$$\times\big[\big(\boldsymbol{\Upsilon}^{IV}\big)'(\boldsymbol{\varsigma}^{(n)}_{IV})^{-1}\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K,\mathrm{g}}(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K})^{-}\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)}_{IV})^{-1}\boldsymbol{\Upsilon}^{IV}\big]^{-}\big(\boldsymbol{\Upsilon}^{IV}\big)'(\boldsymbol{\varsigma}^{(n)}_{IV})^{-1}\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K,\mathrm{g}}(\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K})^{-},$$

where $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K}$ and $\boldsymbol{\Gamma}^{IV}_{\boldsymbol{\vartheta};K,\mathrm{g}}$ are defined in Lemma 7.1 (this includes the $\boldsymbol{\vartheta}_0$-reordering of eigenvalues). Now, by using the facts that $(\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}})'\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}} = 2\mathbf{I}_{k(k-1)/2}$ and $\big(\boldsymbol{\Gamma}^{IV,i}_{\boldsymbol{\vartheta}_0;K}\big)^{-} = \frac{k(k+2)}{\mathcal{J}_k(K_1)}\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}}\boldsymbol{\nu}^{(i)}(\mathbf{G}_k^{\boldsymbol{\beta}^{(i)}})'$, standard algebra yields

$$\underset{\sim}{Q}{}^{(n)}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} = k(k+2)\Big\{\sum_{i=1}^{m}\sum_{1\leq j<j'\leq k}\frac{n_i}{\mathcal{J}_k(K_i)}\,(\boldsymbol{\beta}'_j\,\underset{\sim}{\mathbf{S}}{}_{K;i}\boldsymbol{\beta}_{j'})^2 \quad (7.6)$$
$$- \sum_{i,i'=1}^{m}\sum_{1\leq j<j'\leq k}\frac{n_i n_{i'}}{n}\,\frac{\mathcal{J}_k(K_i,g_i)\mathcal{J}_k(K_{i'},g_{i'})}{\mathcal{J}_k(K_i)\mathcal{J}_k(K_{i'})}\,\frac{\nu_{jj'}(K,\mathrm{g})}{(\nu^{(i)}_{jj'}\nu^{(i')}_{jj'})^{1/2}}\,(\boldsymbol{\beta}'_j\,\underset{\sim}{\mathbf{S}}{}_{K;i}\boldsymbol{\beta}_{j'})\,(\boldsymbol{\beta}'_j\,\underset{\sim}{\mathbf{S}}{}_{K;i'}\boldsymbol{\beta}_{j'})\Big\},$$

where

$$\underset{\sim}{\mathbf{S}}_{K;i} := \frac{1}{n_i}\sum_{j=1}^{n_i} K_i\left(\frac{R_{ij}(\boldsymbol{\vartheta}_0)}{n+1}\right)\mathbf{U}_{ij}(\boldsymbol{\vartheta}_0)\mathbf{U}'_{ij}(\boldsymbol{\vartheta}_0)$$

and

$$\mathrm{diag}(\nu_{12}(K,\mathrm{g}),\ldots,\nu_{(k-1)k}(K,\mathrm{g})) := \left(\sum_{i=1}^{m} r_i^{(n)}\ \frac{\mathcal{J}_k^2(K_i,g_i)}{\mathcal{J}_k(K_i)}\ (\boldsymbol{\nu}^{(i)})^{-1}\right)^{-1} =: \boldsymbol{\nu}(K,\mathrm{g}).$$

As for the pseudo-Gaussian tests of Section 6, obtaining a genuine test statistic requires replacing in (7.6) the parameter value $\boldsymbol{\vartheta}_0$ with an estimator $\boldsymbol{\vartheta}^{(n)}$ satisfying Assumption (E)—here, with $\mathcal{K} = (\mathcal{F}_a)^m$—and replacing the cross-information quantities $\mathcal{J}_k(K_i,g_i)$ with consistent (under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, $\mathrm{g}\in(\mathcal{F}_a)^m$) estimates. The estimates $\hat{\mathcal{J}}_k(K_i,g_i)$ defined at the end of Section 7.1 can be used for that purpose. As for $\boldsymbol{\vartheta}^{(n)}$, many choices are possible. Still avoiding moment assumptions, we propose the following one. Let $\boldsymbol{\theta}_i^{(n)}$ and $\mathbf{V}_i^{(n)}$, $i=1,\ldots,m$ be the location and shape estimators associated with the affine-equivariant multivariate median proposed by Hettmansperger and Randles (2002), which are implicitly defined by

$$\frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{U}_{ij}(\boldsymbol{\theta}_i^{(n)},\mathbf{V}_i^{(n)}) = \mathbf{0} \qquad \text{and} \qquad \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{U}_{ij}(\boldsymbol{\theta}_i^{(n)},\mathbf{V}_i^{(n)})\mathbf{U}'_{ij}(\boldsymbol{\theta}_i^{(n)},\mathbf{V}_i^{(n)}) = \frac{1}{k}\,\mathbf{I}_k,$$

with $\det(\mathbf{V}_i^{(n)})=1$, $i=1,\ldots,m$. Under $\mathcal{H}_0$, the eigenvalue matrices $\boldsymbol{\Lambda}_i^{\mathbf{V}} = \mathrm{diag}(\lambda_{i1}^{\mathbf{V}},\ldots,\lambda_{ik}^{\mathbf{V}})$, $i=1,\ldots,m$ and the matrix $\boldsymbol{\beta}=(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_k)$ of common eigenvectors then can be estimated consistently by using the *plug-in* method as in Boente and Orellana (2001). More precisely, the resulting estimates $\hat{\boldsymbol{\Lambda}}_i^{\mathbf{V}}$, $i=1,\ldots,m$ and $\hat{\boldsymbol{\beta}}$ are obtained by solving the ML-type equations

$$\boldsymbol{\beta}'_j\left(\sum_{i=1}^{m} n_i\frac{\lambda_{ij}^{\mathbf{V}}-\lambda_{il}^{\mathbf{V}}}{\lambda_{ij}^{\mathbf{V}}\lambda_{il}^{\mathbf{V}}}\mathbf{V}_i^{(n)}\right)\boldsymbol{\beta}_l = 0, \quad j\neq l=1,\ldots,k, \tag{7.7}$$

$$\boldsymbol{\beta}'_j\mathbf{V}_i^{(n)}\boldsymbol{\beta}_j = \lambda_{ij}^{\mathbf{V}}, \quad i=1,\ldots,m,\ j=1,\ldots,k, \qquad \boldsymbol{\beta}'_j\boldsymbol{\beta}_l = \delta_{jl}, \quad j,l=1,\ldots,k,$$

where δ_{jl} is the usual Kronecker symbol. As in the pseudo-Gaussian context, the resulting estimators are root-n consistent up to a global permutation (see the comments below (6.3)). Now, the scale parameters σ_i^2, $i=1,\ldots,m$ do not appear in (7.6), so that the resulting "estimators" for $\boldsymbol{\vartheta}$ can be chosen as if they were specified:

$$\boldsymbol{\vartheta}^{(n)} := \left(\hat{\boldsymbol{\theta}}'_1,\ldots,\hat{\boldsymbol{\theta}}'_m,\sigma_1^2,\ldots,\sigma_m^2,(\mathrm{d}\overset{\circ}{\mathrm{vec}}\,(\hat{\boldsymbol{\Lambda}}_1^{\mathbf{V}}))',\ldots,(\mathrm{d}\overset{\circ}{\mathrm{vec}}\,(\hat{\boldsymbol{\Lambda}}_m^{\mathbf{V}}))',\mathbf{1}'_m\otimes(\mathrm{vec}\,\hat{\boldsymbol{\beta}})'\right)'. \tag{7.8}$$

It can be checked that, after appropriate discretization, $\boldsymbol{\vartheta}^{(n)}$ in (7.8) satisfies Assumption (E) with $\mathcal{K}=(\mathcal{F}_1)^m$ (hence without requiring any moment condition), so that it can be used advantageously in our signed-rank tests. Summing up, the signed-rank statistic we propose is

$$\underset{\sim}{Q}^{(n)}_{K} := k(k+2)\Bigg\{ \sum_{i=1}^{m} \sum_{1\leq j<j'\leq k} \frac{n_i}{\mathcal{J}(K_i)} (\hat{\boldsymbol{\beta}}'_j \,\underset{\sim}{\hat{\mathbf{S}}}_{K;i} \hat{\boldsymbol{\beta}}_{j'})^2$$

$$- \sum_{i,i'=1}^{m} \sum_{1\leq j<j'\leq k} \frac{n_i n_{i'}}{n} \frac{\hat{\mathcal{J}}_k(K_i,g_i)\hat{\mathcal{J}}_k(K_{i'},g_{i'})}{\mathcal{J}(K_i)\mathcal{J}(K_{i'})} \frac{\hat{\nu}_{jj'}(K,\mathrm{g})}{(\hat{\nu}^{(i)}_{jj'}\hat{\nu}^{(i')}_{jj'})^{1/2}} (\hat{\boldsymbol{\beta}}'_j \,\underset{\sim}{\hat{\mathbf{S}}}_{K;i} \hat{\boldsymbol{\beta}}_{j'}) (\hat{\boldsymbol{\beta}}'_j \,\underset{\sim}{\hat{\mathbf{S}}}_{K;i'} \hat{\boldsymbol{\beta}}_{j'}) \Bigg\},$$

where

$$\underset{\sim}{\hat{\mathbf{S}}}_{K;i} := \frac{1}{n_i} \sum_{j=1}^{n_i} K_i \left(\frac{R_{ij}(\boldsymbol{\vartheta}^{(n)})}{n+1} \right) \mathbf{U}_{ij}(\boldsymbol{\vartheta}^{(n)}) \mathbf{U}'_{ij}(\boldsymbol{\vartheta}^{(n)})$$

and

$$\mathrm{diag}(\hat{\nu}_{12}(K,\mathrm{g}),\ldots,\hat{\nu}_{(k-1)k}(K,\mathrm{g})) := \left(\sum_{i=1}^{m} r_i^{(n)} \frac{\hat{\mathcal{J}}^2_k(K_i,g_i)}{\mathcal{J}(K_i)} (\hat{\boldsymbol{\nu}}^{(i)})^{-1} \right)^{-1} =: \hat{\boldsymbol{\nu}}(K,\mathrm{g}),$$

all parameters being estimated via the chosen estimator $\boldsymbol{\vartheta}^{(n)}$(given in (7.8), for instance). The resulting test $\underset{\sim}{\phi}^{(n)}_{K}$ rejects the null of CPC, at asymptotic level α, as soon as $\underset{\sim}{Q}^{(n)}_{K}$ exceeds the α-upper quantile of the chi-square distribution with $(m-1)k(k-1)/2$ degrees of freedom. Consider perturbations $\boldsymbol{\tau}^{(n)}$ as described in Proposition 6.1 above. Letting

$$\mathbf{C}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} := \mathrm{diag}\bigg(\frac{\mathcal{J}^2_k(K_1,g_1)}{\mathcal{J}_k(K_1)} (\boldsymbol{\nu}^{(1)})^{-1}, \ldots, \frac{\mathcal{J}^2_k(K_m,g_m)}{\mathcal{J}_k(K_m)} (\boldsymbol{\nu}^{(m)})^{-1} \bigg)$$

and

$$\mathbf{D}^{(n)}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} := \mathbf{C}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} - \mathbf{C}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} \big[((\mathbf{r}^{(n)})^{-1} \mathbf{1}_m \mathbf{1}'_m (\mathbf{r}^{(n)})^{-1}) \otimes \boldsymbol{\nu}(K,\mathrm{g}) \big] \mathbf{C}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}},$$

the quantities characterizing the asymptotic distribution of $\underset{\sim}{Q}^{(n)}_{K}$ under the corresponding local alternatives (see Part (i) of Theorem 7.1 below) are

$$\begin{aligned} l_{\boldsymbol{\vartheta}_0,\boldsymbol{\tau};K,\mathrm{g}} &:= \lim_{n\to\infty} \Big\{ (\boldsymbol{\tau}^{(n)}_{\mathit{IV}})' (\mathbf{I}_m \otimes \mathbf{G}^{\boldsymbol{\beta}}_k) \, \mathbf{D}^{(n)}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} (\mathbf{I}_m \otimes \mathbf{G}^{\boldsymbol{\beta}}_k)' (\boldsymbol{\tau}^{(n)}_{\mathit{IV}}) \Big\} \qquad (7.9) \\ &= \sum_{i,i'=1}^{m} (\mathrm{vec}\,\mathbf{b}^{(i)})' \mathbf{G}^{\boldsymbol{\beta}}_k \Big[\delta_{ii'} \mathbf{T}^{(i,i')}_{K,\mathrm{g}} - (r_i\, r_{i'})^{1/2}\, \mathbf{T}^{(i,i')}_{K,\mathrm{g}}\, \boldsymbol{\nu}(K,\mathrm{g}) \mathbf{T}^{(i,i')}_{K,\mathrm{g}} \Big] \big(\mathbf{G}^{\boldsymbol{\beta}}_k\big)' (\mathrm{vec}\,\mathbf{b}^{(i')}), \end{aligned}$$

where

$$\mathbf{T}^{(i,i')}_{K,\mathrm{g}} := \frac{\mathcal{J}_k(K_i,g_i)\mathcal{J}_k(K_{i'},g_{i'})}{\big(\mathcal{J}(K_i)\,\mathcal{J}(K_{i'})\big)^{1/2}} (\boldsymbol{\nu}^{(i)})^{-1/2} (\boldsymbol{\nu}^{(i')})^{-1/2}.$$

We are now ready to state the main result of this paper.

Proposition 7.1. *Assume that (A), (B), (C), and (D′) hold, and let* $\boldsymbol{\vartheta}^{(n)}$ *be an estimator satisfying Assumption (E) with* $\mathcal{K} = (\mathcal{F}_a)^m$*. Then,*

(i) $\underset{\sim}{Q}_K^{(n)}$ *is asymptotically chi-square with* $(m-1)k(k-1)/2$ *degrees of freedom under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0}\bigcup_{\mathrm{g}\in(\mathcal{F}_a)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}\}$*, and (provided that (B) and (D′) are reinforced into (B′) and (D), respectively) asymptotically noncentral chi-square, still with* $(m-1)k(k-1)/2$ *degrees of freedom, but with noncentrality parameter* $l_{\boldsymbol{\vartheta},\boldsymbol{\tau};K,\mathrm{g}}/k(k+2)$ *under* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\mathrm{g}}$, $\boldsymbol{\vartheta}\in\mathcal{H}_0$, $\boldsymbol{\varsigma\tau} := \lim_{n\to\infty}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$ *as in Proposition 6.1 and* $\mathrm{g}\in(\mathcal{F}_a)^m$*;*

(ii) $\underset{\sim}{\phi}_K^{(n)}$ *has asymptotic level* α *under* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0}\bigcup_{\mathrm{g}\in(\mathcal{F}_a)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}\}$*;*

(iii) *reinforcing (D′) into (D),* $\underset{\sim}{\phi}_{K_{\mathrm{f}}}^{(n)}$, $K_{\mathrm{f}} := (K_{f_1},\ldots,K_{f_m})$*, is locally and asymptotically most stringent, at asymptotic level* α*, for* $\bigcup_{\boldsymbol{\vartheta}\in\mathcal{H}_0}\bigcup_{\mathrm{g}\in(\mathcal{F}_a)^m}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}\}$ *against alternatives of the form* $\bigcup_{\boldsymbol{\vartheta}\notin\mathcal{H}_0}\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{f}}\}$ *with* $\mathrm{f} := (f_1,\ldots,f_m)$*.*

The signed-rank test $\underset{\sim}{\phi}_K^{(n)}$ is *asymptotically invariant* with respect to continuous monotone radial transformatons in the sense that it is asymptotically equivalent to a random variable which is invariant under such transformations.

8. Power comparison and simulations.

8.1. Asymptotic relative efficiencies.

The asymptotic relative efficiencies (AREs) of the signed-rank test $\underset{\sim}{\phi}_K^{(n)}$ with respect to the pseudo-Gaussian test $\phi_{\mathcal{N}}^{(n)\dagger}$ (equivalently, with respect to $\phi_{\mathrm{HPV}}^{(n)}$) directly follow as ratios of noncentrality parameters under local alternatives (see Propositions 6.1 and 7.1).

Proposition 8.1. *Assume that (A), (B′), (C), and (D) hold. Then, the asymptotic relative efficiency of* $\underset{\sim}{\phi}_K^{(n)}$ *with respect to* $\phi_{\mathcal{N}}^{(n)\dagger}$*, when testing* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};\mathrm{g}}$ *against* $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)};\mathrm{g}}$*, with* $\boldsymbol{\vartheta}\in\mathcal{H}_0$, $\boldsymbol{\varsigma\tau} := \lim_{n\to\infty}\boldsymbol{\varsigma}^{(n)}\boldsymbol{\tau}^{(n)}$ *as described in Proposition 6.1, and* $\mathrm{g}\in(\mathcal{F}_a^4)^m$*, is*

$$\mathrm{ARE}_{k,\mathrm{g}}(\underset{\sim}{\phi}_K^{(n)}/\phi_{\mathcal{N}}^{(n)\dagger}) = l_{\boldsymbol{\vartheta},\boldsymbol{\tau};K,\mathrm{g}}/k(k+2)l^{\mathcal{N}}_{\boldsymbol{\vartheta},\boldsymbol{\tau};\mathrm{g}}, \tag{8.1}$$

where $l^{\mathcal{N}}_{\boldsymbol{\vartheta},\boldsymbol{\tau};\mathrm{g}}$ *and* $l_{\boldsymbol{\vartheta},\boldsymbol{\tau};K,\mathrm{g}}$ *are defined in (6.5) and (7.9), respectively.*

Note that, if $\mathrm{g} = (g_1,\ldots,g_1)$ (homogeneous elliptical densities) and if the same score function—namely, K_1—is used for the m rankings, (8.1) simplifies into

$$\mathrm{ARE}_{k,\mathrm{g}}(\underset{\sim}{\phi}_{K_1}^{(n)}/\phi_{\mathcal{N}}^{(n)\dagger}) = (1+\kappa_k(g_1))\mathcal{J}_k^2(K_1,g_1)/k(k+2)\mathcal{J}_k(K_1); \tag{8.2}$$

these are the AREs obtained in one-sample shape problems (see Hallin and Paindaveine (2006) and Hallin et al. (2006), in hypothesis testing and point estimation contexts, respectively). The Chernoff-Savage property of Paindaveine (2006) therefore holds: denoting by $\underset{\sim}{\phi}{}^{(n)}_{\text{vdW}}$ the van der Waerden rank test (based on the Gaussian scores $K_1 = \ldots = K_m := \Psi_k^{-1}$, where Ψ_k^{-1} stands for the quantile function of the chi-square distribution with k degrees of freedom), we have that

$$\text{ARE}_{k,\text{g}}(\underset{\sim}{\phi}{}^{(n)}_{\text{vdW}}/\phi^{(n)\dagger}_{\mathcal{N}}) \geq 1$$

for all homogeneous $\text{g} \in (\mathcal{F}_a^4)^m$, with equality in the Gaussian case only.

In the bivariate two-population case ($m = k = 2$) with $\mathbf{b}^{(1)} = \mathbf{0}$ (no perturbation on the eigenvectors of the first population), the ARE under $\text{g} = (g_1, g_2)$ of $\phi^{(n)\dagger}_{\mathcal{N}}$ with respect to the optimal parametric test $\phi^{(n)}_{\text{g}}$(recall that, under (B$'$), $r_i := \lim_{n\to\infty} n_i/n$) is

$$\text{ARE}_{2,\text{g}}(\phi^{(n)\dagger}_{\mathcal{N}}/\phi^{(n)}_{\text{g}}) = \frac{k(k+2)(1+\kappa_2(g_2))^{-1}\left(1 - r_2(\nu^{(2)}_{12})^{-1}(1+\kappa_2(g_2))^{-1}\nu_{12}(\text{g})\right)}{\mathcal{J}_2(g_2)\left(1 - r_2(\nu^{(2)}_{12})^{-1}\mathcal{J}_2(g_2)\nu_{12}(K_{\text{g}},\text{g})\right)}, \tag{8.3}$$

where denoting by K_{g_1} and K_{g_2} the score functions associated with g_1 and g_2 respectively, $K_{\text{g}} = (K_{g_1}, K_{g_2})$ and $\nu_{12}(K_{\text{g}}, \text{g})$ naturally stands for the $\nu_{12}(K, \text{g})$ quantity computed from K_{g} scores. Under the same setting, the ARE of the van der Waerden test $\underset{\sim}{\phi}{}^{(n)}_{\text{vdW}}$(with score $K_\phi := \Psi_k^{-1}$) with respect to $\phi^{(n)}_{\text{g}}$ takes the form

$$\text{ARE}_{2,\text{g}}(\underset{\sim}{\phi}{}^{(n)}_{\text{vdW}}/\phi^{(n)}_{\text{g}}) = \frac{\frac{\mathcal{J}^2(K_\phi,g_2)}{\mathcal{J}(K_\phi)}\left(1 - r_2(\nu^{(2)}_{12})^{-1}\frac{\mathcal{J}^2(K_\phi,g_2)}{\mathcal{J}(K_\phi)}\nu_{12}(K_\phi,\text{g})\right)}{\mathcal{J}_2(g_2)\left(1 - r_2(\nu^{(2)}_{12})^{-1}\mathcal{J}_2(g_2)\nu_{12}(K_{\text{g}},\text{g})\right)}. \tag{8.4}$$

These AREs do not depend on the value $\boldsymbol{\beta}$ of the common eigenvectors under the null, nor on the perturbation $\mathbf{b}^{(2)}$. Tables 1 and 2 provide numerical values of (8.3) and (8.4), respectively, with $r_2 = 1 - r_1 = 120/220$ (the sampling scheme considered in the simulations of Section 8.2), for various choices of bivariate Student t_ν and Gaussian population densities $\text{g} = (g_1, g_2)$. Note that the ARE of the pseudo-Gaussian tests with respect to van der Waerden ones can be as low as .13 under homokurtic bivariate $t_{4.2}$ populations, which demonstrates the severe lack of efficiency robustness of the pseudo-Gaussian tests.

8.2. Monte-Carlo study

In this section, we concentrate on comparing Flury's traditional Gaussian LRT ($\phi^{(n)}_{\text{Flury}}$) for the null hypothesis of CPC with the pseudo-Gaussian test $\phi^{(n)\dagger}_{\mathcal{N}}$ of Section 6 and the signed-rank tests of Section 7. First, we generated $N = 1,000$ independent replications of three pairs ($m = 2$) of mutually independent samples (with respective sizes $n_1 = 100$ and $n_2 = 120$) of bivariate ($k = 2$) random vectors

$$\boldsymbol{\varepsilon}_{\ell;1j_1} \quad \text{and} \quad \boldsymbol{\varepsilon}_{\ell;2j_2}, \quad \ell = 1,2,3,4, \quad j_i = 1,\ldots,n_i, \quad i = 1,2,$$

g_1/g_2	$t_{4.2}$	t_5	t_6	t_8	t_{12}	$\mathcal{N}$
$t_{4.2}$	.1202	.1773	.1867	.1896	.1889	.1822
t_5	.1963	.4286	.4987	.5378	.5538	.5528
t_6	.2106	.5159	.6250	.6923	.7241	.7353
t_8	.2158	.5679	.7079	.8000	.8468	.8714
t_{12}	.2155	.5905	.7490	.8571	.9143	.9486
$\mathcal{N}$	.2068	.5918	.7656	.8889	.9563	1.000

Table 1. Asymptotic relative efficiencies (8.3) of the pseudo-Gaussian tests with respect to the optimal parametric (or the optimal rank-based) ones under various bivariate Student t_ν and Gaussian population densities $\mathrm{g} = (g_1, g_2)$, with $r_2 = 1 - r_1 = 120/220$.

g_1/g_2	$t_{4.2}$	t_5	t_6	t_8	t_{12}	$\mathcal{N}$
$t_{4.2}$	.9303	.9367	.9419	.9478	.9526	.9561
t_5	.9380	.9446	.9501	.9564	.9616	.9656
t_6	.9443	.9513	.9570	.9636	.9691	.9738
t_8	.9516	.9589	.9650	.9720	.9779	.9833
t_{12}	.9576	.9652	.9717	.9791	.9854	.9915
$\mathcal{N}$	.9622	.9706	.9776	.9858	.9928	1.000

Table 2. Asymptotic relative efficiencies (8.3) of the van der Waerden tests with respect to the optimal parametric (or the optimal rank-based) ones under various bivariate Student t_ν and Gaussian population densities $\mathrm{g} = (g_1, g_2)$, with $r_2 = 1 - r_1 = 120/220$.

with bivariate standard Gaussian densities ($\boldsymbol{\varepsilon}_{1;1j_1}$ and $\boldsymbol{\varepsilon}_{1;2j_2}$: Gaussian case), bivariate Gaussian ($\boldsymbol{\varepsilon}_{2;1j_1}$) and t_5 ($\boldsymbol{\varepsilon}_{2;2j_2}$) (non-Gaussian heterokurtic case with finite fourth-order moments), bivariate standard t_1 densities ($\boldsymbol{\varepsilon}_{3;1j_1}$ and $\boldsymbol{\varepsilon}_{3;2j_2}$: non-Gaussian homokurtic case with infinite fourth-order moments) and bivariate standard t_5 ($\boldsymbol{\varepsilon}_{4;1j_1}$) and t_1 ($\boldsymbol{\varepsilon}_{4;2j_2}$) (non-Gaussian heterokurtic case with infinite fourth-order moments), respectively. Each replication of the $\boldsymbol{\varepsilon}_{\ell;1j_1}$'s was transformed into

$$\mathbf{X}_{\ell;1j_1} = \boldsymbol{\beta}\boldsymbol{\Lambda}_1^{1/2}\boldsymbol{\varepsilon}_{\ell;1j_1}, \quad \ell = 1,2,3,4, \quad j_1 = 1,\ldots,n_1, \tag{8.5}$$

where

$$\boldsymbol{\beta} = \begin{pmatrix} \cos(\pi/6) & -\sin(\pi/6) \\ \sin(\pi/6) & \cos(\pi/6) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Lambda}_1 = \begin{pmatrix} 16 & 0 \\ 0 & 8 \end{pmatrix},$$

while each replication of the $\boldsymbol{\varepsilon}_{\ell;2j_2}$'s was transformed into

$$\mathbf{X}_{\ell;2j_2;\xi} = \boldsymbol{\beta}\mathbf{B}_\xi\boldsymbol{\Lambda}_2^{1/2}\boldsymbol{\varepsilon}_{\ell;2j_2}, \quad \ell = 1,2,3,4, \quad j_2 = 1,\ldots,n_2, \quad \xi = 0,1,2,3 \tag{8.6}$$

where

$$\mathbf{B}_\xi = \begin{pmatrix} \cos(\pi\xi/15) & -\sin(\pi\xi/15) \\ \sin(\pi\xi/15) & \cos(\pi\xi/15) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Lambda}_2 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

Clearly, the scatter matrices of $\mathbf{X}_{\ell;1j_1}$ and $\mathbf{X}_{\ell;2j_2;0}$ have common eigenvectors $\boldsymbol{\beta}$, with distinct eigenvalue matrices $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$, while the eigenvectors of $\mathbf{X}_{\ell;2j_2;\xi}$, $\xi = 1,2,3$

		ξ				
underlying densities g	test	0	1	2	3	$\text{ARE}_{k,\text{g}}(\cdot/\underset{\sim}{\phi}_{\text{opt}})$
$\mathcal{N}$, $\mathcal{N}$	ϕ_{Flury}	.039	.178	.504	.780	1.0000
	$\phi^{\dagger}_{\mathcal{N}}(=\phi_{\text{HPV}})$	.041	.177	.494	.722	1.0000
	$\underset{\sim}{\phi}_{\text{vdW}}$	.031	.148	.412	.633	1.0000
	$\underset{\sim}{\phi}_{t_5,t_5}$	.035	.147	.426	.626	.9446
	$\underset{\sim}{\phi}_{t_1,t_1}$	.036	.110	.350	.517	.7407
	$\underset{\sim}{\phi}_{t_1,t_5}$	.043	.146	.414	.595	.8213
	$\underset{\sim}{\phi}_{\text{vdW},t_5}$	.039	.149	.433	.631	.9740
t_1, t_1	ϕ_{Flury}	.698	.705	.704	.716	.0000
	$\phi^{\dagger}_{\mathcal{N}}(=\phi_{\text{HPV}})$	.025	.045	.037	.037	.0000
	$\underset{\sim}{\phi}_{\text{vdW}}$	.036	.077	.198	.335	.7407
	$\underset{\sim}{\phi}_{t_5,t_5}$	.041	.088	.261	.416	.8972
	$\underset{\sim}{\phi}_{t_1,t_1}$	.035	.123	.295	.460	1.0000
	$\underset{\sim}{\phi}_{t_1,t_5}$	.043	.111	.282	.436	.9505
	$\underset{\sim}{\phi}_{\text{vdW},t_5}$	.035	.081	.235	.369	.8045
t_1, t_5	ϕ_{Flury}	.478	.517	.543	.519	.0000
	$\phi^{\dagger}_{\mathcal{N}}(=\phi_{\text{HPV}})$	.230	.274	.277	.282	.0000
	$\underset{\sim}{\phi}_{\text{vdW}}$	.034	.093	.278	.414	.8091
	$\underset{\sim}{\phi}_{t_5,t_5}$	.041	.116	.315	.481	.9348
	$\underset{\sim}{\phi}_{t_1,t_1}$	.051	.115	.309	.487	.9571
	$\underset{\sim}{\phi}_{t_1,t_5}$	.059	.141	.345	.528	1.0000
	$\underset{\sim}{\phi}_{\text{vdW},t_5}$	.035	.103	.291	.439	.8243
$\mathcal{N}$, t_5	ϕ_{Flury}	.124	.236	.480	.707	.0000
	$\phi^{\dagger}_{\mathcal{N}}(=\phi_{\text{HPV}})$	.068	.156	.374	.536	.5918
	$\underset{\sim}{\phi}_{\text{vdW}}$	.040	.139	.377	.566	.9706
	$\underset{\sim}{\phi}_{t_5,t_5}$	.049	.146	.400	.600	.9725
	$\underset{\sim}{\phi}_{t_1,t_1}$	.049	.141	.353	.514	.8142
	$\underset{\sim}{\phi}_{t_1,t_5}$	.067	.156	.365	.570	.8556
	$\underset{\sim}{\phi}_{\text{vdW},t_5}$	.039	.144	.401	.595	1.0000

Table 3. Rejection frequencies (out of $N = 1,000$ replications), under the null ($\xi = 0$) and three alternatives ($\xi = 1,\ 2,\ 3$; see Section 8.2 for details), of the Flury test (ϕ_{Flury}), the pseudo-Gaussian tests $\phi^{\dagger}_{\mathcal{N}}(=\phi_{\text{HPV}})$, the signed-rank van der Waerden ($\underset{\sim}{\phi}_{\text{vdW}}$—Gaussian scores in both samples), homogeneous t_ν-score ($\underset{\sim}{\phi}_{t_5,t_5}$ and $\underset{\sim}{\phi}_{t_1,t_1}$—identical Student scores in both samples), heterogeneous t_ν-score test ($\underset{\sim}{\phi}_{t_1,t_5}$—$t_1$-scores in sample one and t_5-scores in sample two) and heterogeneous Gaussian and t_5-scores ($\underset{\sim}{\phi}_{\text{vdW},t_5}$—Gaussian scores in sample one, t_5-scores in sample two). Sample sizes are $n_1 = 100$ and $n_2 = 120$. In the last column, we give the AREs with respect to the optimal (for the densities under reference) rank-based test.

differ from those of $\mathbf{X}_{\ell;1j_1}$, thus characterizing increasingly distant alternatives to the null hypothesis of CPC.

Rejection frequencies (based on the asymptotic chi-square critical values, at nominal 5% level) are reported in Table 3, the inspection of which reveals several well expected facts:

(i) $\phi^{(n)}_{\text{Flury}}$ and $\phi^{(n)\dagger}_{\mathcal{N}}$ yield similar behaviors under Gaussian densities, but completely blow up under densities with infinite fourth-order moments. It is however shown in Hallin et al. (2010a) that $\phi^{(n)}_{\text{HPV}} = \phi^{(n)\dagger}_{\mathcal{N}}$ remains valid under heterokurtic elliptical densities with finite fourth-order moments;

(ii) the signed-rank tests, unlike their Gaussian and pseudo-Gaussian competitors, keep the right nominal size under the null in all designs considered. They furthermore exhibit quite good results in terms of efficiency;

(iii) despite the relatively small sample sizes $n_1 = 100$ and $n_2 = 120$, empirical powers and ARE rankings almost perfectly agree.

9. Appendix.

Proofs of Lemma 7.1 and Lemma 7.2. Part(i) of Lemma 7.1 readily follows from classical asymptotic representation results for signed-rank-based statistics: see, for instance, Lemma 4.1 in Hallin and Paindaveine (2010). Parts (ii) and (iii) are direct consequences of Part (i), the multivariate central limit theorem, and ULAN.

We therefore concentrate on the proof of Lemma 7.2. For $i = 1, \ldots, n$, let $\hat{\mathbf{V}}_i := \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\Lambda}}_i^{\mathbf{V}}\hat{\boldsymbol{\beta}}'$ denote the root-n_i consistent estimator (under $\mathcal{H}_0$) of the shape matrix $\mathbf{V}_i$ resulting from the estimated eigenvalues $\hat{\boldsymbol{\Lambda}}_i^{\mathbf{V}}$ and estimated (common) eigenvectors $\hat{\boldsymbol{\beta}}$. With that estimated shape matrix, we get

$$\underset{\sim}{\boldsymbol{\Delta}}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i} = \frac{\sqrt{n_i}}{2}\mathbf{G}_k^{\hat{\boldsymbol{\beta}}}\mathbf{L}_k^{\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\Lambda}}_i^{\mathbf{V}}}(\hat{\mathbf{V}}_i^{\otimes 2})^{-1/2}\text{vec}(\underset{\sim}{\hat{\mathbf{S}}}_{K;i}),$$

where $\underset{\sim}{\hat{\mathbf{S}}}_{K;i} := \frac{1}{n_i}\sum_{j=1}^{n_i} K_i\left(\frac{R_{ij}(\boldsymbol{\vartheta}^{(n)})}{n+1}\right)\mathbf{U}_{ij}(\boldsymbol{\vartheta}^{(n)})\mathbf{U}'_{ij}(\boldsymbol{\vartheta}^{(n)})$. Similarly define $\underset{\sim}{\mathbf{S}}_{K;i} := \frac{1}{n_i}\sum_{j=1}^{n_i} K_i\left(\frac{R_{ij}(\boldsymbol{\vartheta})}{n+1}\right)\mathbf{U}_{ij}(\boldsymbol{\vartheta})\mathbf{U}'_{ij}(\boldsymbol{\vartheta})$. Letting $\mathbf{J}_k^{\perp} := \mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k$, note that, since $n_i^{1/2}\mathbf{J}_k^{\perp}\underset{\sim}{\hat{\mathbf{S}}}_{K;i}$ is $O_{\text{P}}(1)$ as $n \to \infty$ under $\text{P}^{(n)}_{\boldsymbol{\vartheta}_0;\text{g}}$ and $\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_i^{\mathbf{V}}}(\mathbf{V}_i{}^{-1/2})^{\otimes 2}\mathbf{J}_k = \mathbf{0}$, Slutzky's Lemma entails

$$\underset{\sim}{\boldsymbol{\Delta}}^{IV,i}_{\boldsymbol{\vartheta}^{(n)};K_i} := \frac{\sqrt{n_i}}{2}\mathbf{G}_k^{\boldsymbol{\beta}}\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_i^{\mathbf{V}}}\left(\mathbf{V}_i^{\otimes 2}\right)^{-1/2}\text{vec}(\underset{\sim}{\hat{\mathbf{S}}}_{K;i}) + o_{\text{P}}(1) \tag{9.1}$$

as $n \to \infty$ under $\text{P}^{(n)}_{\boldsymbol{\vartheta}_0;\text{g}}$. From Lemma A1 in Hallin, Oja and Paindaveine (2006) and Lemma 4.4 in Kreiss (1987), we have that, for $\boldsymbol{\vartheta}^{(n)}$ satisfying (E),

$$\mathbf{J}_k^{\perp}\sqrt{n_i}\,\text{vec}\,(\underset{\sim}{\hat{\mathbf{S}}}_{K;i} - \underset{\sim}{\mathbf{S}}_{K;i}) + \frac{\mathcal{J}_k(K,g_1)}{4k(k+2)}\Big[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\Big](\mathbf{V}_i{}^{-1/2})^{\otimes 2}n_i^{1/2}\text{vec}\,(\hat{\mathbf{V}}_i - \mathbf{V}_i) = o_{\text{P}}(1) \tag{9.2}$$

as $n \to \infty$, still under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$. It directly follows from (9.1), (9.2) and the fact that $\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_i^{\mathbf{V}}}(\mathbf{V}_i{}^{-1/2})^{\otimes 2}\mathbf{J}_k = \mathbf{0}$ that

$$\underset{\sim}{\boldsymbol{\Delta}}{}^{\mathit{IV},i}_{\boldsymbol{\vartheta}^{(n)};K_i} - \underset{\sim}{\boldsymbol{\Delta}}{}^{\mathit{IV},i}_{\boldsymbol{\vartheta};K_i} = \frac{\mathcal{J}_k(K,g_1)}{4k(k+2)}\mathbf{G}_k^{\boldsymbol{\beta}}\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_i^{\mathbf{V}}}\left(\mathbf{V}_i^{\otimes 2}\right)^{-1}\Big[\mathbf{I}_{k^2}+\mathbf{K}_k\Big]n_i^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}}_i - \mathbf{V}_i) + o_{\mathrm{P}}(1) \quad (9.3)$$

Next, following the same argument as in the proof of Lemma 4.2 in Hallin, Paindaveine, Verdebout (2010b), we have that

$$n_i^{1/2}\mathrm{vec}\,(\hat{\mathbf{V}}_i - \mathbf{V}_i) = (\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_i^{\mathbf{V}}})'(\mathbf{G}_k^{\boldsymbol{\beta}})'n_i^{1/2}\mathrm{vec}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + \boldsymbol{\beta}^{\otimes 2}\mathbf{H}_k' n_i^{1/2}\mathrm{dvec}(\hat{\boldsymbol{\Lambda}}_i^{\mathbf{V}} - \boldsymbol{\Lambda}_i^{\mathbf{V}}) + o_{\mathrm{P}}(1) \quad (9.4)$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$. The result follows by plugging (9.4) into (9.3), then using the fact that $(\mathbf{L}_k^{\boldsymbol{\beta},\boldsymbol{\Lambda}_i^{\mathbf{V}}})'\left(\mathbf{V}_i^{\otimes 2}\right)^{-1}\Big[\mathbf{I}_{k^2}+\mathbf{K}_k\Big]\boldsymbol{\beta}^{\otimes 2}\mathbf{H}_k' = \mathbf{0}$.

□

Proof of Proposition 7.1 Simple algebra yields, for $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$,

$$\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)})^{-1}\boldsymbol{\Upsilon}^{\mathit{IV}} = 2\,\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)})^{-1}\tilde{\boldsymbol{\Upsilon}},$$

with $\tilde{\boldsymbol{\Upsilon}} := \mathbf{1}_m \otimes \mathbf{I}_{k^2}$. This implies that

$$\underset{\sim}{Q}{}^{(n)}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}} = (\underset{\sim}{\boldsymbol{\Delta}}{}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K})'(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}})^{\perp}\underset{\sim}{\boldsymbol{\Delta}}{}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K},$$

where

$$(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}})^{\perp} = (\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K})^{-} - (\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K})^{-}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)})^{-1}\tilde{\boldsymbol{\Upsilon}} \Big(\tilde{\boldsymbol{\Upsilon}}'(\boldsymbol{\varsigma}^{(n)})^{-1}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K})^{-}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)})^{-1}\tilde{\boldsymbol{\Upsilon}}\Big)^{-}\tilde{\boldsymbol{\Upsilon}}'(\boldsymbol{\varsigma}^{(n)})^{-1}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K})^{-}. \quad (9.5)$$

Using Slutzky's Lemma jointly with Lemma 7.2, we obtain that $\underset{\sim}{Q}{}^{(n)}_{K} - \underset{\sim}{Q}{}^{(n)}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}}$ is $o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\vartheta}_0,\mathrm{g}}$ for $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ and $\mathrm{g} \in (\mathcal{F}_a)^m$ iff (denoting by $\boldsymbol{\beta}$ the common eigenvector matrix under $\boldsymbol{\vartheta}_0$)

$$(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}})^{\perp}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0,K,\mathrm{g}}(\boldsymbol{\varsigma}^{(n)})^{-1}\tilde{\boldsymbol{\Upsilon}}n^{1/2}\mathrm{vec}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \mathbf{0}.$$

In view of (9.5), however, this follows trivially from the fact that $\mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{A} = \mathbf{A}$, a standard properties of Moore-Penrose inverses.

Now, since $\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K}(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}})^{\perp}$ is idempotent with trace $(m-1)k(k-1)/2$, it follows from Theorem 9.2.1 in Rao and Mitra (1971) that $\underset{\sim}{Q}{}^{(n)}_{K}$ is asymptotically chi-square with $(m-1)k(k-1)/2$ degrees of freedom under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\mathrm{g}}$, $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$, and asymptotically noncentral chi-square, still with $(m-1)k(k-1)/2$ degrees of freedom, but with noncentrality parameter

$$\lim_{n\to\infty}\left\{(\boldsymbol{\tau}^{\mathit{IV}})^{(n)\prime}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}}(\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}})^{\perp}\boldsymbol{\Gamma}^{\mathit{IV}}_{\boldsymbol{\vartheta}_0;K,\mathrm{g}}(\boldsymbol{\tau}^{\mathit{IV}})^{(n)}\right\} \quad (9.6)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)};\mathrm{g}}$. Evaluation of the limit in (9.6) yields the desired result.

(ii) The fact that $\underset{\sim}{\phi}_{K}^{(n)}$ has asymptotic level α directly follows from the asymptotic null distribution obtained in *(i)* and the classical Helly-Bray theorem.

(iii) Optimality is a consequence of the asymptotic equivalence of $\underset{\sim}{Q}_{K_{\mathrm{f}}}$ and $Q_{\boldsymbol{\vartheta}_0,\mathrm{f}}$ described in (5.3) under $\mathrm{g} = \mathrm{f} = (f_1, \ldots, f_m) \in (\mathcal{F}_a)^m$. □

Proof of Proposition 6.1 *(i)* It follows from Theorem 4.1 in Hallin, Paindaveine and Verdebout (2010a) that $Q_{\mathcal{N}}^{(n)\dagger}$ is asymptotically chi-square with $(m-1)k(k-1)/2$ degrees of freedom under $\mathrm{P}_{\boldsymbol{\vartheta}_0;\mathrm{g}}^{(n)}$, $\boldsymbol{\vartheta}_0 \in \mathcal{H}_0$ and $\mathrm{g} \in (\mathcal{F}_1^4)^m$. Lemma 6.1 implies that $Q_{\mathcal{N}}^{(n)\dagger}$ is asymptotically noncentral chi-square, still with $(m-1)k(k-1)/2$ degrees of freedom, but with noncentrality parameter

$$\lim_{n\to\infty} (\boldsymbol{\tau}^{IV})^{(n)\prime} \boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}}^{g,IV} (\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0;\phi}^{\mathrm{g},IV})^{\perp} \boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0;\phi,\mathrm{g}}^{g,IV} (\boldsymbol{\tau}^{IV})^{(n)} \tag{9.7}$$

under $\mathrm{P}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)};\mathrm{g}}^{(n)}$ with $\mathrm{g} \in (\mathcal{F}_a^4)^m$. Evaluation of the limit in (9.7) yields the result.

(ii) The fact that $\phi_{\mathcal{N}}^{(n)\dagger}$ has asymptotic level α directly follows from the asymptotic null distribution in *(i)* and the classical Helly-Bray theorem.

(iii) Optimality is a consequence of the asymptotic equivalence under $\mathrm{g} = (\phi, \ldots, \phi)$ of $Q_{\boldsymbol{\vartheta}_0,\mathrm{g}}^{\mathcal{N}(n)}$ and $Q_{\boldsymbol{\vartheta}_0,\phi}$ described in (5.3). □

References

[1] Airoldi, J.P. and Hoffmann, R.S. (1984). Age variation in voles (*Microtus californicus* and *Microtus ochrogaster*) and its significance for systematic studies. *Occasional Papers of the Museum of the Natural History*, University of Kansas, Lawrence, **111**, 1–45.

[2] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd edition. Wiley, New York.

[3] Bentler, P.M. and Dudgeon, P. (1996). Covariance structure analysis: statistical practice, theory, and directions. *Annu. Rev. Psych.* **47**, 563–592.

[4] Boente, G. and Orellana, L. (2001). A robust approach to common principal components. *In: Fernholz, L.T., Morgenthaler, S., Stahel, W. (Eds.), Statistics in Genetics and in the Environmental Sciences*. Birkhäuser Verlag AG, Basel, Switzerland, 117–145.

[5] Boente, G. and Orellana, L. (2004). Robust plug-in estimators in proportional scatter models. *J. Statis. Plann. Inference* **122**, 95–110.

[6] Boente, G., Pires, A.M. and Rodrigues, I.M. (2002). Influence functions and outlier detection under the common principal components model: a robust approach, *Biometrika* **89**, 861–875.

[7] Boente, G., Pires, A.M. and Rodrigues, I.M. (2009). Robust tests for the common principal components model, *J. Statis. Plann. Inference* **139**, 1332–1347.

[8] Boik, J.R. (2002). Spectral models for covariance matrices. *Biometrika* **89**, 159–182.

[9] Browne, M.W. (1984). The decomposition of multitrait-multimethod matrices. *British J. Math. Statist. Psych.* **37**, 1–21.

[10] Flury, B. (1984). Common principal components in k groups. *J. Amer. Statist. Assoc.* **79**, 892–898.

[11] Flury, B. (1986). Asymptotic theory for common principal components analysis. *Ann. Statist.* **14**, 418–430.

[12] Flury, B. (1988). Two generalizations of the common principal component model. *Biometrika* **74**, 59–69.

[13] Flury, B. and Riedwyl, H. (1988). *Multivariate Statistics: a practical approach*. Chapman and Hall, New York.

[14] Flury, B. and Gautschi, W. (1986). An algorithm for simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly diagonal form. *SIAM J. Sci. Statist. Comput.* **7**, 169–184.

[15] Hallin, M. and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity. *Ann. Statist.* **34**, 2707–2756.

[16] Hallin, M. and Paindaveine, D. (2008a). A general method for constructing pseudo-Gaussian tests. *J. Japan Statist. Soc.* **38**, 27–40.

[17] Hallin, M. and Paindaveine, D. (2008b). Optimal tests for homogeneity of covariance, scale, and shape. *J. Multivariate Anal.* **100**, 422–444.

[18] Hallin, M., Oja, H. and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape. II. Optimal R-estimation of shape. *Ann. Statist.* **34**, 2757–2789.

[19] Hallin, M., Paindaveine, D. and Verdebout, T. (2010a). Testing for common principal components under heterokurticity. *J. Nonparametr. Stat.* **22**, 879–895.

[20] Hallin, M., Paindaveine, D. and Verdebout, T. (2010b). Optimal rank-based testing for principal components. *Ann. Statist.* **38**, 3245–3299.

[21] Hallin, M. and Werker, B.J.M. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9**, 137–165.

[22] Hettmansperger, T.P. and Randles, R.H. (2002). A practical affine equivariant multivariate median. *Biometrika* **89**, 851–860.

[23] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.* **24**, 417–441.

[24] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer, New York.

[25] Le Cam, L. and Yang, G.L. (2000). *Asymptotics in Statistics*, 2nd edition. Springer-Verlag, New York.

[26] Muirhead, R.J. and Waternaux, C.M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. *Biometrika* **67**, 31–43.

[27] Paindaveine, D. (2006). A Chernoff-Savage result for shape. On the non-admissibility of pseudo-Gaussian methods. *J. Multivariate Anal.* **97**, 2206–2220.

[28] Paindaveine, D. (2008). A canonical definition of shape. *Statist. Probab. Lett.* **78**, 2240–2247.

[29] Pearson, K. (1901). On lines and planes of closest fit to system of points in space. *Philos. Mag.* **2**, 559–572.

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

[31] Satorra, A. and Bentler, P.M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*, 308–313.

[32] Shapiro, A. and Browne, M.W. (1987). Analysis of covariance structures under elliptical distributions. *J. Amer. Statist. Assoc.* **82**, 1092–1097.

[33] Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.* **9**, 60–62.

[34] Yuan, K.H. and Bentler, P.M. (1997). Mean and covariance structure analysis: theoretical and practical improvements. *J. Amer. Statist. Assoc.* **92**, 766–773.