T E C H N I C A L

R E P O R T

0404

Optimal rank-based tests for sphericity

HALLIN, M., and D. PAINDAVEINE

I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Optimal Rank-Based Tests for Sphericity

Marc Hallin*
and
Davy Paindaveine*

I.S.R.O., E.C.A.R.E.S., and Département de Mathématique
Université Libre de Bruxelles, Brussels, Belgium

Abstract

We propose a class of rank-based procedures for testing that the shape matrix $\mathbf{V}$ of an elliptical distribution (with unspecified center of symmetry, scale, and radial density) has some fixed value $\mathbf{V}_0$; this problem includes the problem of testing for sphericity as an important particular case. The proposed tests are invariant under translations, monotone radial transformations, rotations, and reflections with respect to the estimated center of symmetry. They are valid without any moment assumption. For adequately chosen scores, they are locally asymptotically maximin (in the Le Cam sense) at given densities. They are strictly distribution-free when the center of symmetry is specified, and asymptotically so, when it has to be estimated. The multivariate ranks used throughout are those of the distances—in the metric associated with the null value $\mathbf{V}_0$ of the shape matrix—between the observations and the (estimated) center of the distribution. Local powers and asymptotic relative efficiencies are derived with respect to the adjusted Mauchly test (a modified version proposed by Muirhead and Waternaux (1980) of the Gaussian likelihood ratio procedure) or, equivalently, with respect to (an extension of) John (1971)'s classical test for sphericity. Small sample performances are investigated via a Monte-Carlo study.

AMS 1980 subject classification : 62M15, 62G35.

Key words and phrases : Elliptical densities, Shape matrix, Multivariate ranks and signs, Tests for sphericity, Local asymptotic normality, Locally asymptotically maximin tests, Mauchly's test.

1 Introduction.

1.1 The hypothesis of sphericity.

The distribution of a k-dimensional random vector $\mathbf{X}$ is called *spherical*, if, for some $\boldsymbol{\theta} \in \mathbb{R}^k$, the density of $\mathbf{X} - \boldsymbol{\theta}$ is invariant under orthogonal transformations. For multinormal variables, sphericity is equivalent to the covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{X}$ being proportional to the identity matrix $\mathbf{I}_k$; for elliptical variables, for which finite second-order moments need not exist, sphericity is equivalent to the shape matrix $\mathbf{V}$ being equal to $\mathbf{I}_k$; see Section 1.2 for precise definitions.

*Research supported by a P.A.I. contract of the Belgian Federal Government and an Action de Recherche Concertée of the Communauté française de Belgique.

Sphericity plays a key role in a number of statistical problems. The classical analysis of linear models crucially relies on the assumption that the observations have a covariance matrix proportional to $\mathbf{I}_k$. The analysis of repeated measures data also requires that some transformation of the observations enjoys a spherical structure. Efron (1969) in a seminal paper shows how sphericity allows for performing similar tests for location based on Student's t statistic—an idea that straightforwardly extends to linear models. Adopting a Bayesian point of view, Hill (1969) even argues that sphericity, rather than multinormality, is the adequate basis for least squares in linear models. Besides this role in the technical assumptions underlying very general statistical models, the null hypothesis of sphericity is also of direct interest in some specific domains of applications, such as paleomagnetic studies in geology, animal navigation, astronomy, or wind direction data: see Baringhaus (1991) or Mardia and Jupp (1999) for references.

Because of this importance for applications, the problem of testing the null hypothesis of sphericity has a long history, and generated a considerable body of literature, which we only very briefly summarize here. For normal populations, the asymptotic theory has been thoroughly investigated. The likelihood ratio test was derived by Mauchly (1940), the locally most powerful invariant (under shift, orthogonal transformations, and multiplication by a scalar) test by John (1971, 1972) and Sugiura (1972). In their original versions, these tests unfortunately are valid under Gaussian assumptions only; but, after a slight modification, they remain valid under elliptical populations with finite fourth-order moments: see Section 5.3 of Muirhead and Waternaux (1980) for the Mauchly test, or Section 3.3 of the present paper for John's. It has been shown however (Huynh and Mandeville 1979) that all these tests behave rather badly under heavy tails (a fact that is confirmed by the Monte-Carlo study in Section 5). More robust behaviour can be expected from the test statistics proposed by Tyler (1982 and 1983), who proposes—still under moment assumptions, though—replacing covariance matrices with more robust estimators of scatter.

Non-Gaussian models have been investigated by Kariya and Eaton (1977), where elliptical densities, possibly with infinite variances, are considered. Uniformly most powerful unbiased tests are derived, basically against specified non-spherical shape alternatives, though. Such tests thus are of limited practical value. Other non-Gaussian tests are based on various concepts of multivariate skewness and kurtosis, under finite fourth order moment assumptions: see Mardia (1970), Malkovich and Afifi (1973), or Baringhaus and Henze (1992). Their performances very much depend on the particular skewness or kurtosis concepts adopted.

As a reaction to Gaussian assumptions, nonparametric tests of sphericity also have been constructed. Their main advantage is that, under very general conditions, they enjoy a "universal consistency" property, that is, they are consistent against all possible nonspherical alternatives. The drawback is that they are computationally heavy, and only achieve slow nonparametric consistency rates. Butler (1969), Rothman and Woodroofe (1972), Beran (1979), and Baringhaus (1991) belong to this vein; see Mardia (1972) for a discussion.

Another way of escaping Gaussian or fourth-order moment assumptions consists in basing the tests on statistics that are measurable with respect to invariant or distribution-free quantities, such as the multivariate concepts of signs and ranks developed, mainly, in the robustness literature—see Oja (1999) for a review. This approach has been adopted recently by Ghosh and Sengupta (2001), who propose a test based on multivariate signs, i.e., on cosines of the form $((\mathbf{X}_i - \boldsymbol{\theta})^T/\|\mathbf{X}_i - \boldsymbol{\theta}\|)(\mathbf{X}_j - \boldsymbol{\theta})/\|\mathbf{X}_j - \boldsymbol{\theta}\|$, where $\mathbf{X}_i$, $i = 1, \ldots, n$ denote the k-dimensional observations. Their test however is of a purely heuristic nature: no optimality concerns, and no asymptotic relative efficiency results. Moreover, as we shall see, using signs only (i.e., restricting to directional information) leads, somewhat surprisingly, to a strict loss of efficiency.

The approach we are adopting in the present paper is in the same spirit. However, we throughout combine robustness (distribution-freeness under sphericity, without any moment assumptions) and optimality concerns. Our tests are based on multivariate signs and the ranks of the norms of the observations centered at $\boldsymbol{\theta}$ or an estimate $\hat{\boldsymbol{\theta}}$, with test statistics that have a structure similar to that of John's. According as the center of symmetry $\boldsymbol{\theta}$ is specified or not, these statistics are strictly distribution-free under sphericity, or asymptotically so. With adequate scores, they are asymptotically optimal (in the Le Cam sense) against non-spherical elliptical distributions, at chosen radial densities. Asymptotic relative efficiencies (with respect to John or Mauchly's tests) are derived, and appear (particularly so for the van der Waerden version of our tests) to be surprisingly high.

1.2 Elliptical densities: location, scale, shape, and radial density.

Denote by $\mathbf{X}^{(n)} := (\mathbf{X}_1^{(n)\prime}, \ldots, \mathbf{X}_n^{(n)\prime})'$, $n \in \mathbb{N}$ a triangular array of k-dimensional observations. Throughout, $\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}$ are assumed to be i.i.d., with elliptical density

$$\underline{f}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}(\mathbf{x}) := c_{k,f_1} \frac{1}{\sigma^k |\mathbf{V}|^{1/2}} f_1\left(\frac{1}{\sigma}\left((\mathbf{x}-\boldsymbol{\theta})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\theta})\right)^{1/2}\right), \quad \mathbf{x} \in \mathbb{R}^k, \tag{1}$$

where $\boldsymbol{\theta} \in \mathbb{R}^k$ is a *location parameter*, $\sigma^2 \in \mathbb{R}_0^+$ a *scale parameter*, and $\mathbf{V} := (V_{ij})$, a symmetric positive definite real $k \times k$ matrix with $V_{11} = 1$, a *shape parameter*. The infinite-dimensional parameter $f_1 : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^+$ is an a.e. strictly positive function, the constant c_{k,f_1} a normalization factor depending on the dimension k and f_1.

The function f_1 conveniently but improperly (f_1 is not a probability density) will be called a *radial density*. Denote indeed by $d_i^{(n)} = d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}) := \|\mathbf{Z}_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})\|$ the modulus of the centered and *sphericized* observations $\mathbf{Z}_i^{(n)} = \mathbf{Z}_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}) := \mathbf{V}^{-1/2}(\mathbf{X}_i^{(n)} - \boldsymbol{\theta})$, $i = 1, \ldots, n$. If the $\mathbf{X}_i^{(n)}$'s have density (1), these moduli are i.i.d., with density and distribution function

$$r \mapsto \frac{1}{\sigma} \tilde{f}_{1k}\left(\frac{r}{\sigma}\right) := \frac{1}{\sigma \mu_{k-1;f_1}} \left(\frac{r}{\sigma}\right)^{k-1} f_1\left(\frac{r}{\sigma}\right) I_{[r>0]} \quad \text{and} \quad r \mapsto \tilde{F}_{1k}(r/\sigma) := \int_0^{r/\sigma} \tilde{f}_{1k}(s)\, ds,$$

respectively—provided, however, that

$$\mu_{k-1;f_1} := \int_0^\infty r^{k-1} f_1(r) dr < \infty, \tag{2}$$

an assumption we henceforth always are making on f_1. This function $\tilde{f}_{1k}$ is the *actual* radial density, and (2) thus merely ensures that it be a probability density function; in particular, it does not imply any moment restriction on $\tilde{f}_{1k}$, the $d_i^{(n)}$'s, nor the $\mathbf{X}_i^{(n)}$'s. Any square root $\mathbf{V}^{1/2}$ of $\mathbf{V}$ (satisfying $\mathbf{V}^{1/2}\mathbf{V}^{1/2\prime} = \mathbf{V}$) can be used in the above definitions—provided, of course, it is used in a consistent way. For the sake of simplicity, the symmetric root is used throughout, saving superfluous primes.

Now, if σ and f_1 (or, more precisely σ and $f_1/c_{k,f_1}$) are to be identifiable, a scale constraint is required. Still in order to avoid moment restrictions, we chose to impose that the $d_i^{(n)}$'s, under (1), have common median σ, i.e., that

$$\tilde{F}_{1k}(1) = 1/2, \quad \text{or, equivalently,} \quad (\mu_{k-1;f_1})^{-1} \int_0^1 r^{k-1} f_1(r) dr = 1/2. \tag{3}$$

Special cases are the k-variate multinormal distribution, with radial density $f_1(r) = \phi_1(r) := \exp(-a_k r^2/2)$, and the k-variate Student distributions, with radial densities (for ν degrees of freedom) $f_1(r) = f_{1,\nu}(r) := (1 + a_{k,\nu} r^2/\nu)^{-(k+\nu)/2}$; the constants a_k and $a_{k,\nu} > 0$ are such that (3) is satisfied (note that $a_k = \lim_{\nu\to\infty} a_{k,\nu}$).

Writing $\text{vech}\,\mathbf{M} := (M_{11}, (\overset{\circ}{\text{vech}}\,\mathbf{M})')'$ for the $k(k+1)/2$-dimensional vector obtained by stacking the upper-triangular elements of a $k \times k$ symmetric matrix $\mathbf{M} = (M_{ij})$, we denote by $\text{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$ or $\text{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}$ the distribution of $\mathbf{X}^{(n)}$ under given values of $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\text{vech}}\,\mathbf{V})')'$ and f_1 (f_1 satisfying (2) and (3)).

The notation $R^{(n)}_i = R^{(n)}_i(\boldsymbol{\theta}, \mathbf{V})$ will be used for the rank of $d^{(n)}_i = d^{(n)}_i(\boldsymbol{\theta}, \mathbf{V})$ among $d^{(n)}_1, \ldots, d^{(n)}_n$; under $\text{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$, the vector $(R^{(n)}_1, \ldots, R^{(n)}_n)$ is uniformly distributed over the $n!$ permutations of $(1, \ldots, n)$. Let $\mathbf{U}^{(n)}_i = \mathbf{U}^{(n)}_i(\boldsymbol{\theta}, \mathbf{V}) := \mathbf{Z}^{(n)}_i / d^{(n)}_i$: the vectors $\mathbf{U}^{(n)}_i$ under $\text{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$ are i.i.d. and uniformly distributed over the unit sphere. They are independent of the ranks $R^{(n)}_i$, and usually are considered as *multivariate signs* associated with the centered observations $(\mathbf{X}_i - \boldsymbol{\theta})$—as they are totally insensitive to transformations of $(\mathbf{X}_i - \boldsymbol{\theta})$ that preserve half lines through the origin.

1.3 Outline of the paper.

The problem we are considering actually is that of testing that the shape $\mathbf{V}$ is equal to some given value $\mathbf{V}_0$ (admissible for a shape parameter). The special case $\mathbf{V}_0 = \mathbf{I}_k$, where $\mathbf{I}_k$ stands for the k-dimensional identity matrix, yields the problem of testing for sphericity. The shape matrix $\mathbf{V}$ in this problem is thus the parameter of interest, while $\boldsymbol{\theta}$, σ^2, and f_1 play the role of nuisance parameters. Hence, it is highly desirable that the null distributions of the test statistics to be used remain invariant under variations of $\boldsymbol{\theta}$, σ^2, and f_1.

When $\boldsymbol{\theta}$ is specified, we achieve this objective by basing our tests on the signs $\mathbf{U}^{(n)}_i$ and ranks $R^{(n)}_i$ computed from $\mathbf{Z}^{(n)}_i(\boldsymbol{\theta}, \mathbf{V}_0)$, $i = 1, \ldots, n$. These tests are invariant under monotone radial transformations (including scale transformations), rotations, and reflections of the observations (with respect to $\boldsymbol{\theta}$)—hence distribution-free with respect to such transformations. When $\boldsymbol{\theta}$ is unspecified, the ranks and signs are to be computed from $\mathbf{Z}^{(n)}_i(\hat{\boldsymbol{\theta}}, \mathbf{V}_0)$, $i = 1, \ldots, n$, where $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}^{(n)}$ is an arbitrary root-n consistent estimator of the location parameter $\boldsymbol{\theta}$; we however recommend for $\hat{\boldsymbol{\theta}}$ the (rotation-equivariant) spatial median of Möttönen and Oja (1995), which is itself "sign-based". This issue is treated in Section 4.4.

The tests based on these multivariate signed-rank statistics, whether they are computed from $\boldsymbol{\theta}$ or from $\hat{\boldsymbol{\theta}}$, are locally asymptotically optimal (namely, *locally asymptotically maximin-efficient*, as the non-specification of the scale σ induces a strict loss of efficiency) in the Le Cam sense, under adequately chosen score functions.

The rest of the paper is organized as follows. In Section 2, we establish the local and asymptotic normality (with respect to the location, scale, and shape parameters) result that provides the main theoretical tool of the paper. This result allows for developing asymptotically optimal parametric procedures for $\mathbf{V}$ under specified values of f_1 and σ (with possibly unspecified $\boldsymbol{\theta}$), and asymptotically efficient procedures, still for for $\mathbf{V}$, under specified f_1 and unspecified σ and $\boldsymbol{\theta}$. This is explained in detail in Section 3, where we also derive the asymptotically optimal (efficient, at given f_1) "scale-free" tests for hypotheses of the form $\mathbf{V} = \mathbf{V}_0$ (tests for sphericity being a special case). The Gaussian version of this test is investigated further and its link with some classical tests of sphericity is discussed. In Section 4, we propose non-parametric (signed-rank-based) versions of the optimal procedures defined in Section 3, and study their invariance and asymptotic properties. Asymptotic relative efficiencies with respect to the parametric Gaussian test are derived. All these results are obtained under specified $\boldsymbol{\theta}$ first; then, in Section 4.4, we show that, under minimal regularity assumptions on the actual underlying density (essentially, those ensuring ULAN), $\boldsymbol{\theta}$ safely can be replaced by any root-n consistent estimator $\hat{\boldsymbol{\theta}}^{(n)}$. Section 5 provides some simulation results indicating that finite sample performances reflect the asymptotic performances derived in the previous sections. Finally, the appendix collects some technical proofs.

2 Uniform local asymptotic normality (ULAN).

The main technical tool to be used in the sequel is the uniform local asymptotic normality (ULAN), with respect to $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\text{vech}}\, \mathbf{V})')'$, of the families of distributions $\mathcal{P}_{f_1}^{(n)} := \left\{ \mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)} ; \boldsymbol{\vartheta} \in \boldsymbol{\Theta} \right\}$, where $\boldsymbol{\Theta}$ denotes the open region of $\mathbb{R}^{(k+k(k+1)/2)}$ such that $\sigma^2 > 0$ and $\mathbf{V}$ is positive definite. This issue has not been touched so far in the literature—except for the very particular case of bivariate distributions with finite second-order moments recently treated by Falk (2002) in his investigation of the inefficiency of empirical correlation coefficients.

In order to describe the extremely mild assumptions under which the family $\mathcal{P}_{f_1}^{(n)}$ is *uniformly* locally asymptotically normal (ULAN), we introduce the following definitions. Consider the measure space $(\Omega, \mathbb{B}_\Omega, \lambda)$, where λ is some measure on the open subset $\Omega \subset \mathbb{R}$ equipped with its Borel σ-field $\mathbb{B}_\Omega$. Denote by $L^2(\Omega, \lambda)$ the space of measurable functions $h : \Omega \to \mathbb{R}$ satisfying $\int_\Omega [h(x)]^2 \, d\lambda(x) < \infty$. In particular, consider the space $L^2(\mathbb{R}_0^+, \mu_\ell)$ (resp. $L^2(\mathbb{R}, \nu_\ell)$) of square-integrable functions w.r.t. the Lebesgue measure with weight x^ℓ on $\mathbb{R}_0^+$ (resp. with weight $e^{\ell x}$ on $\mathbb{R}$), i.e. the space of measurable functions $h : \mathbb{R}_0^+ \to \mathbb{R}$ satisfying $\int_0^\infty [h(x)]^2 x^\ell \, dx < \infty$ (resp. $h : \mathbb{R} \to \mathbb{R}$ satisfying $\int_{-\infty}^\infty [h(x)]^2 e^{\ell x} \, dx < \infty$). Recall that $g \in L^2(\Omega, \lambda)$ admits a *weak derivative* T iff $\int_\Omega g(x) \, \varphi'(x) \, dx = -\int_\Omega T(x) \, \varphi(x) \, dx$, for all infinitely differentiable (in the classical sense) compactly supported functions φ over Ω. The mapping T is also called the *derivative of* g *in the sense of distributions* in $L^2(\Omega, \lambda)$. If, moreover, T itself is in $L^2(\Omega, \lambda)$, then g belongs to $W^{1,2}(\Omega, \lambda)$, the Sobolev space of order 1 on $L^2(\Omega, \lambda)$. For the sake of simplicity, we will write $L^2(\Omega)$ and $W^{1,2}(\Omega)$, when λ is the Lebesgue measure on Ω.

The family $\mathcal{P}_{f_1}^{(n)}$ is ULAN under the following assumptions on the radial density f_1.

Assumption (A1). The mapping $x \mapsto f_1^{1/2}(x)$ is in $W^{1,2}(\mathbb{R}_0^+, \mu_{k-1})$.

Letting $\varphi_{f_1}(r) := -2(f_1^{1/2})'(r)/f_1^{1/2}(r)$, where $(f_1^{1/2})'$ stands for the weak derivative of $f_1^{1/2}$ in $L^2(\mathbb{R}_0^+, \mu_{k-1})$, Assumption (A1) ensures the finiteness of the *radial Fisher information for location*

$$\mathcal{I}_k(f_1) := \mathrm{E}[\varphi_{f_1}^2(d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})/\sigma)] = \int_0^1 \varphi_{f_1}^2(\tilde{F}_{1k}^{-1}(u)) \, du$$

(expectation is taken under $\mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)}$).

Assumption (A2). The mapping $x \mapsto f_{1;\exp}^{1/2}(x) := f_1^{1/2}(e^x)$ is in $W^{1,2}(\mathbb{R}, \nu_k)$.

Letting $\psi_{f_1}(r) := -2r^{-1}(f_{1;\exp}^{1/2})'(\ln r)/f_{1;\exp}^{1/2}(\ln r)$, where $(f_{1;\exp}^{1/2})'$ stands for the weak derivative of $f_{1;\exp}^{1/2}$ in $L^2(\mathbb{R}, \nu_k)$, Assumption (A2) ensures the finiteness of the *radial Fisher information* *for shape* (and *scale*)

$$\mathcal{J}_k(f_1) := \mathrm{E}[\psi^2_{f_1}(d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})/\sigma)(d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})/\sigma)^2] = \int_0^1 \psi^2_{f_1}(\tilde{F}^{-1}_{1k}(u))(\tilde{F}^{-1}_{1k}(u))^2\, du$$

(expectation is still taken under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$).

The score functions φ_{f_1} for location and ψ_{f_1} for shape (and scale) in principle differ. However, they do coincide (a.e.) under the following assumption (A1-2), which, though slightly more stringent than (A1) and (A2), holds for most densities considered in practice.

Assumption (A1-2). The radial density f_1 is absolutely continuous, with a.e.-derivative $\dot{f}_1$, and, letting $\varphi_{f_1} = \psi_{f_1} := -\dot{f}_1/f_1$, the integrals

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

are finite.

It should be insisted that none of these assumptions requires the existence of any moment for the radial density $\tilde{f}_{1k}$. They are satisfied, for instance, for all multivariate Student radial densities, including the Cauchy ones. Denoting by $f_{1,\nu}$, as in Section 1.2, the radial density of the k-variate t-distribution with ν degrees of freedom ($\nu \in (0, \infty)$), it can be checked that

$$\mathcal{I}_k(f_{1,\nu}) = a^2_{k,\nu}\frac{k(k+\nu)}{k+\nu+2} \quad \text{and} \quad \mathcal{J}_k(f_{1,\nu}) = \frac{k(k+2)(k+\nu)}{k+\nu+2}. \tag{4}$$

The corresponding values for the k-variate multinormal distribution can be obtained by taking limits as $\nu \to \infty$:

$$\mathcal{I}_k(\phi_1) = a^2_k\, k \quad \text{and} \quad \mathcal{J}_k(\phi_1) = k(k+2).$$

Note that $\lim_{\nu\to 0} \mathcal{J}_k(f_{1,\nu}) = k^2$, which is the lower bound of the radial information for shape/scale, since, by Jensen inequality and integration by parts,

$$(\mathcal{J}_k(f_1))^{1/2} \geq \int_0^1 \psi_{f_1}(\tilde{F}^{-1}_{1k}(u))(\tilde{F}^{-1}_{1k}(u))\, du = k; \tag{5}$$

similarly, assuming that the density in (1) has finite second-order moments, the radial information for location $\mathcal{I}_k(f_1)$ satisfies (using the Cauchy-Schwarz inequality)

$$\mathcal{I}_k(f_1) \geq k^2 \left(\int_0^1 (\tilde{F}^{-1}_{1k}(u))^2\, du\right)^{-1},$$

with equality in the multinormal case only.

The following notation is needed in the statement of ULAN. 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 $k^2 \times k^2$ *commutation matrix*, and put $\mathbf{J}_k := \sum_{i,j=1}^k (\mathbf{e}_i\mathbf{e}_j') \otimes (\mathbf{e}_i\mathbf{e}_j') = (\mathrm{vec}\, \mathbf{I}_k)(\mathrm{vec}\, \mathbf{I}_k)'$ (as usual, $\mathrm{vec}\,(\mathbf{A})$ stands for the vector resulting from stacking the columns of $\mathbf{A}$ on top of each other). Finally, let $\mathbf{M}_k$ be such that $\mathbf{M}_k'(\mathrm{v\overset{\circ}{e}ch}\, \mathbf{v}) = \mathrm{vec}\,(\mathbf{v})$ for any symmetric $k \times k$ matrix $\mathbf{v} = (v_{ij})$ such that $v_{11} = 0$.

Proposition 1 *Under Assumptions (A1) and (A2), the family* $\mathcal{P}_{f_1}^{(n)} := \left\{ \mathrm{P}_{\boldsymbol{\vartheta};f_1}^{(n)} \middle| \boldsymbol{\vartheta} \in \boldsymbol{\Theta} \right\}$ *is ULAN, with (writing* d_i *and* $\mathbf{U}_i$ *for* $d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})$ *and* $\mathbf{U}_i^{(n)}(\boldsymbol{\theta}, \mathbf{V})$*, respectively) central sequence*

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{(n)} := \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1;1}^{(n)} \\ \Delta_{\boldsymbol{\vartheta};f_1;2}^{(n)} \\ \boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1;3}^{(n)} \end{pmatrix} := \begin{pmatrix} n^{-1/2} \frac{1}{\sigma} \sum_{i=1}^{n} \varphi_{f_1}\left(\frac{d_i}{\sigma}\right) \mathbf{V}^{-1/2} \mathbf{U}_i \\ \frac{1}{2} n^{-1/2} \begin{pmatrix} \sigma^{-2} (\mathrm{vec}\, \mathbf{I}_k)' \\ \mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \end{pmatrix} \sum_{i=1}^{n} \mathrm{vec}\left(\psi_{f_1}\left(\frac{d_i}{\sigma}\right) \frac{d_i}{\sigma} \mathbf{U}_i \mathbf{U}_i' - \mathbf{I}_k\right) \end{pmatrix} \tag{6}$$

and full-rank information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1} \quad := \quad \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Gamma_{\boldsymbol{\vartheta};f_1;22} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32}' \\ \mathbf{0} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} \end{pmatrix}, \tag{7}$$

where

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;11} := \frac{1}{k\sigma^2} \mathcal{I}_k(f_1) \mathbf{V}^{-1}, \quad \Gamma_{\boldsymbol{\vartheta};f_1;22} := \frac{1}{4\sigma^4} (\mathcal{J}_k(f_1) - k^2),$$

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32} := \frac{1}{4k\sigma^2} \left(\mathcal{J}_k(f_1) - k^2\right) \mathbf{M}_k\, \mathrm{vec}\left(\mathbf{V}^{-1}\right),$$

and

$$\boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} := \frac{1}{4} \mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \left[\frac{\mathcal{J}_k(f_1)}{k(k+2)} \left(\mathbf{I}_{k^2} + \mathbf{K}_k + \mathbf{J}_k\right) - \mathbf{J}_k\right] \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \mathbf{M}_k'. \tag{8}$$

More precisely, for any $\boldsymbol{\vartheta}^{(n)} = (\boldsymbol{\theta}^{(n)\prime}, \sigma^{2(n)}, (\overset{\circ}{\mathrm{vech}}\, \mathbf{V}^{(n)})')' = \boldsymbol{\vartheta} + O(n^{-1/2})$ *and any uniformly bounded sequence* $\boldsymbol{\tau}^{(n)} := (\mathbf{t}^{(n)\prime}, s^{(n)}, (\overset{\circ}{\mathrm{vech}}\, \mathbf{v}^{(n)})')' = (\boldsymbol{\tau}_1^{(n)\prime}, \tau_2^{(n)}, \boldsymbol{\tau}_3^{(n)\prime})' \in \mathbb{R}^{k+k(k+1)/2}$*, we have*

$$\begin{aligned} \Lambda_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)}/\boldsymbol{\vartheta}^{(n)};f_1}^{(n)} &:= \log\left(d\mathrm{P}_{\boldsymbol{\vartheta}^{(n)}+n^{-1/2}\boldsymbol{\tau}^{(n)};f_1}^{(n)} / d\mathrm{P}_{\boldsymbol{\vartheta}^{(n)};f_1}^{(n)}\right) \\ &= (\boldsymbol{\tau}^{(n)})' \boldsymbol{\Delta}_{\boldsymbol{\vartheta}^{(n)};f_1}^{(n)} - \frac{1}{2} (\boldsymbol{\tau}^{(n)})' \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1} \boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1) \end{aligned}$$

and

$$\boldsymbol{\Delta}_{\boldsymbol{\vartheta}^{(n)};f_1}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1}\right)$$

under $\mathrm{P}_{\boldsymbol{\vartheta}^{(n)};f_1}^{(n)}$*, as* $n \to \infty$*.*

Proof. See Appendix (Section 6.1). □

3 Parametrically efficient tests for shape.

3.1 An efficient central sequence for shape.

The block-diagonal structure of the information matrix (7) and ULAN imply that substituting a (in principle, discretized: see, e.g., Le Cam and Yang 2000, page 125) root-n consistent estimator $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}^{(n)}$ for the unknown $\boldsymbol{\theta}$ has no influence, asymptotically, on the $(\sigma^2, \mathbf{V})$-part of the central sequence $\boldsymbol{\Delta}_{\boldsymbol{\vartheta};f_1}^{(n)}$: optimal inference about $(\sigma^2, \mathbf{V})$ thus can be based, without any loss of (asymptotic) efficiency, on $(\Delta_{\hat{\boldsymbol{\theta}},\sigma^2,\mathbf{V};f_1;2}^{(n)}, \boldsymbol{\Delta}_{\hat{\boldsymbol{\theta}},\sigma^2,\mathbf{V};f_1;3}^{(n)\prime})'$ as if $\hat{\boldsymbol{\theta}}$ were the actual location parameter: see Section 4.4 for details. Therefore, in this section, we throughout assume that $\boldsymbol{\theta}$ is known. Similarly, replacing σ^2 and $\mathbf{V}$ with root-n consistent estimators $\hat{\sigma}^{2(n)}$ and $\hat{\mathbf{V}}^{(n)}$ in the $\boldsymbol{\theta}$-part of the central sequence $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};f_1}$ has no impact, asymptotically, on inference about $\boldsymbol{\theta}$.

Unlike the asymptotic covariances between the location and scatter components of the central sequence $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};f_1}$, the asymptotic covariances between the σ^2-part $\Delta^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1;2}$ and the $\mathbf{V}$-part $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1;3}$ are not zero. This means that a local perturbation of scale has the same asymptotic impact on $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1;3}$ as a local perturbation of $\mathbf{V}$, hence the cost of not knowing the actual value of σ^2 is strictly positive when performing inference on $\mathbf{V}$. Since we hardly can think of any practical problem where the scale (but not the shape) is specified, we concentrate on optimality under unspecified scale σ^2, and explicitly compute the information loss due to the presence of this nuisance.

LAN and the convergence of local experiments to the Gaussian shift experiment

$$\begin{pmatrix} \Delta_2 \\ \boldsymbol{\Delta}_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \Gamma_{\boldsymbol{\vartheta};f_1;22} & \boldsymbol{\Gamma}'_{\boldsymbol{\vartheta};f_1;32} \\ \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} \end{pmatrix} \begin{pmatrix} \tau_2 \\ \boldsymbol{\tau}_3 \end{pmatrix}, \begin{pmatrix} \Gamma_{\boldsymbol{\vartheta};f_1;22} & \boldsymbol{\Gamma}'_{\boldsymbol{\vartheta};f_1;32} \\ \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32} & \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} \end{pmatrix} \right), \quad (\tau_2, \boldsymbol{\tau}'_3)' \in \mathbb{R}^{k(k+1)/2} \tag{9}$$

implies that locally optimal inference on shape, in the presence of an unspecified scale parameter, should be based on the residual of the regression (in (9)), of $\boldsymbol{\Delta}_3$ with respect to Δ_2, computed at $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};f_1;3}$ (the shape part of the central sequence) and $\Delta^{(n)}_{\boldsymbol{\vartheta};f_1;2}$ (the scale part of the same). This residual takes the form $\boldsymbol{\Delta}_3 - \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32}\Gamma^{-1}_{\boldsymbol{\vartheta};f_1;22}\Delta_2$; the resulting f_1-*efficient central sequence for shape* is thus

$$\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1} = \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};f_1;3} - \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32}\Gamma^{-1}_{\boldsymbol{\vartheta};f_1;22}\Delta^{(n)}_{\boldsymbol{\vartheta};f_1;2},$$

which, after some elementary algebra, reduces to

$$\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1} = \frac{1}{2} n^{-1/2}\mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right] \sum_{i=1}^{n} \psi_{f_1}\left(\frac{d_i}{\sigma}\right)\frac{d_i}{\sigma} \operatorname{vec}\left(\mathbf{U}_i\mathbf{U}'_i\right). \tag{10}$$

This efficient central sequence under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta},f_1}$ is asymptotically normal, with mean zero and covariance (the *efficient Fisher information for shape* under radial density f_1)

$$\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta};f_1} = \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;33} - \boldsymbol{\Gamma}_{\boldsymbol{\vartheta};f_1;32}\Gamma^{-1}_{\boldsymbol{\vartheta};f_1;22}\boldsymbol{\Gamma}'_{\boldsymbol{\vartheta};f_1;32}.$$

After routine computation, this efficient information takes the form

$$\begin{aligned}\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta};f_1} &= \frac{1}{4}\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right]\left[\frac{\mathcal{J}_k(f_1)}{k(k+2)}\left(\mathbf{I}_{k^2} + \mathbf{K}_k + \mathbf{J}_k\right) - \mathbf{J}_k\right]\left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}'_k \\ &= \frac{\mathcal{J}_k(f_1)}{4k(k+2)}\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}'_k =: \mathcal{J}_k(f_1)\,\mathbf{G}_k(\mathbf{V}), \end{aligned} \tag{11}$$

a form that is not unfamiliar in the area of robust estimation of covariance matrices: see, for instance, the asymptotic covariances in Tyler (1982, 1983), Ollila, Oja, and Croux (2003), and Ollila, Croux, and Oja (2004) for the covariances of scatter estimates (as in (7), (8)), Tyler (1987) and Ollila, Hettmansperger, and Oja (2004) for covariances of shape estimates (as in (11)).

In the sequel, *optimality* (in the local and asymptotic sense, at radial density f_1) is to be understood in the context of the Gaussian shift experiment associated with efficient central sequences $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$. In particular, a sequence of tests will be called locally and asymptotically *maximin-efficient* (at asymptotic level α) if it is asymptotically maximin in the sequence of experiments associated with $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$.

3.2 Optimal parametric tests for shape.

Consider the problem of testing a null hypothesis of the form $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ in the parametric model where f_1 is known and the scale σ^2 is unspecified. Optimality (in a local and asymptotic sense; see Proposition 2 for a precise statement) is reached by tests based on quadratic forms in the f_1-efficient central sequence for shape. More precisely, the optimal test statistics take the form

$$Q_{f_1} = Q_{f_1}^{(n)} := \left(\boldsymbol{\Delta}^{\star(n)}_{\hat{\boldsymbol{\vartheta}}_0;f_1}\right)' \left(\boldsymbol{\Gamma}^{\star}_{\hat{\boldsymbol{\vartheta}}_0;f_1}\right)^{-1} \boldsymbol{\Delta}^{\star(n)}_{\hat{\boldsymbol{\vartheta}}_0;f_1},$$

where, denoting by $\hat{\sigma}$ a root-n consistent estimator for σ, we let $\hat{\boldsymbol{\vartheta}}_0 := (\boldsymbol{\theta}', \hat{\sigma}^2, (\overset{\circ}{\text{vech}}\,\mathbf{V}_0)')'$. Note that consistent estimation of σ under $\bigcup_{f_1} \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1}\right\}$ is easily achieved, since σ then is the common median of the distances $d_i(\boldsymbol{\theta}, \mathbf{V}_0)$. As we will see in Section 3.3, the Gaussian version of these optimal parametric tests allows for bypassing this estimation of the scale.

Lemma 1 below allows for the explicit form

$$Q_{f_1} = \frac{k(k+2)}{2n\mathcal{J}_k(f_1)} \sum_{i,j=1}^{n} \frac{d_i d_j}{\hat{\sigma}^2}\, \psi_{f_1}\!\left(\frac{d_i}{\hat{\sigma}}\right) \psi_{f_1}\!\left(\frac{d_j}{\hat{\sigma}}\right) \left((\mathbf{U}_i'\mathbf{U}_j)^2 - \frac{1}{k}\right),$$

with $d_i := d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$ and $\mathbf{U}_i := \mathbf{U}_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$.

Lemma 1 *Denote by $\mathbf{e}_{k^2,1}$ the first vector of the canonical basis of $\mathbb{R}^{k^2}$. Then, if $\mathbf{V} = (V_{ij})$ is symmetric with $V_{11} = 1$, we have*

$$\mathbf{M}_k' \left\{\frac{1}{4}\mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \left[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\right] \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \mathbf{M}_k'\right\}^{-1} \mathbf{M}_k \tag{12}$$

$$= [\mathbf{I}_{k^2} + \mathbf{K}_k]\,(\mathbf{V}^{\otimes 2}) - 2(\mathbf{V}^{\otimes 2})\,\mathbf{e}_{k^2,1}\,(\text{vec}\mathbf{V})' - 2\,(\text{vec}\mathbf{V})\,(\mathbf{e}_{k^2,1})'(\mathbf{V}^{\otimes 2}) + 2\,(\text{vec}\mathbf{V})\,(\text{vec}\mathbf{V})'.$$

Proof. See Appendix (Section 6.2). □

Proposition 2 *Let f_1 satisfy Assumptions (A1) and (A2). Then, denoting by $\|\mathbf{A}\| := [\text{tr}(\mathbf{A}\mathbf{A}')]^{1/2}$ the Frobenius norm of the array $\mathbf{A}$,*

(i) $Q_{f_1}^{(n)}$ is asymptotically chi-square with $k(k+1)/2-1$ degrees of freedom under $\bigcup_{\sigma^2} \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1}\right\}$, and asymptotically noncentral chi-square, still with $k(k+1)/2-1$ degrees of freedom but with noncentrality parameter

$$\frac{\mathcal{J}_k(f_1)}{2k(k+2)} \left[\text{tr}\left(\mathbf{V}_0^{-1}\mathbf{v}\right)^2 - \frac{1}{k}\left(\text{tr}\ \mathbf{V}_0^{-1}\mathbf{v}\right)^2\right] = \frac{\mathcal{J}_k(f_1)}{2k(k+2)} \left(\text{tr}\ \mathbf{V}_0^{-1}\mathbf{v}\right)^2 \left\|\left(\frac{\mathbf{V}_0^{-1}\mathbf{v}}{\text{tr}\,\mathbf{V}_0^{-1}\mathbf{v}}\right) - \frac{1}{k}\mathbf{I}_k\right\|^2$$

under $\bigcup_{\sigma^2} \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0+n^{-1/2}\mathbf{v};f_1}\right\}$;

(ii) the sequence of tests $\phi_{f_1}^{(n)}$ which consists in rejecting $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ as soon as $Q_{f_1}^{(n)}$ exceeds the α upper-quantile of a chi-square variable with $k(k+1)/2-1$ degrees of freedom has asymptotic level α under $\bigcup_{\sigma^2} \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1}\right\}$, and is locally and asymptotically maximin-efficient, still at asymptotic level α, for $\bigcup_{\sigma^2} \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1}\right\}$ against alternatives of the form $\bigcup_{\sigma^2} \bigcup_{\mathbf{V}\neq\mathbf{V}_0} \left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}\right\}$.

Proof. See Appendix (Section 6.2). □

In contrast with this unspecified-σ^2 test, the locally and asymptotically optimal procedure for testing $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ under specified radial density f_1, specified $\boldsymbol{\theta}$, *and specified scale* σ^2 rejects $\mathcal{H}_0$ (at asymptotic level α) whenever

$$Q_{\sigma^2,f_1} = Q^{(n)}_{\sigma^2,f_1} := \left(\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1;3}\right)' \left(\boldsymbol{\Gamma}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1;33}\right)^{-1} \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1;3}$$

exceeds the α upper-quantile of a chi-square with $k(k+1)/2-1$ degrees of freedom. The efficiency loss due to an unspecified σ^2 thus can be measured by the difference between the non-centrality parameters in the asymptotic chi-square distributions of Q_{σ^2,f_1} and Q_{f_1} under local alternatives. Along the same lines as in the proof of Proposition 2, one can show that this difference, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0+n^{-1/2}\mathbf{v};f_1}$, is

$$\frac{1}{4k}\left(\mathcal{J}_k(f_1) - k^2\right)\left(\mathrm{tr}\ \mathbf{V}_0^{-1}\mathbf{v}\right)^2.$$

Inequality (5) confirms the non-surprising fact that this loss is nonnegative and an increasing function of the information for shape (or scale) $\mathcal{J}_k(f_1)$. Quite remarkably, it does not depend on σ^2.

3.3 Optimal Gaussian tests for shape.

The parametric tests $\phi^{(n)}_{f_1}$ described in Part *(ii)* of Proposition 2 achieve local and asymptotic optimality at radial density f_1, but in general are not valid when the underlying radial density is $g_1 \neq f_1$. If correctly formulated, the Gaussian version of these tests (obtained for $f_1 = \phi_1$, where ϕ_1 was defined in Section 1.2) is an interesting exception to this rule, and can easily be written under a form that remains valid under the class of all radial densities g_1 with finite fourth-order moments.

Denote by $D_k(g_1) := \mathrm{E}[(\tilde{G}^{-1}_{1k}(U))^2]$ and $E_k(g_1) := \mathrm{E}[(\tilde{G}^{-1}_{1k}(U))^4] < \infty$, where U stands for a random variable with uniform distribution over $]0, 1[$, the second and fourth order moments of $\tilde{g}_{1k}$, respectively, and assume that $E_k(g_1) < \infty$ (hence also $D_k(g_1) < \infty$). These two quantities are closely related to the *kurtosis* of the elliptic distribution under consideration. To be precise, the kurtosis $3\kappa_k(g_1)$ of an elliptically symmetric random k-vector $\mathbf{X} = (X_i)$ with location center $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)'$, scale σ^2, shape matrix $\mathbf{V}$, and radial density g_1 is defined (see, e.g., Muirhead and Waternaux (1980) or Tyler (1982)) by

$$3\kappa_k(g_1) := \frac{\mathrm{E}[(X_i - \theta_i)^4]}{\mathrm{E}^2[(X_i - \theta_i)^2]} - 3.$$

This quantity only depends on the dimension k and the radial density f_1—and not on i, nor on the other parameters characterizing the elliptical distribution (which of course justifies the notation); it is related to $D_k(g_1)$ and $E_k(g_1)$ by the simple relation

$$\kappa_k(g_1) = \frac{k}{k+2}\frac{E_k(g_1)}{D_k^2(g_1)} - 1.$$

At the k-variate Gaussian distribution and t-distribution with ν degrees of freedom ($\nu > 4$), this kurtosis parameter takes values $\kappa_k(\phi_1) = 0$ and $\kappa_k(f_{1,\nu}) = 2/(\nu - 4)$, respectively.

The Gaussian version of the efficient central sequence for shape $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$ can be written as $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};\phi_1} = a_k\sigma^{-2}\mathbf{T}_{\boldsymbol{\theta},\mathbf{V}}$, where

$$\mathbf{T}_{\boldsymbol{\theta},\mathbf{V}} = \mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{V}} := \frac{1}{2}\, n^{-1/2}\mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\sum_{i=1}^n \mathrm{vec}\Big((\mathbf{X}_i-\boldsymbol{\theta})(\mathbf{X}_i-\boldsymbol{\theta})'\Big).$$

Working with $\mathbf{T}_{\boldsymbol{\theta},\mathbf{V}}$ and an estimate $\hat{\boldsymbol{\Gamma}}^{(n)}$ of its asymptotic covariance rather than with $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};\phi_1}$ and an estimate of the corresponding information matrix is convenient, since the scalar factor $a_k\sigma^{-2}$ in the quadratic form in $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};\phi_1}$ cancels out. For optimality (at Gaussian radial densities), it is sufficient for $\hat{\boldsymbol{\Gamma}}^{(n)}$ to consistently estimate the asymptotic covariance of $\mathbf{T}_{\boldsymbol{\theta},\mathbf{V}_0}$ under $\bigcup_{\sigma^2}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;\phi_1}\right\}$.

Letting

$$\hat{\boldsymbol{\Gamma}}^{(n)} := \frac{1}{4k(k+2)}\left(\frac{1}{n}\sum_{i=1}^n d_i^4\right)\mathbf{M}_k\left(\mathbf{V}_0^{\otimes 2}\right)^{-1/2}\left[\mathbf{I}_{k^2}+\mathbf{K}_k-\frac{2}{k}\mathbf{J}_k\right]\left(\mathbf{V}_0^{\otimes 2}\right)^{-1/2}\mathbf{M}'_k,$$

with the same $d_i = d_i^{(n)}(\boldsymbol{\theta},\mathbf{V}_0)$'s as in Section 3.2, it is easy to check that $\hat{\boldsymbol{\Gamma}}^{(n)}$ provides, for all $\boldsymbol{\theta}$, a consistent estimate for the asymptotic variance of $\mathbf{T}_{\boldsymbol{\theta},\mathbf{V}_0}$, not only under $\bigcup_{\sigma^2}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;\phi_1}\right\}$, but also under $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$, where union is taken over the set of all radial densities g_1 such that $E_k(g_1)<\infty$. The Gaussian test statistic then takes the form $Q_{\mathcal{N}} = Q^{(n)}_{\mathcal{N}} := \mathbf{T}^{(n)\prime}_{\boldsymbol{\theta},\mathbf{V}_0}(\hat{\boldsymbol{\Gamma}}^{(n)})^{-1}\mathbf{T}^{(n)}_{\boldsymbol{\theta},\mathbf{V}_0}$. Lemma 1 and standard algebra yield

$$Q_{\mathcal{N}} = \frac{k(k+2)}{2(\sum_{i=1}^n d_i^4)}\sum_{i,j=1}^n d_i^2 d_j^2\left((\mathbf{U}_i'\mathbf{U}_j)^2-\frac{1}{k}\right), \qquad (13)$$

with the same $\mathbf{U}_i = \mathbf{U}_i^{(n)}(\boldsymbol{\theta},\mathbf{V}_0)$ as in Section 3.2. Now, defining

$$\mathbf{S} = \mathbf{S}^{(n)} := \frac{1}{n}\sum_{i,j=1}^n [\mathbf{V}_0^{-1/2}(\mathbf{X}_i-\boldsymbol{\theta})][\mathbf{V}_0^{-1/2}(\mathbf{X}_i-\boldsymbol{\theta})]'$$

and letting $\hat{\kappa}^{(n)} := [k(n^{-1}\sum_{i=1}^n d_i^4)]/[(k+2)(n^{-1}\sum_{i=1}^n d_i^2)^2]-1$ be a consistent estimate of the kurtosis parameter $\kappa_k(g_1)$, (13) can be written under the form

$$Q_{\mathcal{N}} = \frac{n^2k(k+2)}{2(\sum_{i=1}^n d_i^4)}\left(\mathrm{tr}\,\mathbf{S}^2 - \frac{1}{k}\,\mathrm{tr}^2\mathbf{S}\right) = \frac{1}{1+\hat{\kappa}^{(n)}}\,\frac{nk^2}{2}\left\|\frac{\mathbf{S}}{\mathrm{tr}\,\mathbf{S}}-\frac{1}{k}\,\mathbf{I}_k\right\|^2. \qquad (14)$$

It is straightforward to check that $Q_{\mathcal{N}}$ is invariant under rotations, scale transformations, and reflections (with respect to $\boldsymbol{\theta}$, in the metric associated with $\mathbf{V}_0$), but that it is not (even asymptotically) invariant under the group of monotone continuous radial transformations (see Section 4.1 below). The following proposition summarizes the asymptotic properties of the Gaussian procedure based on $Q_{\mathcal{N}}$.

Proposition 3 *Denote by $\phi^{(n)}_{\mathcal{N}}$ the parametric Gaussian test which rejects the null hypothesis $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ whenever $Q^{(n)}_{\mathcal{N}}$ exceeds the α upper-quantile of a chi-square with $k(k+1)/2-1$ degrees of freedom. Then (unions over g_1 are taken over all radial densities such that $\tilde{g}_{1k}$ has finite fourth-order moments),* *(i)* $Q^{(n)}_{\mathcal{N}}$ *is asymptotically chi-square with* $k(k+1)/2-1$ *degrees of freedom under* $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$*, and asymptotically noncentral chi-square, still with* $k(k+1)/2-1$ *degrees of freedom but with noncentrality parameter*

$$\frac{1}{2(1+\kappa_k(g_1))}\left[\mathrm{tr}\left(\mathbf{V}_0^{-1}\mathbf{v}\right)^2-\frac{1}{k}\left(\mathrm{tr}\ \mathbf{V}_0^{-1}\mathbf{v}\right)^2\right]$$

under $\bigcup_{\sigma^2}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0+n^{-1/2}\mathbf{v};g_1}\right\}$*;*

(ii) the sequence of tests $\phi^{(n)}_{\mathcal{N}}$ *has asymptotic level* α *under* $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$*, and is locally and asymptotically maximin-efficient, still at asymptotic level* α*, for* $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$ *against alternatives of the form* $\bigcup_{\sigma^2}\bigcup_{\mathbf{V}\neq\mathbf{V}_0}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};\phi_1}\right\}$.

Proof. See Appendix (Section 6.3). □

For $\mathbf{V}_0=\mathbf{I}_k$, the test statistic $Q_{\mathcal{N}}$ in (14) and Proposition 3 actually appears as a modification of the test statistic

$$Q_{\mathrm{John}}:=\frac{nk^2}{2}\left\|\frac{\mathbf{S}}{\mathrm{tr}\,\mathbf{S}}-\frac{1}{k}\mathbf{I}_k\right\|^2=\frac{nk^2}{2}\,\mathrm{tr}\left[\left(\frac{\mathbf{S}}{\mathrm{tr}\,\mathbf{S}}-\frac{1}{k}\mathbf{I}_k\right)^2\right]\tag{15}$$

proposed by John (1971, 1972). The only difference is that Q_{John} uses the Gaussian value $\kappa=0$ of the kurtosis parameter, whereas $Q_{\mathcal{N}}$ rather involves an estimation $\hat{\kappa}^{(n)}$ of the same, which makes the asymptotic null distribution of $Q_{\mathcal{N}}$ agree, under any elliptical distribution with finite fourth-order moments, with the limiting distribution of Q_{John} in the multinormal case.

This adjustment is very much in the spirit of Muirhead and Waternaux (1980)'s version of Mauchly (1940)'s Gaussian likelihood ratio test—probably the most widely used test of sphericity. Muirhead and Waternaux (1980) actually show that the limiting distribution of $(-2\log\Lambda^{(n)})/(1+\kappa_k(g_1))$, where $-2\log\Lambda^{(n)}$ is the Gaussian likelihood ratio test statistic, is asymptotically chi-square, with $k(k+1)/2-1$ degrees of freedom, under $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{I}_k;g_1}\right\}$ (union, again, is over all g_1 such that $\tilde{g}_{1k}$ has finite fourth-order moments); the population kurtosis parameter $\kappa_k(g_1)$ of course can be replaced by its sample counterpart $\hat{\kappa}^{(n)}$ without modifying the asymptotic chi-square distribution. These results straightforwardly extend to the problem of testing for a specified shape $\mathbf{V}_0$ rather than for sphericity. It also follows from Muirhead and Waternaux (1980)'s treatment that the adjusted version of John's test statistic, namely our Gaussian test statistic $Q_{\mathcal{N}}$, is asymptotically equivalent to their adjusted version of the Mauchly test. In the sequel, the expression "optimal parametric Gaussian test" will refer to any of these tests. Note however that optimality here follows from Proposition 3, and therefore is of an asymptotic nature. Actually, only John (1971)'s original (non-adjusted) test enjoys some finite-sample optimality properties (restricted to the Gaussian case), being *locally most powerful invariant* at the multinormal distribution. Our adjusted tests inherit, under weaker asymptotic form, this optimality property from John's test; on the other hand, they remain valid under non Gaussian densities, which is not the case of John's.

4 Rank-based tests for shape.

4.1 Rank-based versions of efficient central sequences for shape.

As already mentioned, the problem, with tests based on efficient central sequences, is that (with the exception of the adjusted Gaussian tests described in Section 3.3) they are valid under correctly specified radial densities only. A correct specification f_1 of the actual radial density g_1 in practice is rather unrealistic, and the problem thus has to be treated from a semiparametric point of view, where g_1 plays the role of a nuisance.

Within the family of distributions $\bigcup_{\sigma^2} \bigcup_{\mathbf{V}} \bigcup_{g_1} \left\{ \mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1} \right\}$, where $\boldsymbol{\theta}$ is fixed, consider the null hypothesis $\mathcal{H}_0(\boldsymbol{\theta}, \mathbf{V}_0)$ where $\mathbf{V} = \mathbf{V}_0$. Throughout, thus, $\boldsymbol{\theta}$ is fixed, and σ^2 and the radial density g_1 remain unspecified (no moment assumptions here). As we have seen, the scalar nuisance σ^2 can be taken care of by means of a simple projection, yielding the efficient central sequence. In principle, the infinite-dimensional nuisance g_1 can be treated similarly, by projecting central sequences along adequate *tangent spaces*. This approach however is rather technical. Hallin and Werker (2003) have showed that appropriate group invariance structures allow for the same result by conditioning central sequences with respect to maximal invariants such as ranks or signs. This is the approach we also adopt here.

Clearly, the null hypothesis $\mathcal{H}_0(\boldsymbol{\theta}, \mathbf{V}_0)$ is invariant under the following two groups of transformations, acting on the observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$:

(i) the group $\mathcal{G}_{\mathrm{orth}},_\circ := \mathcal{G}^{\boldsymbol{\theta},\mathbf{V}_0}_{\mathrm{orth}},_\circ$ of $\mathbf{V}_0$-*orthogonal transformations* (centered at $\boldsymbol{\theta}$) consisting of all transformations

$$\begin{aligned}
\mathbf{X} \mapsto\ & g_{\mathbf{O}}(\mathbf{X}_1, \ldots, \mathbf{X}_n) \\
=\ & g_{\mathbf{O}}(\boldsymbol{\theta} + d_1(\boldsymbol{\theta}, \mathbf{V}_0)\mathbf{V}_0^{1/2}\mathbf{U}_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, \boldsymbol{\theta} + d_n(\boldsymbol{\theta}, \mathbf{V}_0)\mathbf{V}_0^{1/2}\mathbf{U}_n(\boldsymbol{\theta}, \mathbf{V}_0)) \\
:=\ & (\boldsymbol{\theta} + d_1(\boldsymbol{\theta}, \mathbf{V}_0)\mathbf{V}_0^{1/2}\mathbf{O}\mathbf{U}_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, \boldsymbol{\theta} + d_n(\boldsymbol{\theta}, \mathbf{V}_0)\mathbf{V}_0^{1/2}\mathbf{O}\mathbf{U}_n(\boldsymbol{\theta}, \mathbf{V}_0)),
\end{aligned}$$

where $\mathbf{O}$ is an arbitrary $k \times k$ orthogonal matrix. This group contains in particular "rotations" (in the metric associated with $\mathbf{V}_0$) around $\boldsymbol{\theta}$, as well as *reflection* (still with respect to $\boldsymbol{\theta}$), i.e., the mapping $(\mathbf{X}_1, \ldots, \mathbf{X}_n) \mapsto (\boldsymbol{\theta} - (\mathbf{X}_1 - \boldsymbol{\theta}), \ldots, \boldsymbol{\theta} - (\mathbf{X}_n - \boldsymbol{\theta}))$;

(ii) the group $\mathcal{G},_\circ := \mathcal{G}^{\boldsymbol{\theta},\mathbf{V}_0},_\circ$ of *continuous monotone radial transformations*, of the form

$$\begin{aligned}
\mathbf{X} \mapsto\ & g_g(\mathbf{X}_1, \ldots, \mathbf{X}_n) \\
=\ & g_g(\boldsymbol{\theta} + d_1(\boldsymbol{\theta}, \mathbf{V}_0)\mathbf{V}_0^{1/2}\mathbf{U}_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, \boldsymbol{\theta} + d_n(\boldsymbol{\theta}, \mathbf{V}_0)\mathbf{V}_0^{1/2}\mathbf{U}_n(\boldsymbol{\theta}, \mathbf{V}_0)) \\
:=\ & (\boldsymbol{\theta} + g(d_1(\boldsymbol{\theta}, \mathbf{V}_0))\mathbf{V}_0^{1/2}\mathbf{U}_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, \boldsymbol{\theta} + g(d_n(\boldsymbol{\theta}, \mathbf{V}_0))\mathbf{V}_0^{1/2}\mathbf{U}_n(\boldsymbol{\theta}, \mathbf{V}_0)),
\end{aligned}$$

where $g : \mathbb{R}^+ \to \mathbb{R}^+$ is continuous, monotone increasing, and such that $g(0) = 0$ and $\lim_{r\to\infty} g(r) = \infty$. This group includes in particular the subgroup of all *scale transformations* $(\mathbf{X}_1, \ldots, \mathbf{X}_n) \mapsto (\boldsymbol{\theta} + a(\mathbf{X}_1 - \boldsymbol{\theta}), \ldots, \boldsymbol{\theta} + a(\mathbf{X}_n - \boldsymbol{\theta}))$, $a > 0$.

Clearly, the group $\mathcal{G},_\circ$ of continuous monotone radial transformations is a generating group for the family of distributions $\bigcup_{\sigma^2} \bigcup_{f_1} \left\{ \mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1} \right\}$, that is, a generating group for the null hypothesis $\mathcal{H}_0(\boldsymbol{\theta}, \mathbf{V}_0)$ under consideration. The invariance principle therefore leads to consider test statistics that are measurable with respect to the corresponding maximal invariant, namely the vector $(R_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, R_n(\boldsymbol{\theta}, \mathbf{V}_0), \mathbf{U}_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, \mathbf{U}_n(\boldsymbol{\theta}, \mathbf{V}_0))$, where $R_i(\boldsymbol{\theta}, \mathbf{V}_0)$ denotes the rank of $d_i(\boldsymbol{\theta}, \mathbf{V}_0)$ among $d_1(\boldsymbol{\theta}, \mathbf{V}_0), \ldots, d_n(\boldsymbol{\theta}, \mathbf{V}_0)$. The resulting signed rank test statistics are (strictly) invariant under $\mathcal{G}_{,\circ}$, hence distribution-free under $\mathcal{H}_0(\boldsymbol{\theta}, \mathbf{V}_0)$.

Now, in the construction of the proposed tests for the null hypothesis $\mathcal{H}_0(\boldsymbol{\theta}, \mathbf{V}_0)$, we intend to combine invariance and optimality arguments by considering a (signed-)rank-based version of the f_1-efficient central sequences for shape (recall that central sequences are always defined up to $o_{\mathrm{P}}(1)$—under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}$, as $n \to \infty$—terms). The signed-rank version $\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$ of the shape-efficient central sequence $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$ we plan to use in our non-parametric tests is the f_1-score version (that obtained with $K_{f_1}(u) := \psi_{f_1}(\tilde{F}_1^{-1}(u))\tilde{F}_1^{-1}(u)$) of the statistic

$$\begin{aligned}
\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\vartheta};K} &:= \frac{1}{2} n^{-1/2} \mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right] \sum_{i=1}^{n} K\left(\frac{R_i}{n+1}\right) \operatorname{vec}\left(\mathbf{U}_i \mathbf{U}_i'\right) \\
&= \frac{1}{2} n^{-1/2} \mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \sum_{i=1}^{n} K\left(\frac{R_i}{n+1}\right) \operatorname{vec}\left(\mathbf{U}_i \mathbf{U}_i' - \frac{1}{k}\mathbf{I}_k\right) \\
&= \frac{1}{2} n^{-1/2} \mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \sum_{i=1}^{n} \left(K\left(\frac{R_i}{n+1}\right) \operatorname{vec}\left(\mathbf{U}_i \mathbf{U}_i'\right) - \frac{m_K^{(n)}}{k}\mathbf{I}_k\right),
\end{aligned} \tag{16}$$

where $R_i = R_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$ denotes the rank of $d_i = d_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$ among $d_1, \ldots, d_n$, $\mathbf{U}_i = \mathbf{U}_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$, and $m_K^{(n)} := n^{-1}\sum_{i=1}^{n} K(i/(n+1))$.

Beyond its role in the derivation of the asymptotic distribution of the rank-based random vector (16), the following *asymptotic representation* result shows that $\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$ is indeed another version of the efficient central sequence $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};f_1}$ under radial density f_1.

Lemma 2 *Assume that the score function $K :]0,1[\to \mathbb{R}$ is continuous, square-integrable, and that it can be expressed as the difference of two monotone increasing functions. Then, defining*

$$\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};K;f_1} := \frac{1}{2} n^{-1/2} \mathbf{M}_k \left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right] \sum_{i=1}^{n} K\left(\tilde{F}_{1k}\left(\frac{d_i}{\sigma}\right)\right) \operatorname{vec}\left(\mathbf{U}_i \mathbf{U}_i'\right), \tag{17}$$

we have $\underset{\sim}{\boldsymbol{\Delta}}^{\star(n)}_{\boldsymbol{\vartheta};K} = \boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};K;f_1} + o_{L^2}(1)$, as n goes to infinity, under $\left\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}\right\}$.

Proof. See Appendix (Section 6.3). □

4.2 The proposed class of tests.

Let $K :]0,1[\to \mathbb{R}$ be some *score* function as in Lemma 2: the K-score version of the statistics we propose for testing $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ is

$$\underset{\sim}{Q}_K = \underset{\sim}{Q}_K^{(n)} := \frac{k(k+2)}{2n\mathrm{E}[K^2(U)]} \sum_{i,j=1}^{n} K\left(\frac{R_i}{n+1}\right) K\left(\frac{R_j}{n+1}\right) \left((\mathbf{U}_i'\mathbf{U}_j)^2 - \frac{1}{k}\right), \tag{18}$$

where $R_i = R_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$ and $\mathbf{U}_i = \mathbf{U}_i^{(n)}(\boldsymbol{\theta}, \mathbf{V}_0)$. Letting $\mathbf{S}_K = \mathbf{S}_K^{(n)} = \frac{1}{n}\sum_{i,j=1}^{n} K\left(\frac{R_i}{n+1}\right)\mathbf{U}_i\mathbf{U}_i'$, these test statistics can be rewritten as

$$\underset{\sim}{Q}_K = \frac{nk(k+2)}{2\mathrm{E}[K^2(U)]}\left(\operatorname{tr}\mathbf{S}_K^2 - \frac{1}{k}\operatorname{tr}^2\mathbf{S}_K\right) = \frac{k(k+2)\mathrm{E}^2[K(U)]}{k^2\mathrm{E}[K^2(U)]}\frac{nk^2}{2}\left\|\frac{\mathbf{S}_K}{\operatorname{tr}\mathbf{S}_K} - \frac{1}{k}\mathbf{I}_k\right\|^2 + o_{\mathrm{P}}(1),$$

as n goes to infinity, under any elliptical distribution (compare with (14)). These test statistics are strictly invariant under $\mathcal{G}_{\mathrm{orth},\circ}$ and $\mathcal{G}_{,\circ}$.

The power functions $K_a(u) = u^a$, $a \geq 0$, provide some traditional score functions. The corresponding test statistics are

$$\underset{\sim}{Q}_{K_a} := \frac{(2a+1)k(k+2)}{2n(n+1)^{2a}} \sum_{i,j=1}^{n} R_i^a R_j^a \left((\mathbf{U}_i'\mathbf{U}_j)^2 - \frac{1}{k} \right). \tag{19}$$

Important particular cases are the sign-, Wilcoxon-, and Spearman-type test statistics, defined by $\underset{\sim}{Q}_S := \underset{\sim}{Q}_{K_0}$, $\underset{\sim}{Q}_W := \underset{\sim}{Q}_{K_1}$, and $\underset{\sim}{Q}_{SP} := \underset{\sim}{Q}_{K_2}$, respectively. The resulting tests in general are not optimal at any density (they sometimes are, though: for instance, the Wilcoxon test $\underset{\sim}{Q}_W$ is optimal, in dimension $k = 2$, at Student densities with two degrees of freedom, see Section 4.3), but they nevertheless yield good overall performances, and are simple to compute. The sign test statistic $\underset{\sim}{Q}_S$ coincides with the test statistic proposed in Ghosh and Sengupta (2001).

Local asymptotic optimality under radial density f_1 is achieved by $\underset{\sim}{Q}_{f_1} := \underset{\sim}{Q}_{K_{f_1}}$, with scores $K_{f_1}(u) = \psi_{f_1}(\tilde{F}_{1k}^{-1}(u))\,\tilde{F}_{1k}^{-1}(u)$. The test statistic then takes the form

$$\begin{aligned}\underset{\sim}{Q}_{f_1} = \frac{k(k+2)}{2n\mathcal{J}_k(f_1)} \sum_{i,j=1}^{n} \psi_{f_1}\circ\tilde{F}_1^{-1}\left(\frac{R_i}{n+1}\right) \tilde{F}_1^{-1}\left(\frac{R_i}{n+1}\right) \\ \psi_{f_1}\circ\tilde{F}_1^{-1}\left(\frac{R_j}{n+1}\right) \tilde{F}_1^{-1}\left(\frac{R_j}{n+1}\right) \left((\mathbf{U}_i'\mathbf{U}_j)^2 - \frac{1}{k} \right),\end{aligned} \tag{20}$$

which, letting $\mathbf{S}_{f_1} = \mathbf{S}_{f_1}^{(n)} := (1/n)\sum_{i,j=1}^{n} \psi_{f_1}\circ\tilde{F}_1^{-1}(R_i/(n+1))\tilde{F}_1^{-1}(R_i/(n+1))\mathbf{U}_i\mathbf{U}_i'$, simplifies to

$$\underset{\sim}{Q}_{f_1} = \frac{nk(k+2)}{2\mathcal{J}_k(f_1)} \left(\mathrm{tr}\,\mathbf{S}_{f_1}^2 - \frac{1}{k}\,\mathrm{tr}^2\mathbf{S}_{f_1} \right) = \frac{k(k+2)}{\mathcal{J}_k(f_1)} \frac{nk^2}{2} \left\| \frac{\mathbf{S}_{f_1}}{\mathrm{tr}\,\mathbf{S}_{f_1}} - \frac{1}{k}\mathbf{I}_k \right\|^2 + o_{\mathrm{P}}(1),$$

as n goes to infinity, still under any elliptical distribution. The van der Waerden (Gaussian scores $f_1 = \phi_1$) test, for instance, is based on the statistic

$$\underset{\sim}{Q}_{\mathrm{vdW}} := \frac{1}{2n} \sum_{i,j=1}^{n} \Psi_k^{-1}\left(\frac{R_i}{n+1}\right) \Psi_k^{-1}\left(\frac{R_j}{n+1}\right) \left((\mathbf{U}_i'\mathbf{U}_j)^2 - \frac{1}{k} \right), \tag{21}$$

where Ψ_k stands for the chi-square distribution function with k degrees of freedom. See (22) for the rank-based test statistics based on Student scores.

In order to describe the asymptotic behaviour of $\underset{\sim}{Q}_K$ and $\underset{\sim}{Q}_{f_1}$, we will need the following quantities:

$$\mathcal{J}_k(K; g_1) := \int_0^1 K(u)\,\psi_{g_1}(\tilde{G}_{1k}^{-1}(u))\,\tilde{G}_{1k}^{-1}(u)\,du, \quad \text{and}$$

$$\mathcal{J}_k(f_1, g_1) := \int_0^1 \psi_{f_1}(\tilde{F}_{1k}^{-1}(u))\,\tilde{F}_{1k}^{-1}(u)\,\psi_{g_1}(\tilde{G}_{1k}^{-1}(u))\,\tilde{G}_{1k}^{-1}(u)\,du.$$

Denote by $\underset{\sim}{\phi}_K^{(n)}$ (resp. by $\underset{\sim}{\phi}_{f_1}^{(n)}$) the rank-based test which consists in rejecting $\mathcal{H}_0 : \mathbf{V} = \mathbf{V}_0$ as soon as $\underset{\sim}{Q}_K^{(n)}$, defined in (18) (resp. $\underset{\sim}{Q}_{f_1}^{(n)}$, defined in (20)) exceeds the α-upper-quantile of a chi-square with $k(k+1)/2 - 1$ degrees of freedom. We now can state the main result of this paper. Note that the unions over g_1 here extend over *all* possible radial densities: contrary to the Gaussian tests described in Section 3.3, where finite fourth-order moments are required, the tests $\underset{\sim}{\phi}_K^{(n)}$ and $\underset{\sim}{\phi}_{f_1}^{(n)}$ are valid without any moment restrictions.

Proposition 4 *Let K be a continuous square-integrable score function defined on $]0,1[$, that can be expressed as the difference of two monotone increasing functions. Similarly, assume that f_1 (satisfying (A1) and (A2)) is such that $K_{f_1}(u) = \psi_{f_1}(\tilde{F}_{1k}^{-1}(u))\,\tilde{F}_{1k}^{-1}(u)$ is continuous and can be expressed as the difference of two monotone increasing functions. Then,*

(i) $\underset{\sim}{Q}{}_K^{(n)}$ *and* $\underset{\sim}{Q}{}_{f_1}^{(n)}$ *are asymptotically chi-square with $k(k+1)/2-1$ degrees of freedom under* $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$, *and asymptotically noncentral chi-square, still with $k(k+1)/2-1$ degrees of freedom but with noncentrality parameters*

$$\frac{\mathcal{J}_k^2(K;g_1)}{2k(k+2)\mathrm{E}[K^2(U)]}\left[\mathrm{tr}\left(\mathbf{V}_0^{-1}\mathbf{v}\right)^2 - \frac{1}{k}\left(\mathrm{tr}\ \mathbf{V}_0^{-1}\mathbf{v}\right)^2\right]$$

and

$$\frac{\mathcal{J}_k^2(f_1,g_1)}{2k(k+2)\mathcal{J}_k(f_1)}\left[\mathrm{tr}\left(\mathbf{V}_0^{-1}\mathbf{v}\right)^2 - \frac{1}{k}\left(\mathrm{tr}\ \mathbf{V}_0^{-1}\mathbf{v}\right)^2\right],$$

respectively, under $\bigcup_{\sigma^2}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0+n^{-1/2}\mathbf{v};g_1}\right\}$;

(ii) the sequences of tests $\underset{\sim}{\phi}{}_K^{(n)}$ *and* $\underset{\sim}{\phi}{}_{f_1}^{(n)}$ *have asymptotic level α under* $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$, *and*

(iii) the sequence of tests $\underset{\sim}{\phi}{}_{f_1}^{(n)}$ *is locally and asymptotically maximin-efficient, still at asymptotic level α, for* $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$ *against alternatives of the form* $\bigcup_{\sigma^2}\bigcup_{\mathbf{V}\neq\mathbf{V}_0}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}\right\}$.

Proof. See Appendix (Section 6.3). □

4.3 Asymptotic relative efficiencies.

Propositions 3 and 4 allow for computing ARE values for $\underset{\sim}{\phi}{}_K^{(n)}$ (hence, for $\underset{\sim}{\phi}{}_{f_1}^{(n)}$) with respect to the adjusted John test $\phi_{\mathcal{N}}^{(n)}$ (therefore, also with respect to the adjusted Mauchly test) as ratios of the noncentrality parameters in the asymptotic distributions of their respective test statistics under local alternatives, for various radial densities g_1. These adjusted tests still are not valid unless $\kappa_k(g_1) < \infty$, and our ARE values therefore also require finite fourth-order moments. Recall however that the signed rank tests $\underset{\sim}{\phi}{}_K^{(n)}$ remain valid without such moment assumption, so that, when g_1 is such that $\kappa_k(g_1) = \infty$, the asymptotic relative efficiency of any $\underset{\sim}{\phi}{}_K^{(n)}$ with respect to $\phi_{\mathcal{N}}^{(n)}$ actually can be considered as being infinite.

Proposition 5 *Let K satisfy the assumptions of Proposition 4. Then, the asymptotic relative efficiency of $\underset{\sim}{\phi}{}_K$ with respect to the parametric Gaussian test $\phi_{\mathcal{N}}$, under radial density g_1 satisfying (A1), (A2), and $\kappa_k(g_1) < \infty$, is*

$$\mathrm{ARE}_{k,g_1}(\underset{\sim}{\phi}{}_K/\phi_{\mathcal{N}}) = \frac{1}{k(k+2)}\,(1+\kappa_k(g_1))\,\frac{\mathcal{J}_k^2(K;g_1)}{\mathrm{E}[K^2(U)]}.$$

For K of the form K_{f_1}, this yields

$$\mathrm{ARE}_{k,g_1}(\underset{\sim}{\phi}{}_{f_1}/\phi_{\mathcal{N}}) = \frac{1}{k(k+2)}\,(1+\kappa_k(g_1))\,\frac{\mathcal{J}_k^2(f_1,g_1)}{\mathcal{J}_k(f_1)}.$$

In order to investigate the numerical values of these AREs, we consider the tests $\underset{\sim}{\phi}_{f_{1,\nu}}$ based on t_ν-scores, i.e., the scores associated with the Student radial densities introduced in Section 1.2. One can easily check that $\psi_{f_{1,\nu}}(r) = (k+\nu)a_{k,\nu}r/(\nu + a_{k,\nu}r^2)$. Also, since $a_{k,\nu}^2\|\mathbf{X}_1\|^2/k$, under $\mathrm{P}^{(n)}_{\mathbf{0},1,\mathbf{I}_k;f_{1,\nu}}$, is Fisher-Snedecor with k and ν degrees of freedom, one can show that the test statistic $\underset{\sim}{Q}_{f_{1,\nu}}$ is

$$\underset{\sim}{Q}_{f_{1,\nu}} = \frac{k^2(k+\nu)(k+\nu+2)}{2n} \sum_{i,j=1}^{n} \frac{T_i^{(n)}}{\nu + k\,T_i^{(n)}} \frac{T_j^{(n)}}{\nu + k\,T_j^{(n)}} \left((\mathbf{U}_i'\mathbf{U}_j)^2 - \frac{1}{k} \right) \tag{22}$$

(see (4)), where, denoting by $G_{k,\nu}$ the Fisher-Snedecor distribution function with k and ν degrees of freedom, we let $T_i^{(n)} := G_{k,\nu}^{-1}(R_i/(n+1))$. Note that the sign test and the van der Waerden test are obtained by letting $\nu \to 0$ and $\nu \to \infty$, respectively. An easy calculation also shows that, for $\nu = 2$, $\underset{\sim}{Q}_{t_\nu}$ and $\underset{\sim}{Q}_{K_a}$ coincide for $a = 2/k$, $k = 2, 3, 4, \ldots$ Hence, for $k = 2$, the Wilcoxon test statistic $\underset{\sim}{Q}_W$ is optimal at Student densities with two degrees of freedom.

Numerical values of the AREs, under various t_ν and normal densities, of several of the proposed rank-based tests with respect to the Gaussian test are given in Table 1. For the sign test $\underset{\sim}{\phi}_S$, closed-form expressions are

$$\mathrm{ARE}_{k,f_{1,\nu}}[\underset{\sim}{\phi}_S/\phi_\mathcal{N}] = \frac{k(\nu-2)}{(k+2)(\nu-4)} \quad \text{and} \quad \mathrm{ARE}_{k,\phi_1}[\underset{\sim}{\phi}_S/\phi_\mathcal{N}] = \frac{k}{k+2}.$$

(recall that $\kappa_k(f_{1,\nu}) < \infty$ iff $\nu > 4$, which is the condition for a Student radial density to satisfy $E_k(f_{1,\nu}) < \infty$). Also, the highest ARE with respect to the Gaussian test $\phi_\mathcal{N}$ that can be achieved is

$$\mathrm{ARE}_{k,f_{1,\nu}}[\underset{\sim}{\phi}_{f_{1,\nu}}/\phi_\mathcal{N}] = \frac{(k+\nu)(\nu-2)}{(k+\nu+2)(\nu-4)}.$$

Note that these numerical AREs are all uniformly good, especially for the van der Waerden test $\underset{\sim}{\phi}_{vdW}$, for which the ARE values are not only uniformly larger than 1, but even uniformly larger than the AREs corresponding to the location problem (e.g., those obtained when testing that the center of symmetry of an elliptical distribution is equal to some fixed k-vector, as in Hallin and Paindaveine (2002)). This (Pitman) dominance of $\underset{\sim}{\phi}_{vdW}$ over $\phi_\mathcal{N}$ also holds under elliptical distributions with ligther-than-Gaussian tails, as can be checked numerically by considering radial densities of the form $g_{1\eta}(r) := \exp(-b_{k,\eta}r^{2\eta})$ ($b_{k,\eta} > 0$ is a scalar determined by Condition (3) again). For instance, in the problem of testing for trivariate sphericity, the corresponding AREs are 1.166, 1.014, 1.000, 1.039, 1.108 and 1.183 for $\eta = .5, .8, 1, 1.5, 2$, and 2.5, respectively.

4.4 Unspecified location $\boldsymbol{\theta}$.

In practice the center of symmetry $\boldsymbol{\theta}$ is seldom specified, and has to be replaced, in test statistics, with an estimator $\hat{\boldsymbol{\theta}}^{(n)}$. Under very mild conditions, any root-n consistent estimator will be adequate (in principle, after due discretization), but we recommend the (rotation-equivariant) spatial median of Möttönen and Oja (1995), which is itself "sign-based".

The asymptotic impact of this substitution on the validity of the signed-rank tests proposed in Section 4.2 could be studied directly (see, e.g., Randles 1982), but is more conveniently handled via Le Cam's third Lemma, which allows for deriving the asymptotic distribution under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$

of the test statistic $\underset{\sim}{Q}_K^{(n)} =: \underset{\sim}{Q}_{K;\boldsymbol{\theta}}^{(n)}$ considered in Section 4.2 but computed at $\hat{\boldsymbol{\theta}}^{(n)}$ instead of $\boldsymbol{\theta}$. This lemma applies in the parametric location experiment $\mathcal{E}_g^{(n)} := \left\{ \mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1} \middle| \boldsymbol{\theta} \in \mathbb{R}^k \right\}$, provided that it is ULAN, which essentially requires that g_1 satisfies Assumption (A1) (see Hallin and Paindaveine 2002).

The asymptotic distribution, as $n \to \infty$, of $\underset{\sim}{Q}_{K;\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)}}^{(n)}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$ for any bounded sequence $\boldsymbol{\tau}^{(n)}$ is the same as under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},\sigma^2,\mathbf{V};g_1}$ (namely, in view of Part *(i)* of Proposition 4, chi-square with $k(k+1)/2 - 1$ degrees of freedom) provided that the asymptotic joint distribution, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$, of $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};K;g_1}$ (defined in (17)) and the central sequence for location $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};g_1;1}$ of $\mathcal{E}_g^{(n)}$ (defined in (6)) is normal with block-diagonal asymptotic covariance. Now, this is automatically satisfied, under the assumptions made on K: indeed, both $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};K;g_1}$ and $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};g_1;1}$ are sums of i.i.d. vectors with finite variances, and, in view of the independence under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$ between $d_i^{(n)}(\boldsymbol{\theta},\mathbf{V})$ and $\mathbf{U}_i^{(n)}(\boldsymbol{\theta},\mathbf{V})$, have a cross-covariance matrix proportional to $\mathrm{E}\left[\mathrm{vec}(\mathbf{U}_i\mathbf{U}_i')\mathbf{U}_i'\right] = \mathbf{0}$. A classical reasoning then extends this to random sequences of the form $\boldsymbol{\tau}^{(n)} = n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta})$, where $n^{1/2}(\hat{\boldsymbol{\theta}}^{(n)} - \boldsymbol{\theta})$ is $O_{\mathrm{P}}(1)$ and $\hat{\boldsymbol{\theta}}^{(n)}$ is *locally discrete*, i.e., such that the number, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$, of its possible values in balls of the form $\{\mathbf{z} \in \mathbb{R}^k \,|\, \|\mathbf{z} - \boldsymbol{\theta}\|^2 \le b^2\}$ remains bounded as $n \to \infty$. It is well known that this latter assumption has no practical consequences (see, e.g., Le Cam and Yang 2000). The null distribution of $\underset{\sim}{Q}_{K;\hat{\boldsymbol{\theta}}^{(n)}}^{(n)}$ is thus the same, then, as that of $\underset{\sim}{Q}_{K;\boldsymbol{\theta}}^{(n)}$.

Le Cam's third Lemma however provides asymptotic equivalence in distribution results, not asymptotic equivalence in probability. Asymptotic equivalence in probability (that is, a result of the form $\underset{\sim}{Q}_{K;\hat{\boldsymbol{\theta}}^{(n)}}^{(n)} - \underset{\sim}{Q}_{K;\boldsymbol{\theta}}^{(n)} = o_{\mathrm{P}}(1)$) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$ requires slightly more stringent *asymptotic linearity* results for $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta};K;g_1}$ (the only exception being the case of $\underset{\sim}{Q}_{g_1;\hat{\boldsymbol{\theta}}^{(n)}}^{(n)}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};g_1}$, which readily follows from ULAN). Proving such results is long and tedious. In order not to overload the paper, we do not elaborate a formal proof here, and either refer to similar proofs given, in a similar context, by Hallin and Paindaveine (2003), or to more general methods such as the one recently developed by Andreou and Werker (2003). The latter still requires $\mathcal{E}_g^{(n)}$ to be ULAN.

Note that $\underset{\sim}{Q}_{K;\hat{\boldsymbol{\theta}}^{(n)}}^{(n)}$ is no longer strictly invariant nor distribution-free, but remains asymptotically so, in the sense of being asymptotically equivalent to its genuinely invariant and distribution-free counterpart $\underset{\sim}{Q}_{K;\boldsymbol{\theta}}^{(n)}$. This asymptotic equivalence carries on to contiguous alternatives, so that local optimality properties also are preserved. Incidentally, note that $\underset{\sim}{Q}_{K;\hat{\boldsymbol{\theta}}^{(n)}}^{(n)}$ is translation-invariant as soon as $\hat{\boldsymbol{\theta}}$ is translation-equivariant.

5 Simulation results.

The asymptotic relative efficiencies of the tests described in Sections 3.3 and 4.2 do not depend on the null value $\mathbf{V}_0$ of the shape matrix. Therefore, in this section, we concentrate on the particular case ($\mathbf{V}_0 = \mathbf{I}_k$) of testing for sphericity. We generated $N = 2,500$ independent samples $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_{500}$ of size $n = 500$ from various bivariate spherical densities (the bivariate normal and bivariate t-distributions with .2, 1, and 6 degrees of freedom, respectively), with symmetry center $\boldsymbol{\theta} = (0,0)'$. From each of these samples, we constructed four series of 500 spherical (for $m = 0$) or elliptical ($m = 1, 2, 3$) observations $\mathbf{X}_1, \ldots, \mathbf{X}_{500}$ characterized by

$$\mathbf{X}_i = (\mathbf{I}_k + m\mathbf{v})\,\boldsymbol{\varepsilon}_i, \quad m = 0, 1, 2, 3, \tag{23}$$

with $\overset{\circ}{\mathrm{vech}}\,\mathbf{v} = (0, .14)'$.

Although designed against elliptical alternatives, our tests also perform quite well under non-elliptical alternatives. In order to show this, we considered the following skew populations: Population $\mathcal{SN}$ refers to samples of 500 observations $\mathbf{X}_1, \ldots, \mathbf{X}_{500}$ characterized by

$$\mathbf{X}_i = (\text{Sign } V_{m;i})\mathbf{W}_{m;i} - \text{E}[(\text{Sign } V_{m;i})\mathbf{W}_{m;i}], \quad m = 0, 1, 2, 3, \tag{24}$$

where the i.i.d. vectors $(V_{m;i}, \mathbf{W}'_{m;i})'$ are drawn from the trivariate standard normal distribution with mean $\mathbf{0}$ and covariance matrix

$$\begin{pmatrix} 1 & \delta' \\ \delta & \mathbf{I}_2 \end{pmatrix}, \quad \delta = (1 + m^2\mathbf{v}'\mathbf{v})^{-1/2} m\mathbf{v},$$

with $\mathbf{v} = (.15, 0)'$. The distribution of the resulting $\mathbf{X}_i$'s is the so-called *bivariate skew normal distribution* with parameters $\mathbf{0}$, $\mathbf{I}_2$, and $m\mathbf{v}$ (see, e.g., Azzalini and Capitanio 1999 or 2003). Population $\mathcal{S}t_2$ is obtained in the same way, but with trivariate t_2-distributed vectors $(V_{m;i}, \mathbf{W}'_{m;i})'$ with the same mean and covariance matrix as in the Gaussian case above, but with $\mathbf{v} = (.25, 0)'$ (see Azzalini and Capitanio 2003).

On each of these samples, we performed the following eleven tests for sphericity (all at asymptotic probability level $\alpha = 5\%$): John's test (based on (15)), the Gaussian test $\phi_{\mathcal{N}}$ (based on (13)), the sign, Wilcoxon, and Spearman tests (based on $\underset{\sim}{Q}_{K_0}$, $\underset{\sim}{Q}_{K_1}$, and $\underset{\sim}{Q}_{K_2}$ in (19), respectively), the van der Waerden test $\underset{\sim}{\phi}_{vdW}$ (based on (21)), and several t_ν-score tests $\underset{\sim}{\phi}_{f_{1,\nu}}$ ($\nu = .2$, $.5$, 1, 2, and 6) (based on (22)). Rejection frequencies are reported in Table 2. The corresponding individual confidence intervals (for $N = 2,500$ replications), at confidence level .95, have half-widths .0044, .0080, and .0100, for frequencies of the order of .05 (.95), .20 (.80), and .50, respectively.

Inspection of Table 2 reveals that the Gaussian test $\phi_{\mathcal{N}}$ collapses under the heavy-tailed distributions $t_{0.2}$ and t_1 (which have infinite fourth-order moments), and confirms the fact that John's test is valid under normal distributions only. All rank-based tests apparently satisfy the 5% probability level constraint. Power rankings are essentially consistent with the corresponding ARE values, which we also report in Table 2. In particular, the asymptotic optimality of $\underset{\sim}{\phi}_{f_{1,\nu}}$ under the Student distribution with ν degrees of freedom is confirmed. The performances under elliptical and non-elliptical alternatives of the various procedures seem to be quite similar.

Finally, in order to investigate the performances of our tests in very small samples, we generated $N = 2,500$ independent samples of size $n = 25$ based on (23) (but with $\overset{\circ}{\mathrm{vech}}\,\mathbf{v} = (0, .2)'$). Only Gaussian and $t_{0.2}$ densities were considered. The corresponding rejection frequencies are reported in Table 3. Similar conclusions as in the first Monte-Carlo study above hold in this small sample simulation. However, note that, for such a small sample size, the asymptotic approximation seems to produce strictly conservative critical values for the van der Waerden- and t_6-score versions of our tests.

6 Appendix.

6.1 Proof of Proposition 1.

Our proof relies on Swensen (1985)'s Lemma 1 (more precisely, on its extension by Garel and Hallin (1995)). The sufficient conditions for LAN given in Swensen's result follow from standard arguments once it is shown that $(\boldsymbol{\theta}, \sigma^2, \mathbf{V}) \mapsto \underline{f}^{1/2}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}(\mathbf{x})$, where $\underline{f}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}$ is the density in (1), is differentiable in quadratic mean, and we therefore focus on this. The main step in establishing this quadratic mean differentiability is the following:

Lemma 3 *Let Assumptions (A1) and (A2) hold. Define* $\underline{g}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) := c_{k,f_1}|\boldsymbol{\Sigma}|^{-1/2} f_1(\|\mathbf{x}-\boldsymbol{\theta}\|_{\boldsymbol{\Sigma}})$, $\mathbf{x} \in \mathbb{R}^k$,

$$D_{\boldsymbol{\theta}}\underline{g}^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) := \frac{1}{2}\, \underline{g}^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x})\ \varphi_{f_1}(\|\mathbf{x}-\boldsymbol{\theta}\|_{\boldsymbol{\Sigma}})\ \boldsymbol{\Sigma}^{-1/2}\mathbf{u}(\boldsymbol{\theta},\boldsymbol{\Sigma}),$$

and

$$D_{\boldsymbol{\Sigma}}\underline{g}^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) := \frac{1}{4}\, \underline{g}^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x})\ \mathbf{P}_k \left(\boldsymbol{\Sigma}^{\otimes 2}\right)^{-1/2} \mathrm{vec}\Big(\psi_{f_1}(\|\mathbf{x}-\boldsymbol{\theta}\|_{\boldsymbol{\Sigma}})\,\|\mathbf{x}-\boldsymbol{\theta}\|_{\boldsymbol{\Sigma}}\,\mathbf{u}(\boldsymbol{\theta},\boldsymbol{\Sigma})\mathbf{u}'(\boldsymbol{\theta},\boldsymbol{\Sigma}) - \mathbf{I}_k\Big),$$

where $\|\mathbf{z}\|_{\boldsymbol{\Sigma}} := (\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z})^{1/2}$, $\mathbf{u}(\boldsymbol{\theta},\boldsymbol{\Sigma}) := \boldsymbol{\Sigma}^{-1/2}(\mathbf{x}-\boldsymbol{\theta})/\|\mathbf{x}-\boldsymbol{\theta}\|_{\boldsymbol{\Sigma}}$, *and* $\mathbf{P}_k$ *is such that* $\mathbf{P}'_k(\mathrm{vech}\,\mathbf{H}) = \mathrm{vec}\,\mathbf{H}$ *for any symmetric* $k \times k$ *matrix* $\mathbf{H} = (H_{ij})$. *Then*

(i) $\int \left\{ g^{1/2}_{\boldsymbol{\theta}+\mathbf{t},\boldsymbol{\Sigma};f_1}(\mathbf{x}) - g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) - \mathbf{t}'\left(D_{\boldsymbol{\theta}} g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x})\right)\right\}^2 d\mathbf{x} = o(\|\mathbf{t}\|^2)$,

(ii) $\int \left\{ g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma}+\mathbf{H};f_1}(\mathbf{x}) - g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) - (\mathrm{vech}\,\mathbf{H})'\left(D_{\boldsymbol{\Sigma}} g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x})\right)\right\}^2 d\mathbf{x} = o(\|\mathbf{H}\|^2)$, *and*

(iii) $\int \left\{ g^{1/2}_{\boldsymbol{\theta}+\mathbf{t},\boldsymbol{\Sigma}+\mathbf{H};f_1}(\mathbf{x}) - g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) - \begin{pmatrix} \mathbf{t} \\ \mathrm{vech}\,\mathbf{H} \end{pmatrix}' \begin{pmatrix} D_{\boldsymbol{\theta}} g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) \\ D_{\boldsymbol{\Sigma}} g^{1/2}_{\boldsymbol{\theta},\boldsymbol{\Sigma};f_1}(\mathbf{x}) \end{pmatrix} \right\}^2 d\mathbf{x} = o\left(\left\| \begin{pmatrix} \mathbf{t} \\ \mathrm{vech}\,\mathbf{H} \end{pmatrix} \right\|^2 \right)$.

To prove Lemma 3, we will need the following reformulation of Assumption (A2).

Lemma 4 *Assumption (A2) holds iff (i)* $f^{1/2}_{1;\exp} \in L^2(\mathbb{R},\nu_k)$ *and (ii) there exists* $Df^{1/2}_{1;\exp} \in L^2(\mathbb{R},\nu_k)$ *such that*

$$\int \left[f^{1/2}_{1;\exp}(x+h) - f^{1/2}_{1;\exp}(x) - h\,(Df^{1/2}_{1;\exp})(x) \right]^2 e^{kh}\, dh = o(h^2).$$

In that case, $Df^{1/2}_{1;\exp}$ *and* $(f^{1/2}_{1;\exp})'$ *are equal in* $L^2(\mathbb{R},\nu_k)$.

The proof of this lemma relies on the following result by Schwartz (see Schwartz (1973), pages 186-188).

Lemma 5 (Schwartz) *The real function* g *is in* $W^{1,2}(\mathbb{R})$ *(with weak derivative* g'*, say) iff (i)* $g \in L^2(\mathbb{R})$ *and (ii) there exists* $Dg \in L^2(\mathbb{R})$ *such that* $x \mapsto g(x+h) - g(x) - h\,(Dg(x))$ *is* $o(h)$ *in* $L^2(\mathbb{R})$ *(as* $h \to 0$*), i.e.,* $\int [g(x+h) - g(x) - h\,(Dg(x))]^2\, dx = o(h^2)$ *as* $h \to 0$. *In that case,* Dg *and* g' *are equal in* $L^2(\mathbb{R})$.

Proof of Lemma 4. Throughout this proof, we write f instead of $f_{1;\exp}^{1/2}$ and all $o(h)$'s are taken as $h \to 0$.

(Necessity) It is easy to show that the real function $x \mapsto g(x) := f(x)\, e^{kx/2}$ admits the weak derivative $x \mapsto g'(x) = f'(x)\, e^{kx/2} + (k/2)\, g(x)$, where f' denotes the weak derivative of f. In view of the assumptions on f, both g and g' are in $L^2(\mathbb{R})$. Lemma 5 therefore yields that $x \mapsto M_h(x) := g(x+h) - g(x) - h\, g'(x)$ is $o(h)$ in $L^2(\mathbb{R})$. But $M_h = I_h + J_h + K_h + L_h$, where

$$I_h(x) := (f(x+h) - f(x) - h\, f'(x))\, e^{kx/2}, \;\; J_h(x) := f(x+h)\, e^{k(x+h)/2} e^{-kh/2}(e^{kh/2} - 1 - hk/2),$$

$$K_h(x) := (f(x+h)\, e^{k(x+h)/2} - f(x)\, e^{kx/2}) hk/2, \;\text{ and }\; L_h(x) := f(x+h)\, e^{k(x+h)/2}(e^{-kh/2} - 1) hk/2.$$

Since J_h, K_h, and L_h also are $o(h)$ in $L^2(\mathbb{R})$, so is I_h.

(Sufficiency) Assume now that $f \in L^2(\mathbb{R}, \nu_k)$ satisfies $x \mapsto I_h(x) := (f(x+h) - f(x) - h\, Df(x))\, e^{kx/2}$ is $o(h)$ in $L^2(\mathbb{R})$ for some $Df \in L^2(\mathbb{R}, \nu_k)$, and define again $x \mapsto g(x) := f(x)\, e^{kx/2}$ ($g \in L^2(\mathbb{R})$). With $Dg(x) := Df(x)\, e^{kx/2} + (k/2)\, g(x)$ ($Dg \in L^2(\mathbb{R})$), we have that

$$x \mapsto \widetilde{M}_h(x) := g(x+h) - g(x) - h\, Dg(x) = (f(x+h) - f(x) - h\, Df(x))\, e^{kx/2} + J_h(x) + K_h(x) + L_h(x)$$

is $o(h)$ in $L^2(\mathbb{R})$. Lemma 5 thus yields that Dg is the weak derivative of g; this implies that, for all infinitely differentiable compactly supported function φ,

$$\int [\varphi(x)\, e^{-kx/2}]\, [Df(x)\, e^{kx/2} + (k/2) g(x)]\, dx = -\int [\varphi'(x)\, e^{-kx/2} - (k/2)\varphi(x)\, e^{-kx/2}]\, [f(x)\, e^{kx/2}]\, dx,$$

i.e., that Df is the weak derivative of f. □

Proof of Lemma 3. *(i)* See Hallin and Paindaveine (2002).

(ii) Using $(\mathbf{C}' \otimes \mathbf{A}) \operatorname{vec} \mathbf{B} = \operatorname{vec}(\mathbf{ABC})$ and letting $\mathbf{y} = \mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\theta})$, the left-hand side of *(ii)* takes the form

$$c_{k,f_1} \int \Big\{ \frac{1}{|\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}|^{1/4}} f_1^{1/2}(\|\mathbf{y}\|_{\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}}) - f_1^{1/2}(\|\mathbf{y}\|) - \frac{1}{4} f_1^{1/2}(\|\mathbf{y}\|)$$
$$\times (\operatorname{vec} \mathbf{H}_{\mathbf{\Sigma}})' \operatorname{vec}\Big(\psi_{f_1}(\|\mathbf{y}\|)\, \frac{\mathbf{y}\mathbf{y}'}{\|\mathbf{y}\|} - \mathbf{I}_k \Big) \Big\}^2 d\mathbf{y} \leq C(T_1 + T_2 + T_3),$$

where $\mathbf{H}_{\mathbf{\Sigma}} := \mathbf{\Sigma}^{-1/2}\mathbf{H}\mathbf{\Sigma}^{-1/2}$,

$$T_1 := \int \Big\{ \frac{1}{|\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}|^{1/4}} - 1 + \frac{1}{4} (\operatorname{vec} \mathbf{H}_{\mathbf{\Sigma}})'(\operatorname{vec} \mathbf{I}_k) \Big\}^2 f_1(\|\mathbf{y}\|_{\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}})\, d\mathbf{y},$$

$$T_2 := \int \frac{1}{16} [(\operatorname{vec} \mathbf{H}_{\mathbf{\Sigma}})'(\operatorname{vec} \mathbf{I}_k)]^2 \Big\{ f_1^{1/2}(\|\mathbf{y}\|_{\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}}) - f_1^{1/2}(\|\mathbf{y}\|) \Big\}^2 d\mathbf{y}, \text{ and}$$

and

$$T_3 := \int \Big\{ f_1^{1/2}(\|\mathbf{y}\|_{\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}}) - f_1^{1/2}(\|\mathbf{y}\|) - \frac{1}{4} f_1^{1/2}(\|y\|)\, (\operatorname{vec} \mathbf{H}_{\mathbf{\Sigma}})' \operatorname{vec}\Big(\psi_{f_1}(\|\mathbf{y}\|)\, \frac{\mathbf{y}\mathbf{y}'}{\|\mathbf{y}\|} \Big) \Big\}^2 d\mathbf{y}.$$

Since $(\operatorname{vec} \mathbf{A})'(\operatorname{vec} \mathbf{B}) = \operatorname{tr}(\mathbf{A}'\mathbf{B})$ and $|\mathbf{A} + \mathbf{B}|^a = |\mathbf{A}|^a + a\, |\mathbf{A}|^a \operatorname{tr}(\mathbf{A}^{-1}\mathbf{B}) + o(\|\mathbf{B}\|)$ for all a (see, e.g., Magnus and Neudecker 1999, page 149),

$$T_1 = \frac{|\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}|^{1/2}}{c_{k,f_1}} \Big\{ |\mathbf{I}_k + \mathbf{H}_{\mathbf{\Sigma}}|^{-1/4} - 1 + \frac{1}{4} (\operatorname{tr} \mathbf{H}_{\mathbf{\Sigma}}) \Big\}^2 = o(\|\mathbf{H}\|^2).$$

Now, working in spherical coordinates $(r, \mathbf{u}) := (\mathbf{y}, \mathbf{y}/\|\mathbf{y}\|)$, we obtain

$$\begin{aligned}
T_3 &= C \iint \left\{ f_1^{1/2}\left(r\|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}\right) - f_1^{1/2}(r) - \frac{1}{4} f_1^{1/2}(r)\, \psi_{f_1}(r)\, r\, \left[\mathbf{u}'\mathbf{H}_{\mathbf{\Sigma}}\mathbf{u}\right] \right\}^2 r^{k-1}\, dr\, d\sigma(\mathbf{u}) \\
&= C \iint \left\{ f_{1;\exp}^{1/2}\left((\ln r) + (\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}})\right) - f_{1;\exp}^{1/2}(\ln r) + (f_{1;\exp}^{1/2})'(\ln r) \left[\frac{1}{2}\mathbf{u}'\mathbf{H}_{\mathbf{\Sigma}}\mathbf{u}\right] \right\}^2 r^{k-1}\, dr\, d\sigma(\mathbf{u}) \\
&= C \iint \left\{ f_{1;\exp}^{1/2}\left(s + (\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}})\right) - f_{1;\exp}^{1/2}(s) + (f_{1;\exp}^{1/2})'(s) \left[\frac{1}{2}\mathbf{u}'\mathbf{H}_{\mathbf{\Sigma}}\mathbf{u}\right] \right\}^2 e^{ks}\, ds\, d\sigma(\mathbf{u}) \\
&\leq C\,(T_{3a} + T_{3b}),
\end{aligned}$$

where

$$T_{3a} := \iint \left\{ f_{1;\exp}^{1/2}\left(s + (\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}})\right) - f_{1;\exp}^{1/2}(s) - (f_{1;\exp}^{1/2})'(s)\left[\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}\right] \right\}^2 e^{ks}\, ds\, d\sigma(\mathbf{u})$$

and

$$T_{3b} := \iint \left\{ \left[\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}\right] + \left[\frac{1}{2}\mathbf{u}'\mathbf{H}_{\mathbf{\Sigma}}\mathbf{u}\right] \right\}^2 \left[(f_{1;\exp}^{1/2})'(s)\right]^2 e^{ks}\, ds\, d\sigma(\mathbf{u}).$$

By using Lemma 4 and the fact that $\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}} = O(\|\mathbf{H}\|)$ for all $\mathbf{u}$, we obtain that

$$\int \left\{ f_{1;\exp}^{1/2}\left(s + (\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}})\right) - f_{1;\exp}^{1/2}(s) - (f_{1;\exp}^{1/2})'(s)\left[\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}\right] \right\}^2 e^{ks}\, ds = o(\|\mathbf{H}\|^2),$$

for all $\mathbf{u}$. Therefore, Lebesgue's dominated convergence theorem entails $T_{3a} = o(\|\mathbf{H}\|^2)$. As for T_{3b}, we have that

$$T_{3b} \leq \sup_{\mathbf{u}\in\mathcal{S}^{k-1}} \left\{ \left[\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}\right] + \left[\frac{1}{2}\mathbf{u}'\mathbf{H}_{\mathbf{\Sigma}}\mathbf{u}\right] \right\}^2 = o(\|\mathbf{H}\|^2),$$

since $[\ln \|\mathbf{u}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}] + [\frac{1}{2}\mathbf{u}'\mathbf{H}_{\mathbf{\Sigma}}\mathbf{u}] = o(\|\mathbf{H}\|)$, uniformly for $\mathbf{u} \in \mathcal{S}^{k-1}$ (see, e.g., Magnus and Neudecker 1999, page 151). Consequently, $T_3 = o(\|\mathbf{H}\|^2)$.

By using the fact that $T_3 = o(1)$ as $\|\mathbf{H}\|$ goes to zero, we obtain

$$\begin{aligned}
T_2 &\leq C\,\|\mathbf{H}_{\mathbf{\Sigma}}\|^2 \int \left\{ f_1^{1/2}\left(\|\mathbf{y}\|_{\mathbf{I}_k+\mathbf{H}_{\mathbf{\Sigma}}}\right) - f_1^{1/2}\left(\|\mathbf{y}\|\right) \right\}^2 d\mathbf{y} \\
&\leq C\,\|\mathbf{H}_{\mathbf{\Sigma}}\|^2 \int \left\{ \frac{1}{4} f_1^{1/2}\left(\|y\|\right) (\mathrm{vec}\,\mathbf{H}_{\mathbf{\Sigma}})' \mathrm{vec}\Big(\psi_{f_1}\left(\|\mathbf{y}\|\right) \frac{\mathbf{y}\mathbf{y}'}{\|\mathbf{y}\|}\Big) \right\}^2 d\mathbf{y} + o(\|\mathbf{H}\|^2),
\end{aligned}$$

which shows that $T_2 = o(\|\mathbf{H}\|^2)$. This proves *(ii)*.

(iii) The left-hand side in *(iii)* is bounded by $C(S_1 + S_2 + S_3 + S_4)$, where

$$S_1 := \int \left\{ g_{\boldsymbol{\theta}+\mathbf{t},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x}) - g_{\boldsymbol{\theta},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x}) - \mathbf{t}'\left(D_{\boldsymbol{\theta}} g_{\boldsymbol{\theta},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x})\right) \right\}^2 d\mathbf{x},$$

$$S_2 := \int \left\{ g_{\boldsymbol{\theta},\mathbf{\Sigma}+\mathbf{H};f_1}^{1/2}(\mathbf{x}) - g_{\boldsymbol{\theta},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x}) - (\mathrm{vech}\,\mathbf{H})'\left(D_{\mathbf{\Sigma}} g_{\boldsymbol{\theta},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x})\right) \right\}^2 d\mathbf{x},$$

and

$$S_3 := \int \left\{ g_{\boldsymbol{\theta}+\mathbf{t},\mathbf{\Sigma}+\mathbf{H};f_1}^{1/2}(\mathbf{x}) - g_{\boldsymbol{\theta}+\mathbf{t},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x}) \right\}^2 d\mathbf{x}, \text{ and } S_4 := \int \left\{ g_{\boldsymbol{\theta},\mathbf{\Sigma}+\mathbf{H};f_1}^{1/2}(\mathbf{x}) - g_{\boldsymbol{\theta},\mathbf{\Sigma};f_1}^{1/2}(\mathbf{x}) \right\}^2 d\mathbf{x}.$$

Now, from *(i)* and *(ii)*, respectively, S_1 and S_2 are $o(\|(\mathbf{t}' \vdots (\mathrm{vech}\,\mathbf{H})')'\|^2)$. As for $S_3 = S_4$, it follows from *(ii)* that it is also $o(\|(\mathbf{t}' \vdots (\mathrm{vech}\,\mathbf{H})')'\|^2)$. The result follows. □

Lemma 6 *Let* $\mathbf{x} \mapsto G_{\boldsymbol{\eta}}(\mathbf{x})$ *be differentiable in quadratic mean at* $\boldsymbol{\eta}_0$*, with gradient* $\mathbf{x} \mapsto DG_{\boldsymbol{\eta}_0}(\mathbf{x})$*, say. Let* h *be a diffeomorphism in a neigbourhood of* $\boldsymbol{\xi}_0 := h^{-1}(\boldsymbol{\eta}_0)$*. Then* $\mathbf{x} \mapsto G_{h(\boldsymbol{\xi})}(\mathbf{x})$ *is differentiable in quadratic mean at* $\boldsymbol{\xi}_0$*, with gradient* $\mathbf{x} \mapsto (Dh_{\boldsymbol{\xi}_0})'(DG_{h(\boldsymbol{\xi}_0)}(\mathbf{x}))$*, where* $Dh_{\boldsymbol{\xi}_0} := (\frac{\partial h_i}{\partial \boldsymbol{\xi}_j}(\boldsymbol{\xi}_0))$ *denotes the Jacobian matrix of* h *at* $\boldsymbol{\xi}_0$*.*

Proof of Lemma 6. Trivial. □

Applied to Lemma 3 *(iii)*, this latter result implies that $\mathbf{x} \mapsto \underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x}) = \underline{f}^{1/2}_{\boldsymbol{\theta},\sigma^2,\mathbf{V};f_1}(\mathbf{x}) = \underline{g}^{1/2}_{\boldsymbol{\theta},\sigma^2\mathbf{V};f_1}(\mathbf{x})$ is differentiable in quadratic mean, with gradient

$$D\underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x}) = \begin{pmatrix} D_{\boldsymbol{\theta}}\underline{g}^{1/2}_{\boldsymbol{\theta},\sigma^2\mathbf{V};f_1}(\mathbf{x}) \\ \begin{pmatrix} 1 & (\overset{\circ}{\mathrm{vech}}\,\mathbf{V})' \\ \mathbf{0} & \sigma^2\mathbf{I} \end{pmatrix} D_{\boldsymbol{\Sigma}}\underline{g}^{1/2}_{\boldsymbol{\theta},\sigma^2\mathbf{V};f_1}(\mathbf{x}) \end{pmatrix} = \frac{1}{2}\,\underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x})\,W_{\boldsymbol{\vartheta};f_1}(\mathbf{x}),$$

where

$$W_{\boldsymbol{\vartheta};f_1}(\mathbf{x}) := \begin{pmatrix} \frac{1}{\sigma}\,\varphi_{f_1}\left(\frac{\|\mathbf{x}-\boldsymbol{\theta}\|_{\mathbf{V}}}{\sigma}\right)\mathbf{V}^{-1/2}\mathbf{u}(\boldsymbol{\theta},\mathbf{V}) \\ \frac{1}{2}\begin{pmatrix} \sigma^{-2}(\mathrm{vec}\,\mathbf{I}_k)' \\ \mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2} \end{pmatrix} \mathrm{vec}\left(\psi_{f_1}\left(\frac{\|\mathbf{x}-\boldsymbol{\theta}\|_{\mathbf{V}}}{\sigma}\right)\frac{\|\mathbf{x}-\boldsymbol{\theta}\|_{\mathbf{V}}}{\sigma}\,\mathbf{u}(\boldsymbol{\theta},\mathbf{V})\mathbf{u}'(\boldsymbol{\theta},\mathbf{V}) - \mathbf{I}_k\right) \end{pmatrix}.$$

Checking Swensen's sufficient conditions for LAN is then a routine task. For example, letting $\nu_i^{(n)} := (\underline{f}^{1/2}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};f_1}(\mathbf{X}_i)/\underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{X}_i)) - 1$ and $Z_i^{(n)} := (1/2)\,(\boldsymbol{\tau}^{(n)})'n^{-1/2}W_{\boldsymbol{\vartheta};f_1}(\mathbf{X}_i)$, $i = 1, \ldots, n$, we have

$$\begin{aligned}\mathrm{E}\left[\sum_{i=1}^n(\nu_i^{(n)} - Z_i^{(n)})^2\right] &= n\int\left\{\underline{f}^{1/2}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};f_1}(\mathbf{x}) - \underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x}) - (1/2)\,(\boldsymbol{\tau}^{(n)})'n^{-1/2}\underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x})W_{\boldsymbol{\vartheta};f_1}(\mathbf{x})\right\}^2 d\mathbf{x} \\ &= n\int\left\{\underline{f}^{1/2}_{\boldsymbol{\vartheta}+n^{-1/2}\boldsymbol{\tau}^{(n)};f_1}(\mathbf{x}) - \underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x}) - (n^{-1/2}\boldsymbol{\tau}^{(n)})'(D\underline{f}^{1/2}_{\boldsymbol{\vartheta};f_1}(\mathbf{x}))\right\}^2 d\mathbf{x},\end{aligned}$$

which is $o(1)$. The other conditions follow easily. Now, the linear term in the LAQ decomposition of the local log-likelihood ratio is $2\sum_{i=1}^n Z_i^{(n)} = (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};f_1}$, where $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\vartheta};f_1}$ is the central sequence announced in Proposition 1. □

6.2 Proofs of Lemma 1 and Proposition 2.

Proof of Lemma 1. Denote by $\mathbf{Q}_k(\mathbf{V})$ the matrix in the right hand side of (12), and characterize $\mathbf{N}_k$ as the $(k(k+1)/2 - 1) \times k^2$ real matrix such that $\mathbf{N}_k(\mathrm{vec}\,\mathbf{v}) = \overset{\circ}{\mathrm{vech}}\,\mathbf{v}$ for any $k \times k$ matrix $\mathbf{v}$. Tedious but routine algebra then yields

$$\mathbf{N}_k\mathbf{Q}_k(\mathbf{V})\mathbf{N}_k' = \left\{\frac{1}{4}\,\mathbf{M}_k\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\left[\mathbf{I}_{k^2} + \mathbf{K}_k - \frac{2}{k}\mathbf{J}_k\right]\left(\mathbf{V}^{\otimes 2}\right)^{-1/2}\mathbf{M}_k'\right\}^{-1}.$$

It is therefore sufficient, in order to prove the lemma, to show that $\mathbf{M}_k'\mathbf{N}_k\mathbf{Q}_k(\mathbf{V}) = \mathbf{Q}_k(\mathbf{V})$. Now, it is easily seen that

$$\mathbf{Q}_k(\mathbf{V}) = \left[\mathbf{I}_{k^2} - (\mathrm{vec}\mathbf{V})\,(\mathbf{e}_{k^2,1})'\right]\left[\mathbf{I}_{k^2} + \mathbf{K}_k\right](\mathbf{V}^{\otimes 2})\left[\mathbf{I}_{k^2} - (\mathrm{vec}\mathbf{V})\,(\mathbf{e}_{k^2,1})'\right]'.$$

But, letting $\mathbf{E}_{ij} := \mathbf{e}_i\mathbf{e}_j' + \mathbf{e}_j\mathbf{e}_i'$ (where $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ stands for the canonical basis of $\mathbb{R}^k$), we have

$$\left[\mathbf{I}_{k^2} - (\mathrm{vec}\mathbf{V})\,(\mathbf{e}_{k^2,1})'\right][\mathbf{I}_{k^2} + \mathbf{K}_k] = \mathbf{I}_{k^2} + \mathbf{K}_k - 2\,(\mathrm{vec}\mathbf{V})\,(\mathbf{e}_{k^2,1})'$$
$$= \frac{1}{2} \sum_{\substack{i,j=1 \\ (i,j)\neq(1,1)}}^{k} (\mathrm{vec}\,\mathbf{E}_{ij})\,(\mathrm{vec}\,\mathbf{E}_{ij})' + 2\left(\mathrm{vec}\,(\mathbf{e}_1\mathbf{e}_1' - \mathbf{V})\right)(\mathbf{e}_{k^2,1})'.$$

The result follows, since $\mathbf{M}_k'\mathbf{N}_k\,(\mathrm{vec}\mathbf{W}) = (\mathrm{vec}\mathbf{W})$ for all symmetric $k \times k$ matrix $\mathbf{W} = (W_{ij})$ such that $W_{11} = 0$ (recall that it is assumed that $\mathbf{V} = (V_{ij})$ be symmetric with $V_{11} = 1$). □

Proof of Proposition 2. Under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$, for any fixed $\boldsymbol{\vartheta}_0' := (\boldsymbol{\theta}', \sigma^2, (\mathrm{v\mathring{e}ch}\,\mathbf{V}_0)')$, we have

$$Q^{(n)}_{f_1} = \left(\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;f_1}\right)' \left(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;f_1}\right)^{-1} \boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;f_1} + o_{\mathrm{P}}(1),$$

as $n \to \infty$. The proof of the first statement in part *(i)* of Proposition 2 follows, since $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;f_1}$ is asymptotically $\mathcal{N}_{k(k+1)/2-1}(\mathbf{0}, \boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;f_1})$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$. On the other hand, it is easy to see that, still under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f_1}$, $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;f_1}$ and the local log-likelihood ratio $\Lambda^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau}/\boldsymbol{\vartheta}_0;f_1}$, where $\boldsymbol{\tau}' := (\mathbf{t}', s, (\mathrm{v\mathring{e}ch}\,\mathbf{v})')$, are jointly multinormal, with asymptotic covariance $\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;f_1}(\mathrm{v\mathring{e}ch}\,\mathbf{v})$. Le Cam's third Lemma thus implies that $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;f_1}$ is asymptotically $\mathcal{N}_{k(k+1)/2-1}(\boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;f_1}(\mathrm{v\mathring{e}ch}\,\mathbf{v}), \boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;f_1})$ under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau};f_1}$, which establishes the second statement in part *(i)* of the lemma.

As for part *(ii)*, the fact that $\phi^{(n)}_{f_1}$ has asymptotic level α directly follows from the asymptotic null distribution given in part *(i)* and the classical Helly-Bray theorem, while local asymptotic maximinity is a consequence of the weak convergence to Gaussian shifts of local shape experiments (see, e.g., Section 11.9 of Le Cam 1986). □

6.3 Proofs of Propositions 3 and 4.

Proof of Proposition 3. Under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;\phi_1}$, for any fixed $\boldsymbol{\vartheta}_0' := (\boldsymbol{\theta}', \sigma^2, (\mathrm{v\mathring{e}ch}\,\mathbf{V}_0)')$, we have

$$Q^{(n)}_{\mathcal{N}} = \left(\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;\phi_1}\right)' \left(E_k(g_1)\,\mathbf{G}_k(\mathbf{V}_0)\right)^{-1} \boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;\phi_1} + o_{\mathrm{P}}(1),$$

as $n \to \infty$, where $\mathbf{G}_k(\mathbf{V}_0)$ was defined in (11). The result then follows—as in Proposition 2—by proving that, under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+n^{-1/2}\boldsymbol{\tau};g_1}$ (with $\boldsymbol{\tau}' := (\mathbf{t}', s, (\mathrm{v\mathring{e}ch}\,\mathbf{v})')$), we have

$$\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;\phi_1} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathrm{E}[\psi_{g_1}(\tilde{G}_1^{-1}(u))(\tilde{G}_1^{-1}(u))^3]\,\mathbf{G}_k(\mathbf{V}_0)(\mathrm{v\mathring{e}ch}\,\mathbf{v}), E_k(g_1)\,\mathbf{G}_k(\mathbf{V}_0)\right)$$

(also note that integration by parts yields $\mathrm{E}[\psi_{g_1}(\tilde{G}_1^{-1}(u))(\tilde{G}_1^{-1}(u))^3] = (k+2)\,D_k(g_1)$). As for the optimality statement in part *(ii)* of the proposition, it is obtained as in the proof of Proposition 2, and by noting that $E_k(\phi_1)\,\mathbf{G}_k(\mathbf{V}_0) = \boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;\phi_1}$. □

Proof of Lemma 2. Let

$$\underset{\sim}{\mathbf{T}}^{(n)}_{\boldsymbol{\vartheta};K} := n^{-1/2}\left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right] \sum_{i=1}^{n} K\left(\frac{R_i}{n+1}\right) \mathrm{vec}\,(\mathbf{U}_i\mathbf{U}_i')$$

and

$$\mathbf{T}^{(n)}_{\boldsymbol{\vartheta};K;f_1} := n^{-1/2}\left[\mathbf{I}_{k^2} - \frac{1}{k}\mathbf{J}_k\right]\sum_{i=1}^{n} K\left(\tilde{F}_{1k}\left(\frac{d_i}{\sigma}\right)\right) \text{vec}\left(\mathbf{U}_i\mathbf{U}_i'\right).$$

Clearly, it is sufficient to prove that $\underset{\sim}{\mathbf{T}}^{(n)}_{\boldsymbol{\vartheta};K} - \mathbf{T}^{(n)}_{\boldsymbol{\vartheta};K;f_1}$ goes to zero in quadratic mean under $\left\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f_1}\right\}$, as $n \to \infty$. For all $l = 1, 2, \ldots, k^2$, we have

$$\mathrm{E}\left[\left(\underset{\sim}{\mathbf{T}}^{(n)}_{\boldsymbol{\vartheta};K} - \mathbf{T}^{(n)}_{\boldsymbol{\vartheta};K;f_1}\right)^2_\ell\right] = C_{\ell,k}\; n^{-1}\sum_{i=1}^{n}\left(K\left(\frac{R_i}{n+1}\right) - K\left(\tilde{F}_{1k}\left(\frac{d_i}{\sigma}\right)\right)\right)^2,$$

where, denoting by $U_{i,j}$ the jth component of $\mathbf{U}_i$, $C_{\ell,k} = \mathrm{Var}[U^2_{1,1}] = 2(k-1)/(k^2(k+2))$ for $\ell \in \mathcal{L}_k := \{mk + m + 1,\; m = 0, 1, \ldots, k-1\}$ and $C_{\ell,k} = \mathrm{Var}\left[U_{1,1}U_{1,2}\right] = 1/k^2$ for $\ell \notin \mathcal{L}_k$. Hájek's classical projection result for signed rank linear statistics (see, e.g., Puri and Sen 1985, Chapter 3) thus yields the desired result. □

Proof of Proposition 4. From Lemma 2, we easily obtain (for any fixed $\boldsymbol{\vartheta}_0' := (\boldsymbol{\theta}', \sigma^2, (\overset{\circ}{\mathrm{vech}}\,\mathbf{V}_0)')$)

$$\underset{\sim}{Q}^{(n)}_K = \left(\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;K}\right)'\left(\mathrm{E}[K^2(U)]\,\mathbf{G}_k(\mathbf{V}_0)\right)^{-1}\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;K} + o_{\mathrm{P}}(1),$$

as $n \to \infty$, under $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$. Part *(i)* of Proposition 4 follows, since

$$\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;K} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathcal{J}_k(K;g_1)\,\mathbf{G}_k(\mathbf{V}_0)(\overset{\circ}{\mathrm{vech}}\,\mathbf{v}), \mathrm{E}[K^2(U)]\,\mathbf{G}_k(\mathbf{V}_0)\right),$$

as $n \to \infty$, under $\bigcup_{\sigma^2}\bigcup_{g_1}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;g_1}\right\}$, with $\boldsymbol{\tau}' := (\mathbf{t}', s, (\overset{\circ}{\mathrm{vech}}\,\mathbf{v})')$. Again Part *(ii)* follows as in the proof of Proposition 2 by noting that the asymptotic variance of $\boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;K_{f_1}} = \boldsymbol{\Delta}^{\star(n)}_{\boldsymbol{\vartheta}_0;f_1}$ under $\bigcup_{\sigma^2}\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\sigma^2,\mathbf{V}_0;f_1}\right\}$ is $\mathcal{J}_k(f_1)\,\mathbf{G}_k(\mathbf{V}) = \boldsymbol{\Gamma}^{\star}_{\boldsymbol{\vartheta}_0;f_1}$. □

References

[1] Andreou, E. and B.J.M. Werker (2003). A simple asymptotic analysis of residual-based statistics. Preprint, CentER, Tilburg University, the Netherlands.

[2] Azzalini, A. and A. Capitanio (1999). Statistical applications of the multivariate skew-normal distribution, *Journal of the Royal Statistical Society Series B* **61**, 579-602.

[3] Azzalini, A. and A. Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution, *Journal of the Royal Statistical Society Series B* **65**, 367-389.

[4] Baringhaus, L. (1991). Testing for spherical symmetry of a multivariate distribution, *Annals of Statistics* **19**, 889-917.

[5] Baringhaus, L. and N. Henze (1992). Limit distributions for Mardia's measure of multivariate skewness, *Annals of Statistics* **20**, 1889-1902.

[6] Beran, R. (1979). Testing for elliptical symmetry of a multivariate density, *Annals of Statistics* **7**, 150-162.

[7] Butler, C.C. (1969). A test for symmetry using the sample distribution function. *Annals of Mathematical Statistics* **40**, 2209-2210.

[8] Efron, B. (1969). Student's t test under symmetry conditions, *Journal of the American Statistical Association* **64**, 1278-1302.

[9] Falk, M. (2002). The sample covariance is not efficient for elliptical distributions, *Journal of Multivariate Analysis* **80**, 358-377.

[10] Garel, B., and M. Hallin (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend, *Annals of the Institute of Statistical Mathematics* **47**, 551-579.

[11] Ghosh, S.K. and D. Sengupta (2001). Testing for proportionality of multivariate dispersion structures using interdirections, *Journal of Nonparametric Statistics* **13**, 331-349.

[12] Hájek, I., Z. Šidák, and P.K. Sen (1999). *Theory of Rank Tests*, 2nd edition, Academic Press, New York.

[13] Hallin, M. and D. Paindaveine (2002). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks, *Annals of Statistics* **30**, 1103-1133.

[14] Hallin, M. and D. Paindaveine (2003). Asymptotic linearity of serial and nonserial multivariate signed rank statistics, submitted.

[15] Hallin, M., and B. J. M. Werker (2003). Semiparametric efficiency, distribution-freeness, and invariance, *Bernoulli* **9**, 55-65.

[16] Hill, B.M. (1969). Foundations for the theory of least squares, *Journal of the Royal Statistical Society Series B* **31**, 89-97.

[17] Huynh, H. and G.K. Mandeville (1979). Validity conditions in repeated- measures designs, *Psychological Bulletin* **86**, 964-973.

[18] John, S. (1971). Some optimal multivariate tests, *Biometrika* **58**, 123-127.

[19] John, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions, *Biometrika* **59**, 169-174.

[20] Kariya, T. and M.L. Eaton (1977). Robust tests for spherical symmetry, *Annals of Statistics* **5**, 206-215.

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

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

[23] Magnus, J.R. and H. Neudecker (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, revised edition, John Wiley, New York.

[24] Malkovich, J.F. and A.A. Afifi (1973). On tests for multivariate normality, *Biometrika* **70**, 713-718.

[25] Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications, *Biometrika* **51**, 519-530.

[26] Mardia, K.V. (1972). *Statistics of Directional Data*, Academic Press, New York.

[27] Mardia, K.V. and P.E. Jupp (1999). *Directional Statistics*, John Wiley, New York.

[28] Mauchly, J.W. (1940). Test for sphericity of a normal n-variate distribution, *Annals of Mathematical Statistics* **11**, 204-209.

[29] Möttönen, J. and H. Oja (1995). Multivariate spatial sign and rank methods, *Journal of Nonparametric Statistics* **5**, 201-213.

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

[31] Oja, H. (1999). Affine invariant multivariate sign and rank tests and the corresponding estimates: a review, *Scandinavian Journal of Statistics* **26**, 319-343.

[32] Ollila, E., C. Croux, and H. Oja (2004). Influence function and asymptotic efficiency of the affine equivariant rank covariance matrix. *Statistica Sinica*, to appear.

[33] Ollila, E., T.P. Hettmansperger, and H. Oja (2004). Affine equivariant multivariate sign methods. Preprint, University of Jyväskylä.

[34] Ollila, E., H. Oja, and C. Croux (2003). The Affine equivariant sign covariance matrix: asymptotic behavior and efficiencies, *Journal of Multivariate Analysis* **87**, 328-355.

[35] Puri, M.L. and P.K. Sen (1985). *Nonparametric Methods in General Linear Models*, John Wiley, New York.

[36] Randles, R. H. (1982). On the asymptotic normality of statistics with estimated parameters, *Annals of Statistics* **10**, 462-474.

[37] Rothman, E.D. and Woodroofe, M. (1972). A Cramér-von Mises type statistic for testing symmetry, *Annals of Mathematical Statistics* **43**, 2035-2038.

[38] Schwartz, L. (1973). *Théorie des Distributions*. Hermann, Paris.

[39] Sugiura, N. (1972). Locally best invariant test for sphericity and the limiting distributions, *Annals of Mathematical Statistics* **43**, 1312-1316.

[40] Swensen, A.R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend, *Journal of Multivariate Analysis* **16**, 54-70.

[41] Tyler, D. E. (1982). Radial estimates and the test for sphericity, *Biometrika* **69**, 429-436.

[42] Tyler, D. E. (1983). Robustness and efficiency properties of scatter matrices, *Biometrika* **70**, 411-420.

[43] Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter, *Annals of Statistics* **15**, 234-251.

Marc Hallin
Département de Mathématique,
I.S.R.O., and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
mhallin@ulb.ac.be

Davy Paindaveine
Département de Mathématique,
I.S.R.O., and E.C.A.R.E.S.
Université Libre de Bruxelles
Campus de la Plaine CP 210
B-1050 Bruxelles Belgium
dpaindav@ulb.ac.be

		degrees of freedom of the underlying t density							
ν	k	1	3	4	5	8	15	20	∞
	2	$+\infty$	$+\infty$	$+\infty$	2.331	1.248	1.045	1.013	0.957
		$(+\infty)$	(2.067)	(1.484)	(1.294)	(1.107)	(1.009)	(0.986)	(0.927)
	3	$+\infty$	$+\infty$	$+\infty$	2.398	1.267	1.052	1.018	0.957
		$(+\infty)$	(2.174)	(1.540)	(1.331)	(1.124)	(1.014)	(0.988)	(0.919)
$\underset{\sim}{\phi}_{t_6}$	4	$+\infty$	$+\infty$	$+\infty$	2.453	1.284	1.058	1.023	0.958
		$(+\infty)$	(2.258)	(1.584)	(1.361)	(1.139)	(1.019)	(0.990)	(0.913)
	6	$+\infty$	$+\infty$	$+\infty$	2.537	1.311	1.070	1.031	0.959
		$(+\infty)$	(2.382)	(1.652)	(1.408)	(1.163)	(1.028)	(0.995)	(0.905)
	10	$+\infty$	$+\infty$	$+\infty$	2.646	1.349	1.087	1.044	0.963
		$(+\infty)$	(2.534)	(1.736)	(1.468)	(1.196)	(1.043)	(1.005)	(0.896)
	2	$+\infty$	$+\infty$	$+\infty$	2.204	1.215	1.047	1.025	1.000
		$(+\infty)$	(1.729)	(1.301)	(1.171)	(1.060)	(1.016)	(1.009)	(1.000)
	3	$+\infty$	$+\infty$	$+\infty$	2.270	1.233	1.052	1.028	1.000
		$(+\infty)$	(1.798)	(1.336)	(1.194)	(1.069)	(1.019)	(1.011)	(1.000)
$\underset{\sim}{\phi}_{vdW}$	4	$+\infty$	$+\infty$	$+\infty$	2.326	1.249	1.057	1.031	1.000
		$(+\infty)$	(1.853)	(1.364)	(1.212)	(1.077)	(1.022)	(1.012)	(1.000)
	6	$+\infty$	$+\infty$	$+\infty$	2.413	1.275	1.066	1.036	1.000
		$(+\infty)$	(1.935)	(1.408)	(1.242)	(1.092)	(1.027)	(1.016)	(1.000)
	10	$+\infty$	$+\infty$	$+\infty$	2.531	1.312	1.080	1.045	1.000
		$(+\infty)$	(2.041)	(1.467)	(1.283)	(1.112)	(1.035)	(1.021)	(1.000)
	2	$+\infty$	$+\infty$	$+\infty$	1.500	0.750	0.591	0.563	0.500
		$(+\infty)$	(2.000)	(1.388)	(1.185)	(0.984)	(0.877)	(0.851)	(0.785)
	3	$+\infty$	$+\infty$	$+\infty$	1.800	0.900	0.709	0.675	0.600
		$(+\infty)$	(2.162)	(1.500)	(1.281)	(1.063)	(0.947)	(0.920)	(0.849)
$\underset{\sim}{\phi}_{S}$	4	$+\infty$	$+\infty$	$+\infty$	2.000	(1.000	0.788	0.750	0.667
		$(+\infty)$	(2.250)	(1.561)	(1.333)	(1.107)	(0.986)	(0.958)	(0.884)
	6	$+\infty$	$+\infty$	$+\infty$	2.250	(1.125	0.886	0.844	0.750
		$(+\infty)$	(2.344)	(1.626)	(1.389)	(1.153)	(1.027)	(0.997)	(0.920)
	10	$+\infty$	$+\infty$	$+\infty$	2.500	1.250	0.985	0.938	0.833
		$(+\infty)$	(2.422)	(1.681)	(1.436)	(1.192)	(1.062)	(1.031)	(0.951)
	2	$+\infty$	$+\infty$	$+\infty$	2.258	1.174	0.956	0.919	0.844
		$(+\infty)$	(1.748)	(1.317)	(1.185)	(1.066)	(1.015)	(1.005)	(0.985)
	3	$+\infty$	$+\infty$	$+\infty$	2.386	1.246	1.022	0.985	0.913
		$(+\infty)$	(1.621)	(1.233)	(1.117)	(1.019)	(0.983)	(0.978)	(0.975)
$\underset{\sim}{\phi}_{W}$	4	$+\infty$	$+\infty$	$+\infty$	2.432	1.273	1.048	1.012	0.945
		$(+\infty)$	(1.533)	(1.171)	(1.064)	(0.979)	(0.954)	(0.952)	(0.961)
	6	$+\infty$	$+\infty$	$+\infty$	2.451	1.283	1.060	1.026	0.969
		$(+\infty)$	(1.422)	(1.090)	(0.994)	(0.921)	(0.908)	(0.911)	(0.938)
	10	$+\infty$	$+\infty$	$+\infty$	2.426	1.264	1.045	1.013	0.970
		$(+\infty)$	(1.315)	(1.007)	(0.919)	(0.855)	(0.851)	(0.857)	(0.907)

Table 1: AREs of the t_6-, van der Waerden-, sign-, and Wilcoxon-score rank-based tests for specified shape and location (in parentheses), with respect to the corresponding parametric Gaussian tests, under k-dimensional Student (1, 3, 4, 5, 8, 15, and 20 degrees of freedom) and normal densities, respectively, for $k = 2, 3, 4, 6$, and 10.

				m						m		
test		0	1	2	3	ARE		0	1	2	3	ARE
ϕ_{John}		.0504	.2380	.6856	.9492	1.000		.9868	.9872	.9848	.9840	ND
$\phi_{\mathcal{N}}$		.0492	.2348	.6824	.9492	1.000		.0060	.0052	.0064	.0088	ND
$\underset{\sim}{\phi}_{vdW}$		.0460	.2208	.6652	.9432	1.000		.0432	.1244	.3620	.6508	ND
$\underset{\sim}{\phi}_{f_{1,6}}$		.0468	.2260	.6644	.9404	0.957		.0456	.1492	.4256	.7376	ND
$\underset{\sim}{\phi}_{f_{1,2}} = \underset{\sim}{\phi}_{W}$	$\mathcal{N}$	.0544	.2052	.6036	0.9028	0.844	t_1	.0480	.1636	.4668	.7936	ND
$\underset{\sim}{\phi}_{f_{1,1}}$		.0544	.1900	.5532	.8600	0.741		.0468	.1632	.4724	.8028	ND
$\underset{\sim}{\phi}_{f_{1,0.5}}$		.0560	.1732	.5000	.8024	0.648		.0460	.1636	.4700	.7964	ND
$\underset{\sim}{\phi}_{f_{1,0.2}}$		.0560	.1628	.4536	.7476	0.568		.0428	.1548	.4404	.7644	ND
$\underset{\sim}{\phi}_{S}$		.0568	.1484	.4016	.6908	0.500		.0452	.1408	.4020	.7064	ND
$\underset{\sim}{\phi}_{SP}$		.0460	.2180	.6576	.9356	0.934		.0488	.1444	.4092	.7240	ND
ϕ_{John}		.1928	.3712	.7016	.9092	ND		.9468	.9460	.9460	.9500	ND
$\phi_{\mathcal{N}}$		.0480	.1580	.4528	.7608	1.000		.0196	.0184	.0252	.0352	ND
$\underset{\sim}{\phi}_{vdW}$		.0428	.1816	.5708	.8800	1.531		.0412	.0924	.2468	.4644	ND
$\underset{\sim}{\phi}_{f_{1,6}}$		.0460	.1956	.5916	.8956	1.600		.0452	.1144	.2996	.5572	ND
$\underset{\sim}{\phi}_{f_{1,2}} = \underset{\sim}{\phi}_{W}$	t_6	.0520	.1904	.5832	.8860	1.531	$t_{0.2}$	.0528	.1284	.3460	.6220	ND
$\underset{\sim}{\phi}_{f_{1,1}}$		.0500	.1836	.5444	.8588	1.408		.0544	.1348	.3760	.6672	ND
$\underset{\sim}{\phi}_{f_{1,0.5}}$		.0464	.1708	.4980	.8148	1.269		.0476	.1356	.3908	.6996	ND
$\underset{\sim}{\phi}_{f_{1,0.2}}$		.0468	.1480	.4432	.7648	1.172		.0500	.1372	.3940	.7016	ND
$\underset{\sim}{\phi}_{S}$		.0488	.1284	.3884	.7064	1.000		.0468	.1296	.3724	.6764	ND
$\underset{\sim}{\phi}_{SP}$		.0480	.1980	.5956	.8888	1.579		.0468	.1056	.2752	.5100	ND
ϕ_{John}		.0520	.0624	.2596	.8000	?		.8640	.8616	.9044	.9520	?
$\phi_{\mathcal{N}}$		.0528	.0664	.2600	.8000	?		.0196	.0188	.0640	.1896	?
$\underset{\sim}{\phi}_{vdW}$		.0472	.0608	.2488	.7828	?		.0536	.0740	.4144	.8504	?
$\underset{\sim}{\phi}_{f_{1,6}}$		.0508	.0620	.2456	.7808	?		.0536	.0724	.4184	.8276	?
$\underset{\sim}{\phi}_{f_{1,2}} = \underset{\sim}{\phi}_{W}$	$\mathcal{SN}$	.0492	.0620	.2304	.7336	?	$\mathcal{S}t_2$	.0512	.0744	.3592	.6964	?
$\underset{\sim}{\phi}_{f_{1,1}}$		.0488	.0608	.2012	.6784	?		.0472	.0724	.2964	.5048	?
$\underset{\sim}{\phi}_{f_{1,0.5}}$		.0476	.0620	.1796	.6112	?		.0484	.0720	.2324	.3280	?
$\underset{\sim}{\phi}_{f_{1,0.2}}$		.0492	.0568	.1568	.5540	?		.0464	.0688	.1744	.2076	?
$\underset{\sim}{\phi}_{S}$		.0512	.0544	.1412	.4972	?		.0468	.0604	.1524	.1556	?
$\underset{\sim}{\phi}_{SP}$		.0528	.0652	.2504	.7752	?		.0552	.0756	.4592	.8820	?

Table 2: Rejection frequencies (out of $N = 2,500$ replications), under various null and non-null distributions (see (23) and (24) for details), of John's test (ϕ_{John}), the Gaussian parametric test ($\phi_{\mathcal{N}}$), and the signed-rank van der Waerden ($\underset{\sim}{\phi}_{vdW}$), t_ν-score ($\underset{\sim}{\phi}_{f_{1,\nu}}$, $\nu = .2, .5, 1, 2, 6$), sign ($\underset{\sim}{\phi}_{S}$), Wilcoxon-type ($\underset{\sim}{\phi}_{W}$), and Spearman-type ($\underset{\sim}{\phi}_{SP}$) tests, respectively; the sample size is 500 ("ND" means "not defined", which occurs as soon as one out of the two tests involved is not valid under the distribution under consideration; "?" means that no theoretical ARE values are available under non elliptical alternatives).

		m						m				
test		0	1	2	3	ARE		0	1	2	3	ARE
ϕ_{John}		.0412	.6032	.9252	.9860	1.000		.8652	.9076	.9360	.9484	ND
$\phi_{\mathcal{N}}$		.0424	.5848	.8924	.9708	1.000		.0004	.0008	.0016	.0020	ND
$\underset{\sim}{\phi}_{vdW}$		.0172	.4136	.8088	.9408	1.000		.0148	.1476	.3608	.5192	ND
$\underset{\sim}{\phi}_{f_{1,6}}$		.0356	.5280	.8684	.9628	0.957		.0308	.2492	.5080	.6844	ND
$\underset{\sim}{\phi}_{f_{1,2}} = \underset{\sim}{\phi}_{W}$	$\mathcal{N}$	.0416	.5400	.8612	.9584	0.844	$t_{0.2}$	.0452	.3288	.6168	.7968	ND
$\underset{\sim}{\phi}_{f_{1,1}}$		.0468	.5036	.8316	.9432	0.741		.0496	.3592	.6784	.8376	ND
$\underset{\sim}{\phi}_{f_{1,0.5}}$		.0496	.4500	.7924	.9132	0.648		.0488	.3824	.7172	.8584	ND
$\underset{\sim}{\phi}_{f_{1,0.2}}$		.0484	.4016	.7328	.8724	0.568		.0508	.3892	.7272	.8692	ND
$\underset{\sim}{\phi}_{S}$		.0480	.3580	.6736	.8216	0.500		.0480	.3752	.7044	.8504	ND
$\underset{\sim}{\phi}_{SP}$		.0396	.5600	.8856	.9696	0.934		.0348	.2320	.4620	.6352	ND

Table 3: Rejection frequencies (still out of $N = 2,500$ replications), under spherical and elliptic Gaussian and $t_{0.2}$ distributions, of the same tests as in Table 2; the sample size is now 25.