

OPTIMAL SIGNED-RANK TESTS BASED ON HYPERPLANES

Hannu Oja and Davy Paindaveine

University of Jyväskylä, Finland,
Université Libre de Bruxelles, Belgium

Abstract

For analysing k -variate data sets, Randles (1989) considered hyperplanes going through $k - 1$ data points and the origin. He then introduced an empirical angular distance between two k -variate data vectors based on the number of hyperplanes (the so-called *interdirections*) that separate these two points, and proposed a multivariate sign test based on those interdirections. In this paper, we present an analogous concept (namely, *lift-interdirections*) to measure the regular distances between data points. The empirical distance between two k -variate data vectors is again determined by the number of hyperplanes that separate these two points; in this case, however, the considered hyperplanes are going through k distinct data points. The invariance and convergence properties of the empirical distances are considered. We show that the lift-interdirections together with Randles' interdirections allow for building hyperplane-based versions of the optimal testing procedures developed in Hallin and Paindaveine (2002a, b, c, and 2004a) for a broad class of location and time series problems. The resulting procedures, which generalize the univariate signed-rank procedures, are affine-invariant and asymptotically invariant under a group of monotone radial transformations (acting on the standardized residuals). Consequently, they are asymptotically distribution-free under the class of elliptical distributions. They are optimal under correctly specified radial densities and, in several cases, enjoy a uniformly good efficiency behavior. These asymptotic properties are confirmed by a Monte-Carlo study, and, finally, a simple robustness study is conducted. It is remarkable that, in the test construction, the value of the test statistic depends on the data cloud only through the geometrical notions of data vectors and oriented hyperplanes, and their relations "above" and "below".

MSC: 62G10; 62G35

Keywords: Multivariate analysis; Nonparametric inference; Hypothesis testing; Interdirections; Lift-interdirections

1 Introduction.

Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be i.i.d. k -vectors with a common unknown symmetry center $\boldsymbol{\theta}$, and consider the one-sample location problem $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. In the case of *central symmetry* ($\mathbf{X}_i - \boldsymbol{\theta} =_d -(\mathbf{X}_i - \boldsymbol{\theta})$, where $=_d$ stands for equality in distribution), multivariate rank-based tests were developed and studied by Möttonen et al. (1995, 1997, and 1998), Hetmansperger et al. (1994, 1997), and Oja (1999). These procedures are either based on *spatial signs and ranks* or on *Oja signs and ranks*. While the spatial procedures are only rotation-invariant, the ones based on Oja signs and ranks are affine-invariant. This latter invariance feature, which is highly

desirable (affine-invariant procedures are coordinate-free), is the main advantage of Oja signs and ranks over the classical *componentwise signs and ranks* (see Puri and Sen 1971).

In the sequel, we will rather consider a more restrictive concept of multivariate symmetry, namely the concept of *elliptical symmetry* (i.e., $\mathbf{X}_i =_{\text{d}} \boldsymbol{\theta} + \boldsymbol{\Sigma}^{1/2} \mathbf{V}_i$, where the \mathbf{V}_i 's are spherically symmetric k -vectors). Denoting by $\mathbf{Z}_i(\boldsymbol{\theta})$ the residual $\mathbf{X}_i - \boldsymbol{\theta}$ associated with observation \mathbf{X}_i and value $\boldsymbol{\theta}$ of the parameter of interest, this means that the collection of residuals $\mathbf{Z}^{(n)}(\boldsymbol{\theta}_0) := (\mathbf{Z}_1(\boldsymbol{\theta}_0), \dots, \mathbf{Z}_n(\boldsymbol{\theta}_0))$ is, under $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$, an *elliptically symmetric white-noise* (see Section 2.1 for a precise definition), that is, essentially, a zero-mean elliptical error process.

The classical (Gaussian) procedure for the multivariate one-sample location problem is Hotelling's T^2 test, which rejects the null hypothesis whenever

$$\frac{n-k}{k(n-1)} T^2 := \frac{n(n-k)}{k(n-1)} (\bar{\mathbf{X}} - \boldsymbol{\theta}_0)^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\theta}_0) > F_{k, n-k; 1-\alpha},$$

where

$$\bar{\mathbf{X}} := \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i, \quad \mathbf{S} := \frac{1}{n-1} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T,$$

and $F_{k, n-k; 1-\alpha}$ stands for the α -upper quantile of a Fisher distribution with k and $(n-k)$ degrees of freedom. Under asymptotic form, the same test rejects the null hypothesis whenever $T^2 > \chi_{k; 1-\alpha}^2$, where $\chi_{k; 1-\alpha}^2$ denotes the α -upper quantile of a chi-square distribution with k degrees of freedom; this latter version is asymptotically valid under any finite-variance distribution (while the exact test requires the underlying distribution to be Gaussian). Hotelling's test is optimal (in its asymptotic version, asymptotically optimal) under Gaussian densities.

Randles (1989) based a multivariate sign test on *interdirections*, i.e., on a hyperplane-based concept of signs (the interdirection $q_{ij}^{(n)}(\boldsymbol{\theta}_0)$ is essentially the number of hyperplanes in \mathbb{R}^k passing through the origin and $(k-1)$ out of the $(n-2)$ points $\mathbf{Z}_1(\boldsymbol{\theta}_0), \dots, \mathbf{Z}_{i-1}(\boldsymbol{\theta}_0), \mathbf{Z}_{i+1}(\boldsymbol{\theta}_0), \dots, \mathbf{Z}_{j-1}(\boldsymbol{\theta}_0), \mathbf{Z}_{j+1}(\boldsymbol{\theta}_0), \dots, \mathbf{Z}_n(\boldsymbol{\theta}_0)$, that are separating $\mathbf{Z}_i(\boldsymbol{\theta}_0)$ and $\mathbf{Z}_j(\boldsymbol{\theta}_0)$; see Section 2.3). His test (ϕ_S , say) consists in rejecting the null hypothesis \mathcal{H}_0 (at asymptotic level α) as soon as

$$Q_S^{(n)} := \frac{k}{n} \sum_{i,j=1}^n \cos(\pi p_{ij}) > \chi_{k; 1-\alpha}^2,$$

where $p_{ij} := p_{ij}^{(n)}(\boldsymbol{\theta}_0) := \binom{n}{k-1}^{-1} q_{ij}^{(n)}(\boldsymbol{\theta}_0)$. This test is valid—that is, it keeps its nominal asymptotic level α —provided that the common distribution of the \mathbf{X}_i 's has elliptical directions (see Randles 1989).

Only directions of the observations \mathbf{X}_i from $\boldsymbol{\theta}_0$ (or equivalently, directions of the residuals $\mathbf{Z}_i(\boldsymbol{\theta}_0)$ from the origin in \mathbb{R}^k) are taken into account in $Q_S^{(n)}$. Information on the radii can also be considered by using the ranks $\hat{R}_i := \hat{R}_i^{(n)}(\boldsymbol{\theta}_0)$ of the Mahalanobis distances $\|\hat{\boldsymbol{\Sigma}}^{-1/2} \mathbf{Z}_i(\boldsymbol{\theta}_0)\|$ (where $\hat{\boldsymbol{\Sigma}}$ is some estimate of the shape matrix $\boldsymbol{\Sigma}$) between the residuals and the origin. This allows for the generalized Wilcoxon-type signed-rank statistic (see Peters and Randles 1990)

$$\hat{Q}_W^{(n)} := \frac{3k}{n^3} \sum_{i,j=1}^n \hat{R}_i \hat{R}_j \cos(\pi p_{ij}).$$

The associated test $\hat{\phi}_W$ (that consists, still at asymptotic level α , in rejecting the null as soon as $\hat{Q}_W^{(n)} > \chi_{k; 1-\alpha}^2$) has higher power than the previous one under light-tailed distributions,

but requires strict ellipticity. By weighting the (pseudo-)Mahalanobis ranks with ad hoc score functions, Hallin and Paindaveine (2002a) developed a class of locally and asymptotically optimal procedures based on (pseudo-)Mahalanobis ranks and interdirections (*pseudo* here means that $\widehat{\Sigma}$ does not have to be the regular covariance matrix estimate, but could be any estimate of Σ satisfying very mild assumptions; see Hallin and Paindaveine (2002a) for details). The K -score version $\widehat{\phi}_K$ of their tests rejects the null when

$$\widehat{Q}_K^{(n)} := \frac{k}{n \mathbb{E}[K^2(U)]} \sum_{i,j=1}^n K\left(\frac{\widehat{R}_i}{n+1}\right) K\left(\frac{\widehat{R}_j}{n+1}\right) \cos(\pi p_{ij}) > \chi_{k;1-\alpha}^2, \quad (1)$$

where U is uniform over $(0, 1)$, and $K: (0, 1) \rightarrow \mathbb{R}$ is some square-integrable score function (see Hallin and Paindaveine (2002a, b) for the technical assumptions on these score functions). When $K = K(g) = \varphi_g \circ \widetilde{G}_k^{-1}$ for some radial density g (we let $\varphi_g := -2(g^{1/2})'/g^{1/2}$, where $(g^{1/2})'$ denotes the weak derivative of $g^{1/2}$; see Section 2.1 for the definition of \widetilde{G}_k), the test $\widehat{\phi}_g$ based on $\widehat{Q}_g^{(n)} := \widehat{Q}_{K(g)}^{(n)}$ is locally and asymptotically optimal under radial density g . Provided that the estimator $\widehat{\Sigma}$ that is used to compute pseudo-Mahalanobis ranks is root- n consistent under arbitrarily heavy tails (such as, e.g., Tyler (1987)'s estimator), the resulting tests remain valid under any elliptical distribution *without any moment assumption* (whereas, typically, standard Mahalanobis ranks—based on the sample covariance matrix—require finite second-order moments). Those tests also often exhibit a uniformly good efficiency behavior. In particular, the van der Waerden version of these tests (based on Gaussian score functions) uniformly dominates the classical Gaussian procedure, namely the Hotelling T^2 test.

In the sequel, we will call *hyperplane-based* statistic (resp., *shape-based* statistic) a statistic that depends on the sample only through the hyperplanes going through the observations (resp., a statistic that relies on the estimation of the underlying shape matrix Σ). With this terminology, one can think of the above signed-rank procedures as *hybrid* ones, since they rely on a hyperplane-based concept of signs (interdirections) and on shape-based ranks (the pseudo-Mahalanobis ranks). Hallin and Paindaveine (2002b) proposed a *purely shape-based* alternative to the above class of hybrid procedures. The associated signed-rank tests consist (at asymptotic level α) in rejecting the null as soon as

$$\frac{k}{n \mathbb{E}[K^2(U)]} \sum_{i,j=1}^n K\left(\frac{\widehat{R}_i}{n+1}\right) K\left(\frac{\widehat{R}_j}{n+1}\right) \mathbf{U}_i^T(\widehat{\Sigma}) \mathbf{U}_j(\widehat{\Sigma}) > \chi_{k;1-\alpha}^2, \quad (2)$$

where $\mathbf{U}_i(\Sigma) := \Sigma^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta}_0) / \|\Sigma^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta}_0)\|$ denotes the spatial sign of the standardized (under the null) observation \mathbf{X}_i , and $\widehat{\Sigma}$ is the estimator of the shape Σ that was used to compute the pseudo-Mahalanobis ranks \widehat{R}_i . Both classes of procedures are asymptotically equivalent and enjoy the same invariance properties. See Hallin and Paindaveine (2002b) for finite-sample comparisons.

The goal of this paper is to propose a *fully hyperplane-based* class of procedures for the one-sample problem, as well as for a much more general class of testing problems. The main step to achieve this is to define a hyperplane-based concept of distances (namely, *lift-interdirections*). We define those new distances, investigate their finite-sample and asymptotic properties, and show that optimal purely hyperplane-based signed-rank procedures can be based on interdirections and on (the ranks of) lift-interdirections. This will be illustrated in the following generic examples: the multivariate one-sample location problem, and the serial problem of testing for multivariate randomness against VARMA dependence.

The paper is organized as follows. In Section 2, we introduce the notation to be used in the sequel, review the concepts of elliptical white-noise and interdirections, and introduce the concepts of lift-interdirections and symmetrized lift-interdirections; we also establish some finite-sample properties of these new statistics. In Section 3, we investigate the asymptotic properties of lift-interdirections and of their ranks. Section 4 shows how lift-interdirections can be used jointly with interdirections to build optimal hyperplane-based testing procedures for a broad class of location and time series problems. In Section 5, two Monte-Carlo studies are conducted in order to investigate, in the one-sample location case, the finite-sample performances and the robustness properties of the proposed hyperplane-based procedures. In Section 6, we state the conclusion and give several final comments. Technical proofs are concentrated in the Appendix.

2 Elliptical symmetry, hyperplane-based angles and distances.

In this section, we introduce the notation to be used (which is essentially the same as in Koshevoy et al. 2003), review the concept of interdirections, and introduce the proposed hyperplane-based distances.

2.1 Elliptical white-noise.

In the sequel, we will typically be interested in semiparametric models of the form

$$F(\mathbf{X}^{(n)}; \boldsymbol{\theta}; \boldsymbol{\varepsilon}^{(n)}) = 0, \quad (3)$$

where $\mathbf{X}^{(n)} = (\mathbf{X}_1, \dots, \mathbf{X}_n) \in \mathbb{R}^{nk}$ is the observation, $\boldsymbol{\theta}$ is the parameter of interest, and $\boldsymbol{\varepsilon}^{(n)} = (\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_n) \in \mathbb{R}^{nk}$ is (a finite realization of) an *elliptical white-noise*.

We will say that $\boldsymbol{\varepsilon}^{(n)}$ is (a finite realization of) an elliptical white-noise if there exist a symmetric positive definite $k \times k$ matrix $\boldsymbol{\Sigma}$ and a positive function $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$, such that the joint probability density of $\boldsymbol{\varepsilon}^{(n)}$ at $(\mathbf{z}_1, \dots, \mathbf{z}_n) \in \mathbb{R}^{nk}$ is of the form $\prod_{i=1}^n \underline{f}(\mathbf{z}_i; \boldsymbol{\Sigma}, f)$, where

$$\underline{f}(\mathbf{z}_i; \boldsymbol{\Sigma}, f) := c_{k,f} \frac{1}{(\det \boldsymbol{\Sigma})^{1/2}} f((\mathbf{z}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{z}_i)^{1/2}), \quad \mathbf{z}_i \in \mathbb{R}^k; \quad (4)$$

we will denote the associated hypothesis by $\mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$. The constant $c_{k,f}$ is the normalization factor $(\omega_k \mu_{k-1;f})^{-1}$, where ω_k stands for the $(k-1)$ -dimensional Lebesgue measure of the unit sphere $\mathcal{S}^{k-1} \subset \mathbb{R}^k$, and $\mu_{\ell;f} := \int_0^\infty r^\ell f(r) dr$. Of course, we require that $\mu_{k-1;f} < \infty$. Under $\mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$, $\|\boldsymbol{\Sigma}^{-1/2} \boldsymbol{\varepsilon}_i\|$ has density $\tilde{f}_k(r) := (\mu_{k-1;f})^{-1} r^{k-1} f(r) \mathbb{I}_{[r>0]}$, where \mathbb{I}_A denotes the indicator function associated with the set A (this justifies that we throughout refer to f as to the *radial density* of the considered elliptical distribution). In the sequel, \tilde{F}_k stands for the distribution function associated with \tilde{f}_k . Similarly, \tilde{G}_k (resp., \tilde{g}_k) denotes the cdf (resp., the pdf) of $\|\boldsymbol{\Sigma}^{-1/2} \boldsymbol{\varepsilon}_i\|$ under $\mathcal{H}^{(n)}(\boldsymbol{\Sigma}, g)$.

We will assume that (3) can be inverted with respect to $\boldsymbol{\varepsilon}^{(n)}$. More precisely, we assume that, for every sample $\mathbf{X}^{(n)}$ and value $\boldsymbol{\theta}$ of the parameter, there is a unique vector of *residuals* $\mathbf{Z}^{(n)}(\boldsymbol{\theta}) = \mathbf{Z}^{(n)}(\mathbf{X}^{(n)}, \boldsymbol{\theta}) = (\mathbf{Z}_1(\mathbf{X}^{(n)}, \boldsymbol{\theta}), \mathbf{Z}_2(\mathbf{X}^{(n)}, \boldsymbol{\theta}), \dots, \mathbf{Z}_n(\mathbf{X}^{(n)}, \boldsymbol{\theta}))$ that satisfies $F(\mathbf{X}^{(n)}; \boldsymbol{\theta}; \mathbf{Z}^{(n)}(\boldsymbol{\theta})) = 0$. Denoting by $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\Sigma}, f)$ the hypothesis under which the observation $\mathbf{X}^{(n)}$ is generated by (3) with value $\boldsymbol{\theta}$ of the parameter of interest, and values $\boldsymbol{\Sigma}$ and f for the white-noise parameters, the distribution of $\mathbf{X}^{(n)}$ lies in $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\Sigma}, f)$ (resp., in $\cup_{\boldsymbol{\Sigma}} \cup_f \mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\Sigma}, f)$) iff that of $\mathbf{Z}^{(n)}(\boldsymbol{\theta})$ belongs to $\mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$ (resp., to $\cup_{\boldsymbol{\Sigma}} \cup_f \mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$).

For instance, in the k -variate one-sample location problem, we have $F(\mathbf{X}^{(n)}; \boldsymbol{\theta}; \boldsymbol{\varepsilon}^{(n)}) = (\mathbf{X}_i - \boldsymbol{\theta} - \boldsymbol{\varepsilon}_i, i = 1, \dots, n)$, and $\mathbf{Z}_i(\mathbf{X}^{(n)}, \boldsymbol{\theta}) = \mathbf{X}_i - \boldsymbol{\theta}$ for all i . In the problem of testing the adequacy of a k -variate VARMA model against other VARMA models, $F(\mathbf{X}^{(n)}; \boldsymbol{\theta}; \boldsymbol{\varepsilon}^{(n)}) = (\mathbf{A}(L)\mathbf{X}_i - \mathbf{B}(L)\boldsymbol{\varepsilon}_i, i = 1, \dots, n)$, where, denoting by L the lag operator, $\mathbf{A}(L) := \mathbf{I}_k - \sum_{r=1}^p \mathbf{A}_r L^r$, $\mathbf{B}(L) := \mathbf{I}_k + \sum_{s=1}^q \mathbf{B}_s L^s$ for some $k \times k$ arrays \mathbf{A}_r and \mathbf{B}_s (\mathbf{I}_k denotes the k -dimensional identity matrix); we let

$$\boldsymbol{\theta} := ((\text{vec}\mathbf{A}_1)^T, \dots, (\text{vec}\mathbf{A}_p)^T, (\text{vec}\mathbf{B}_1)^T, \dots, (\text{vec}\mathbf{B}_q)^T)^T.$$

In that case, $\mathbf{Z}_i(\mathbf{X}^{(n)}, \boldsymbol{\theta}) = \mathbf{H}(L)\mathbf{A}(L)\mathbf{X}_i$ for all i , where $\mathbf{H}(L) := \mathbf{I}_k + \sum_{s=1}^{\infty} \mathbf{H}_s L^s = (\mathbf{B}(L))^{-1}$ is the operator associated with the Green's matrices of operator $\mathbf{B}(L)$. Existence of this inverse operator requires the so-called *invertibility* of the model $\mathbf{A}(L)\mathbf{X}_i = \mathbf{B}(L)\boldsymbol{\varepsilon}_i$.

In view of the above remarks, the elliptical version of the problem of testing $\mathcal{H}_0^{(n)} : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $\mathcal{H}_1^{(n)} : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, i.e.,

$$\begin{cases} \mathcal{H}_0^{(n)} : \bigcup_{\Sigma} \bigcup_f \mathcal{H}_{\boldsymbol{\theta}_0}^{(n)}(\Sigma, f) \\ \mathcal{H}_1^{(n)} : \bigcup_{\boldsymbol{\theta} \neq \boldsymbol{\theta}_0} \bigcup_{\Sigma} \bigcup_f \mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\Sigma, f), \end{cases}$$

is equivalent to that of testing that the distribution of the residuals $\mathbf{Z}^{(n)}(\boldsymbol{\theta}_0)$ lies in $\bigcup_{\Sigma} \bigcup_f \mathcal{H}^{(n)}(\Sigma, f)$. Consequently, we mainly focus on that problem in the sequel.

2.2 Hyperplanes.

Consider some observed k -dimensional series $(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n)$ (typically, the above series of (null) residuals), and write $I := (i_1, i_2, \dots, i_{k-1})$ ($1 \leq i_1 < i_2 < \dots < i_{k-1} \leq n$) and $J := (j_1, j_2, \dots, j_k)$ ($1 \leq j_1 < j_2 < \dots < j_k \leq n$) for ordered sets of indices with size $k-1$ and k , respectively. We will look at I (resp., at J) either as a set of indices or as the set of observations that are indexed by I (resp., by J).

Let $\mathbf{e}(I)$ and $(d_0(J), \mathbf{d}^T(J))^T$ be the vectors whose components are the cofactors of the last column in the arrays

$$(\mathbf{Z}_{i_1} \ \mathbf{Z}_{i_2} \ \dots \ \mathbf{Z}_{i_{k-1}} \ \mathbf{z}) \quad \text{and} \quad \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ \mathbf{Z}_{j_1} & \mathbf{Z}_{j_2} & \dots & \mathbf{Z}_{j_k} & \mathbf{z} \end{pmatrix},$$

respectively (for $k=1$, the only cofactor of some 1×1 matrix is defined as 1). Observations in I point to the hyperplane in \mathbb{R}^k going through those $k-1$ observations along with the origin. Analogously, each J determines the hyperplane going through $\mathbf{Z}_{j_1}, \dots, \mathbf{Z}_{j_{k-1}}$, and \mathbf{Z}_{j_k} . The equations of these hyperplanes are respectively

$$\mathbf{e}^T(I)\mathbf{z} = 0 \quad \text{and} \quad d_0(J) + \mathbf{d}^T(J)\mathbf{z} = 0;$$

so the vectors $\mathbf{e}(I)$ and $\mathbf{d}(J)$ are normal to these hyperplanes. Denoting by $\text{sign}(z)$ the sign function at z (defined as $\text{sign}(z) := \mathbb{I}_{[z>0]} - \mathbb{I}_{[z<0]}$), note that the quantities

$$\text{sign}(\mathbf{e}^T(I)\mathbf{z}) \quad \text{and} \quad \text{sign}(d_0(J) + \mathbf{d}^T(J)\mathbf{z})$$

give a precise meaning to “ \mathbf{z} lies above or below the given hyperplane” (the order of components in I and J determines what is “above” and “below”).

The distances between some point $\mathbf{z} \in \mathbb{R}^k$ and those hyperplanes are respectively given by

$$d(I; \mathbf{z}) := \frac{|\mathbf{e}^T(I)\mathbf{z}|}{\|\mathbf{e}(I)\|} \quad \text{and} \quad d(J; \mathbf{z}) := \frac{|d_0(J) + \mathbf{d}^T(J)\mathbf{z}|}{\|\mathbf{d}(J)\|}. \quad (5)$$

For all $\mathbf{z} \in \mathbb{R}^k$, we can associate with I (resp., with J) the k -simplex whose vertices are $\mathbf{Z}_{i_1}, \mathbf{Z}_{i_2}, \dots, \mathbf{Z}_{i_{k-1}}, \mathbf{z}$, along with the origin (resp., the k -simplex whose vertices are $\mathbf{Z}_{j_1}, \mathbf{Z}_{j_2}, \dots, \mathbf{Z}_{j_k}$, along with \mathbf{z}). The volumes of these simplices are respectively

$$\frac{1}{k!} |\mathbf{e}^T(I)\mathbf{z}| \quad \text{and} \quad \frac{1}{k!} |d_0(J) + \mathbf{d}^T(J)\mathbf{z}|.$$

Finally, I (resp., J) also identifies the $(k-1)$ -subsimplex with vertices $\mathbf{Z}_{i_1}, \mathbf{Z}_{i_2}, \dots, \mathbf{Z}_{i_{k-1}}$, along with the origin (resp., the $(k-1)$ -subsimplex with vertices $\mathbf{Z}_{j_1}, \mathbf{Z}_{j_2}, \dots, \mathbf{Z}_{j_k}$). It can be shown that the “area” of these subsimplices—that is, their $(k-1)$ -dimensional Lebesgue measure—are $\|\mathbf{e}(I)\|/(k-1)!$ and $\|\mathbf{d}(J)\|/(k-1)!$, respectively. So $\mathbf{e}(I)$ and $\mathbf{d}(J)$ are normal to subsimplices considered and their norms are proportional to the volumes of those subsimplices.

2.3 Interdirections.

Hyperplane-based empirical measures of angular distances between k -vectors \mathbf{u} and \mathbf{v} , within some observed k -dimensional series $(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n)$, can be defined as

$$q^{(n)}(\mathbf{u}, \mathbf{v}) = \sum_I \frac{1 - \text{sign}(\mathbf{e}^T(I)\mathbf{u}) \text{sign}(\mathbf{e}^T(I)\mathbf{v})}{2}$$

(see Randles 1989), that is, as the number of hyperplanes in \mathbb{R}^k passing through the origin and $(k-1)$ out of the n points $\mathbf{Z}_1, \dots, \mathbf{Z}_n$, that are separating \mathbf{u} and \mathbf{v} .

The interdirection $q_{ij}^{(n)}$ associated with the pair of observations $(\mathbf{Z}_i, \mathbf{Z}_j)$ in the k -variate sample $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ is defined as

$$q_{ij}^{(n)} = q^{(n)}(\mathbf{Z}_i, \mathbf{Z}_j),$$

that is, up to a small-sample correction, as the number of hyperplanes in \mathbb{R}^k passing through the origin and $(k-1)$ out of the $(n-2)$ points $\mathbf{Z}_1, \dots, \mathbf{Z}_{i-1}, \mathbf{Z}_{i+1}, \dots, \mathbf{Z}_{j-1}, \mathbf{Z}_{j+1}, \dots, \mathbf{Z}_n$, that are separating \mathbf{Z}_i and \mathbf{Z}_j . Interdirections provide affine-invariant estimations of the Euclidean angles between the standardized observations $\Sigma^{-1/2}\mathbf{Z}_i$, that is, they estimate the scalar products of the corresponding spatial signs. More precisely, one can show the following result (see Hallin and Paindaveine (2002a) for a proof based on U-statistics):

Lemma 1 Under $\cup_f \mathcal{H}^{(n)}(\Sigma, f)$,

$$\binom{n}{k-1}^{-1} q_{ij}^{(n)} = \pi^{-1} \arccos \left(\mathbf{U}_i^T(\Sigma) \mathbf{U}_j(\Sigma) \right) + o_P(1),$$

as $n \rightarrow \infty$, where we let $\mathbf{U}_i(\Sigma) := \Sigma^{-1/2}\mathbf{Z}_i / \|\Sigma^{-1/2}\mathbf{Z}_i\|$ denote the spatial sign of $\Sigma^{-1/2}\mathbf{Z}_i$.

2.4 Lift-interdirections.

The proposed hyperplane-based concept of distances is the following. The empirical distance between k -vectors \mathbf{u} and \mathbf{v} , in some observed k -dimensional series $(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n)$, is defined as

$$l^{(n)}(\mathbf{u}, \mathbf{v}) := \sum_J \frac{1 - \text{sign}(d_0(J) + \mathbf{d}^T(J)\mathbf{u}) \text{sign}(d_0(J) + \mathbf{d}^T(J)\mathbf{v})}{2}, \quad (6)$$

that is, as the number of hyperplanes in \mathbb{R}^k passing through k out of the n points $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$, that are separating \mathbf{u} and \mathbf{v} . The distance was first mentioned (but not analysed) in Hettmansperger et al. (1998). We speak throughout of distances even though, for fixed sample size n , $l^{(n)}$ is not a distance measure in the mathematical sense (in particular, $l^{(n)}(\mathbf{u}, \mathbf{v}) = 0$ does not imply that $\mathbf{u} = \mathbf{v}$; note, however, that $l^{(n)}$ obeys the triangular inequality). Nevertheless, as we will see in the sequel, as n goes to infinity, the “distance measure” $l^{(n)}$ (properly normalized) converges to a *genuine* distance over \mathbb{R}^k .

It will be sufficient for our purposes to consider the particular case

$$l^{(n)}(\mathbf{v}) := l^{(n)}(\mathbf{v}, -\mathbf{v}), \quad (7)$$

which is a measure of the distance between \mathbf{v} and $-\mathbf{v}$, or equivalently, between \mathbf{v} and the origin in \mathbb{R}^k . Since the number of arguments of $l^{(n)}$ removes any ambiguity, we use the same notation for the general distance in (6) and the distance from the origin in (7); we will call the latter the *empirical lift-interdirection function*. Note that, in the univariate case, $l^{(n)}(v)$ coincides almost everywhere with the absolute rank function $v \mapsto \sum_{j=1}^n \mathbb{I}_{[|v| \geq |Z_j|]}$.

Quite analogously to interdirections, we define the *lift-interdirection* associated with observation \mathbf{Z}_i , in the k -variate sample $\mathbf{Z}_1, \dots, \mathbf{Z}_n$, as

$$l_i^{(n)} := l^{(n)}(\mathbf{Z}_i),$$

that is, up to a small-sample correction, as the number of hyperplanes in \mathbb{R}^k passing through k out of the $(n - 1)$ points $\mathbf{Z}_1, \dots, \mathbf{Z}_{i-1}, \mathbf{Z}_{i+1}, \dots, \mathbf{Z}_n$, that are separating \mathbf{Z}_i and $-\mathbf{Z}_i$. The terminology is justified by the following remark. Define by $\mathbf{Z}_{l;i} := (1, \mathbf{Z}_i^T)^T$ the *lifted observation* associated with observation \mathbf{Z}_i . Then the lift-interdirection $l_i^{(n)}$ is the (usual) interdirection associated with the pair $((1, \mathbf{Z}_i^T)^T, (1, -\mathbf{Z}_i^T)^T)$ in the lifted sample $\mathbf{Z}_{l;1}, \dots, \mathbf{Z}_{l;n}$.

Now, denote by $(d_{0\mathbf{M}}(J), \mathbf{d}_{\mathbf{M}}^T(J))^T$ the vector $(d_0(J), \mathbf{d}^T(J))^T$ computed from the transformed sample $\mathbf{M}\mathbf{Z}_1, \mathbf{M}\mathbf{Z}_2, \dots, \mathbf{M}\mathbf{Z}_n$ for some arbitrary non-singular $k \times k$ matrix \mathbf{M} . Then, for all k -vector \mathbf{v} ,

$$\begin{aligned} d_{0\mathbf{M}}(J) + \mathbf{d}_{\mathbf{M}}^T(J)(\mathbf{M}\mathbf{v}) &= \det \begin{pmatrix} 1 & \dots & 1 & 1 \\ \mathbf{M}\mathbf{Z}_{j_1} & \dots & \mathbf{M}\mathbf{Z}_{j_k} & \mathbf{M}\mathbf{v} \end{pmatrix} \\ &= \det \begin{pmatrix} 0 & \dots & 0 & 1 \\ \mathbf{M}(\mathbf{Z}_{j_1} - \mathbf{v}) & \dots & \mathbf{M}(\mathbf{Z}_{j_k} - \mathbf{v}) & \mathbf{M}\mathbf{v} \end{pmatrix} \\ &= (\det \mathbf{M}) \det \begin{pmatrix} 0 & \dots & 0 & 1 \\ \mathbf{Z}_{j_1} - \mathbf{v} & \dots & \mathbf{Z}_{j_k} - \mathbf{v} & \mathbf{v} \end{pmatrix} \\ &= (\det \mathbf{M})(d_0(J) + \mathbf{d}^T(J)\mathbf{v}), \end{aligned}$$

so that $d_{0\mathbf{M}}(J) = (\det \mathbf{M}) d_0(J)$ and

$$\mathbf{d}_{\mathbf{M}}(J) = (\det \mathbf{M})(\mathbf{M}^{-1})^T \mathbf{d}(J). \quad (8)$$

This shows that $\text{sign}(d_{0\mathbf{M}}(J) \pm \mathbf{d}_{\mathbf{M}}^T(J)(\mathbf{M}\mathbf{v})) = \text{sign}(\det \mathbf{M}) \text{sign}(d_0(J) \pm \mathbf{d}^T(J)\mathbf{v})$, which proves that lift-interdirections are affine-invariant.

2.5 Symmetrized lift-interdirections.

As we will see in Section 4, the lift-interdirections introduced in the previous subsection suffer from a certain lack of symmetry. That is why we need to consider the *symmetrized empirical lift-interdirection function*, which we define as

$$\underline{l}^{(n)}(\mathbf{v}) := \sum_J \sum_{\mathbf{s}} \frac{1 - \text{sign}(d_{0\mathbf{s}}(J) + \mathbf{d}_{\mathbf{s}}^T(J)\mathbf{v}) \text{sign}(d_{0\mathbf{s}}(J) - \mathbf{d}_{\mathbf{s}}^T(J)\mathbf{v})}{2},$$

where, for some $J = (j_1, \dots, j_k)$ and some $\mathbf{s} \in \{-1, 1\}^k$ (where $\{-1, 1\}^k$ denotes the set of the 2^k k -vectors with entries 1 or -1), $(d_{0\mathbf{s}}(J), \mathbf{d}_{\mathbf{s}}^T(J))^T$ stands for the vector of cofactors of the last column in the array

$$\begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ s_1 \mathbf{Z}_{j_1} & s_2 \mathbf{Z}_{j_2} & \dots & s_k \mathbf{Z}_{j_k} & \mathbf{z} \end{pmatrix}.$$

As its non-symmetrized counterpart, the function $\underline{l}^{(n)}(\mathbf{v})$ is symmetric under permutations of the observations; but unlike $l^{(n)}(\mathbf{v})$, it is also invariant under reflections of the observations with respect to the origin in \mathbb{R}^k (which of course justifies the terminology). Consequently, the *symmetrized lift-interdirections*

$$\underline{l}_i^{(n)} := \underline{l}^{(n)}(\mathbf{Z}_i), \quad i = 1, \dots, n,$$

are invariant under permutations of the observations, as well as under reflections of the observations with respect to the origin in \mathbb{R}^k . Proceeding as above, one can also show that the symmetrized lift-interdirections are affine-invariant.

3 Asymptotics of lift-interdirections.

Since $\binom{n}{k}^{-1} l^{(n)}(\mathbf{v})$ is a U-statistic with a bounded kernel, it does converge in quadratic mean, as $n \rightarrow \infty$, to the expectation of this kernel, i.e., to the *theoretical lift-interdirection function*

$$l_{\Sigma, f}(\mathbf{v}) := \mathbb{E}_{\Sigma, f} \left[\binom{n}{k}^{-1} l^{(n)}(\mathbf{v}) \right] = (1/2) \mathbb{E}_{\Sigma, f} \left[1 - \text{sign}(d_0(J) + \mathbf{d}^T(J)\mathbf{v}) \text{sign}(d_0(J) - \mathbf{d}^T(J)\mathbf{v}) \right],$$

where the expectation $\mathbb{E}_{\Sigma, f}$ is taken under $\mathcal{H}^{(n)}(\Sigma, f)$. In other words, $\binom{n}{k}^{-1} l^{(n)}(\mathbf{v})$ converges (in quadratic mean, as $n \rightarrow \infty$) to the probability that a random hyperplane J separates \mathbf{v} and $-\mathbf{v}$. Note that, in this elliptic setup, $1 - l_{\Sigma, f}(\mathbf{v})$ gives the so-called “majority depth” of \mathbf{v} (see Zuo and Serfling (2000); the original definition was given by Singh (1991)).

As soon as the underlying distribution is symmetric in the sense that $\mathbf{Z}_i \stackrel{d}{=} -\mathbf{Z}_i$ for all i , the theoretical counterparts of the empirical symmetrized and non-symmetrized lift-interdirection functions do clearly coincide, i.e., $\underline{l}^{(n)}(\mathbf{v})/[2^k \binom{n}{k}]$ and $l^{(n)}(\mathbf{v})/\binom{n}{k}$ share the same expectation. In particular, this holds for elliptically symmetric distributions. Accordingly, under $\bigcup_{\Sigma} \bigcup_f \mathcal{H}^{(n)}(\Sigma, f)$, $\underline{l}^{(n)}(\mathbf{v})/[2^k \binom{n}{k}] = l^{(n)}(\mathbf{v})/\binom{n}{k} + o_{L^2}(1)$, as $n \rightarrow \infty$.

Since affine-invariance implies that $l_{\Sigma, f}(\mathbf{v}) = l_{\mathbf{I}_k, f}(\Sigma^{-1/2}\mathbf{v})$, it is sufficient to study the function $l_f(\mathbf{v}) := l_{\mathbf{I}_k, f}(\mathbf{v})$. Clearly, $\mathbf{v} \mapsto l_f(\mathbf{v})$ is spherically symmetric and is a monotone increasing function in $\|\mathbf{v}\|$. The following proposition characterizes further that monotone function and allows for an intuitive interpretation of the theoretical lift-interdirection function.

Proposition 1 Let \mathbf{Z} be the $k \times k$ array $(\mathbf{Z}_1, \dots, \mathbf{Z}_k)$, where the \mathbf{Z}_i 's are i.i.d. spherically symmetric with radial density f . Then the theoretical lift-interdirection function l_f at \mathbf{v} is given by

$$l_f(\mathbf{v}) = \mathbb{P} \left[(\mathbf{1}^T (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{1})^{-1/2} \leq |\mathbf{U}^T \mathbf{v}| \right], \quad (9)$$

where $\mathbf{1} := (1, \dots, 1)^T \in \mathbb{R}^k$ and \mathbf{U} is a k -vector (independent of $\mathbf{1}^T (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{1}$) that is uniformly distributed over the unit sphere $\mathcal{S}^{k-1} \subset \mathbb{R}^k$.

See the Appendix for the proof. Note that, in the univariate case, this just yields $l_f(v) = P[|Z_1| \leq |v|] = \tilde{F}_1(|v|)$, where \tilde{F}_1 denotes the cumulative distribution function of $|Z_1|$ (see the notation in Section 2.1).

Let $J = (1, 2, \dots, k)$. As shown in the proof of Proposition 1, $(\mathbf{1}^T (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{1})^{-1/2}$ is equal to the distance $d(J; \mathbf{0}) = |d_0(J)| / \|\mathbf{d}(J)\|$ between hyperplane J and the origin in \mathbb{R}^k ; see (5) (in particular, this gives an expression of this distance as a function of the scalar products $\mathbf{Z}_i^T \mathbf{Z}_j$ only). The proof of Proposition 1 also shows that the vector \mathbf{U} in (9) is nothing but $\mathbf{u}(J) := \mathbf{d}(J) / \|\mathbf{d}(J)\|$, which gives the normal direction to hyperplane J . The unit normal $\mathbf{u}(J)$ and the distance $d(J; \mathbf{0})$ are independent in the spherical case, and the probability that hyperplane J separates \mathbf{v} and $-\mathbf{v}$ is thus

$$\mathbb{P} \left[d(J; \mathbf{0}) \leq |\mathbf{u}^T(J) \mathbf{v}| \right].$$

This allows for the following simple interpretations.

- (i) $\mathbb{P}[d(J; \mathbf{0}) \leq |\mathbf{u}^T(J) \mathbf{v}| \mid |\mathbf{u}^T(J) \mathbf{v}| = \rho \|\mathbf{v}\|]$ is a monotone increasing function of $\rho \in (0, 1)$, which means that this conditional probability increases as the normal direction to hyperplane J gets closer to the direction of \mathbf{v} , and
- (ii) $\mathbb{P} \left[d(J; \mathbf{0}) \leq |\mathbf{u}^T(J) \mathbf{v}| \mid d(J; \mathbf{0}) = d \right]$ is a monotone decreasing function of d , which means that this conditional probability decreases as the hyperplane J moves away from the origin in \mathbb{R}^k .

This is illustrated in Figures 1 and 2, respectively.

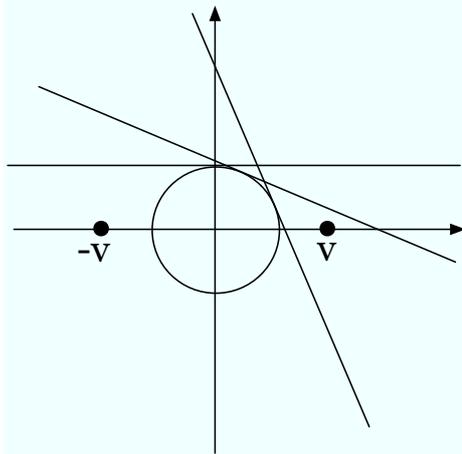


Figure 1: Three hyperplanes J with same distances $d(J; \mathbf{0})$ but different unit normals $\mathbf{u}(J)$.

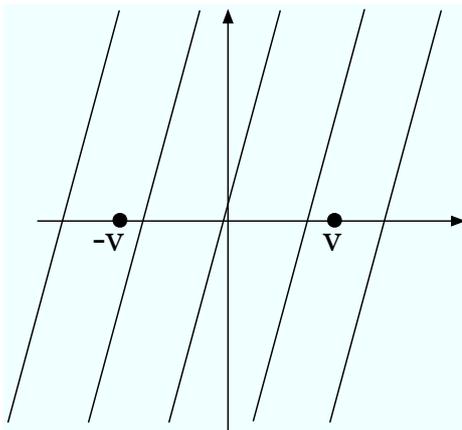


Figure 2: Five hyperplanes J with identical unit normals $\mathbf{u}(J)$ but different distances $d(J; \mathbf{0})$.

We strongly believe that the expression for $l_f(\mathbf{v})$ in Proposition 1 cannot be improved in the general case. However, in the Gaussian case ($f(r) = \phi(r) := \exp(-r^2/2)$), $\mathbf{Z}^T \mathbf{Z} =_d \mathbf{Z} \mathbf{Z}^T = \sum_{i=1}^k \mathbf{Z}_i \mathbf{Z}_i^T$, where the \mathbf{Z}_i 's are i.i.d. $\mathcal{N}_k(\mathbf{0}, \mathbf{I}_k)$. Consequently, $\mathbf{Z}^T \mathbf{Z} \sim W_k(k, \mathbf{I}_k)$, where $W_k(k, \mathbf{I}_k)$ denotes the k -variate standard Wishart distribution with k degrees of freedom. If $\mathbf{W} \sim W_k(k, \mathbf{I}_k)$ and if \mathbf{t} is a k -vector such that $\|\mathbf{t}\| = 1$, we have that

$$\frac{1}{\mathbf{t}^T \mathbf{W}^{-1} \mathbf{t}} \sim \chi_1^2$$

(see, e.g., Bilodeau and Brenner 1999, page 96). This yields $k(\mathbf{1}^T (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{1})^{-1} \sim \chi_1^2$ in the Gaussian case. On the other hand, denoting by \mathbf{e}_1 the first vector of the canonical basis of \mathbb{R}^k , we get, if \mathbf{U} is uniformly distributed over the unit sphere in \mathbb{R}^k , $(\mathbf{U}^T \mathbf{v})^2 = \|\mathbf{v}\|^2 (\mathbf{U}^T \mathbf{v} / \|\mathbf{v}\|)^2 =_d \|\mathbf{v}\|^2 (\mathbf{U}^T \mathbf{e}_1)^2$, where $(\mathbf{U}^T \mathbf{e}_1)^2 \sim \text{beta}(1/2, (k-1)/2)$ for $k > 1$, and $(\mathbf{U}^T \mathbf{e}_1)^2 = 1$ a.s. for $k = 1$. Consequently, in the (spherically symmetric) multinormal case, the theoretical lift-interdirection function is

$$\mathbf{v} \mapsto l_\phi(\mathbf{v}) = P \left[Y \leq k \|\mathbf{v}\|^2 Z \right], \quad (10)$$

where $Y \sim \chi_1^2$ and $Z \sim \text{beta}(1/2, (k-1)/2)$ are independent.

Interestingly, this Gaussian result can be derived in another more direct way. Consider again $\mathbf{Z}_1, \dots, \mathbf{Z}_k$ i.i.d. with k -variate standard normal distribution. Since $l_\phi(\mathbf{v})$ is spherically symmetric, we will only compute $l_\phi(\|\mathbf{v}\| \mathbf{e}_1)$. Using that $\mathbf{Z} =_d \mathbf{Z}^T$ in the Gaussian case, one obtains

$$d_0(J) + \mathbf{d}^T(J)(\|\mathbf{v}\| \mathbf{e}_1) = \det \begin{pmatrix} \mathbf{1}^T & 1 \\ \mathbf{Z} & \|\mathbf{v}\| \mathbf{e}_1 \end{pmatrix} =_d \det \begin{pmatrix} \mathbf{1}^T & 1 \\ \mathbf{Z}^T & \|\mathbf{v}\| \mathbf{e}_1 \end{pmatrix},$$

where $J = (1, \dots, k)$. Denote by $[\mathbf{Z}^T]_1$ the array obtained by deleting the first row in \mathbf{Z}^T , and

fix $I = (2, \dots, k)$. Developing the last determinant along its $(k + 1)$ th column, we then obtain

$$\begin{aligned}
d_0(J) + \mathbf{d}^T(J)(\|\mathbf{v}\|\mathbf{e}_1) &=_{\text{d}} (-1)^k \left[(\det \mathbf{Z}) - \|\mathbf{v}\| \det \begin{pmatrix} \mathbf{1}^T \\ [\mathbf{Z}^T]_1 \end{pmatrix} \right] \\
&= (-1)^k \left[(\det \mathbf{Z}) - \|\mathbf{v}\| \det(\mathbf{1}, \mathbf{Z}_2, \dots, \mathbf{Z}_k) \right] \\
&= -\mathbf{e}^T(I)\mathbf{Z}_1 + \|\mathbf{v}\| \mathbf{e}^T(I)\mathbf{1} \\
&= \|\mathbf{e}(I)\| \left(\mathbf{u}^T(I)(\|\mathbf{v}\|\mathbf{1} - \mathbf{Z}_1) \right),
\end{aligned}$$

where $\mathbf{u}(I) := \mathbf{e}(I)/\|\mathbf{e}(I)\|$ is uniformly distributed over the unit sphere in \mathbb{R}^k (see e.g., the proof of Lemma 1 in Hallin and Paindaveine (2002a)). We deduce that $\text{sign}(d_0(J) + \mathbf{d}^T(J)(\|\mathbf{v}\|\mathbf{e}_1)) =_{\text{d}} \text{sign}(\mathbf{u}^T(I)(\|\mathbf{v}\|\mathbf{1} - \mathbf{Z}_1))$ and, similarly, that $\text{sign}(d_0(J) + \mathbf{d}^T(J)(-\|\mathbf{v}\|\mathbf{e}_1)) =_{\text{d}} \text{sign}(\mathbf{u}^T(I)(-\|\mathbf{v}\|\mathbf{1} - \mathbf{Z}_1))$, with the same vector $\mathbf{u}(I)$. Consequently,

$$l_\phi(\|\mathbf{v}\|\mathbf{e}_1) = \text{P} \left[|\mathbf{u}^T(I)\mathbf{Z}_1| \leq \|\mathbf{v}\| |\mathbf{u}^T(I)\mathbf{1}| \right] = \text{P} \left[(\mathbf{Z}_1^T \mathbf{e}_1)^2 \leq k \|\mathbf{v}\|^2 (\mathbf{u}^T(I)\mathbf{e}_1)^2 \right],$$

which coincides with the result in (10). \square

This additional derivation stresses the very particular role played by the multivariate normal distribution in multivariate analysis (the fact that this second derivation *only* holds in the Gaussian case is directly related to the Maxwell-Herschel Theorem, which states that the (spherical) multinormal distribution is the only spherical multivariate distribution with independent marginals).

Denoting by $l^{(n)}(\mathbf{v}; \mathbf{Z}_1, \dots, \mathbf{Z}_n)$ the value of $l^{(n)}(\mathbf{v})$ computed in the sample $\mathbf{Z}_1, \dots, \mathbf{Z}_n$, we have, by using the affine-invariance of the empirical lift-interdirection function, $l^{(n)}(\mathbf{v}; \mathbf{Z}_1, \dots, \mathbf{Z}_n) = l^{(n)}(\boldsymbol{\Sigma}^{-1/2}\mathbf{v}; \boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_1, \dots, \boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_n)$. Under $\mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$, $(\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_1, \dots, \boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_n)$ is a spherically symmetric sample, still with radial density f , so that

$$\binom{n}{k}^{-1} l^{(n)}(\mathbf{v}) = l_f(\boldsymbol{\Sigma}^{-1/2}\mathbf{v}) + o_{L^2}(1),$$

as $n \rightarrow \infty$. This entails that, under $\bigcup_f \mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$, $l^{(n)}(\mathbf{v})$ is consistent for a monotone increasing function of the unknown Mahalanobis distance $\|\boldsymbol{\Sigma}^{-1/2}\mathbf{v}\|$ between \mathbf{v} and the origin in \mathbb{R}^k . It can therefore be expected that the ranks of the lift-interdirections $l_i^{(n)} := l^{(n)}(\mathbf{Z}_i)$ coincide asymptotically with the ranks of the Mahalanobis distances $\|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_i\|$. The following proposition is the main step in the proof that this does indeed hold true (see the Appendix for a proof).

Proposition 2 *Let $\dot{R}^{(n)}(\mathbf{v}) := \sum_{i=1}^n \mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}]}$, $\dot{\underline{R}}^{(n)}(\mathbf{v}) := \sum_{i=1}^n \mathbb{I}_{[\underline{l}^{(n)}(\mathbf{v}) \geq l_i^{(n)}]}$, and $R^{(n)}(\mathbf{v}; \boldsymbol{\Sigma}) := \sum_{i=1}^n \mathbb{I}_{[\|\boldsymbol{\Sigma}^{-1/2}\mathbf{v}\| \geq \|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_i\|]}$. Then, for all \mathbf{v} , $\boldsymbol{\Sigma}$, and f ,*

$$\text{E}_{\boldsymbol{\Sigma}, f} \left[\left(\dot{R}^{(n)}(\mathbf{v}) - R^{(n)}(\mathbf{v}; \boldsymbol{\Sigma}) \right)^2 \right] \text{ and } \text{E}_{\boldsymbol{\Sigma}, f} \left[\left(\dot{\underline{R}}^{(n)}(\mathbf{v}) - R^{(n)}(\mathbf{v}; \boldsymbol{\Sigma}) \right)^2 \right]$$

are $o(n^2)$, as $n \rightarrow \infty$.

Denote by $\dot{R}_i^{(n)} := \dot{R}^{(n)}(\mathbf{Z}_i)$ the rank of the lift-interdirection $l_i^{(n)}$ among $l_1^{(n)}, \dots, l_n^{(n)}$, and by $R_i^{(n)}(\boldsymbol{\Sigma}) := R^{(n)}(\mathbf{Z}_i; \boldsymbol{\Sigma})$ the rank of $\|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_i\|$ among $\|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_1\|, \dots, \|\boldsymbol{\Sigma}^{-1/2}\mathbf{Z}_n\|$. Then, conditioning with respect to \mathbf{Z}_i , we obtain

$$\mathbb{E}_{\boldsymbol{\Sigma}, f} \left[\left(\frac{\dot{R}_i^{(n)}}{n+1} - \frac{R_i^{(n)}(\boldsymbol{\Sigma})}{n+1} \right)^2 \right] = \int_{\mathbb{R}^k} \mathbb{E}_{\boldsymbol{\Sigma}, f} \left[\left(\frac{\dot{R}_i^{(n)}(\mathbf{v})}{n+1} - \frac{R_i^{(n)}(\mathbf{v}; \boldsymbol{\Sigma})}{n+1} \right)^2 \right] \underline{f}(\mathbf{v}; \boldsymbol{\Sigma}, f) d\mathbf{v}.$$

Using Proposition 2 and Lebesgue's dominated convergence theorem, one obtains that the above expectation is $o(1)$, as $n \rightarrow \infty$. One can prove analogously the corresponding equivalence result for the ranks $\underline{\dot{R}}_i^{(n)} := \underline{\dot{R}}^{(n)}(\mathbf{Z}_i)$ of the symmetrized lift-interdirections. Consequently, we proved the following:

Corollary 1 *Under $\cup_f \mathcal{H}^{(n)}(\boldsymbol{\Sigma}, f)$, $\dot{R}_i^{(n)}/(n+1) = [R_i^{(n)}(\boldsymbol{\Sigma})/(n+1)] + o_P(1)$ and $\underline{\dot{R}}_i^{(n)}/(n+1) = [\underline{R}_i^{(n)}(\boldsymbol{\Sigma})/(n+1)] + o_P(1)$, as $n \rightarrow \infty$.*

As we show in the following section, the asymptotic equivalence result in Corollary 1 allows to build multivariate signed-rank procedures based on interdirections and the ranks of lift-interdirections.

4 Optimal hyperplane-based procedures.

Since the ranks of lift-interdirections are asymptotically equivalent to the true ranks, one could think of replacing, in the class of "hybrid" procedures in (1), the pseudo-Mahalanobis ranks \widehat{R}_i by the ranks \dot{R}_i (or $\underline{\dot{R}}_i$) of lift-interdirections. We explain below why this can indeed be done *safely* (i.e., without affecting the asymptotic behavior of the test statistics) in our generic example : the one-sample location problem.

4.1 The one-sample location problem.

Let us consider again the k -variate one-sample location model $\mathbf{X}_i = \boldsymbol{\theta} + \boldsymbol{\varepsilon}_i$, $i = 1, \dots, n$ where the $\boldsymbol{\varepsilon}_i$'s are elliptically symmetric i.i.d. random vectors. We want to test $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, for some fixed k -vector $\boldsymbol{\theta}_0$. In this section, $q_{ij} := q_{ij}^{(n)}(\boldsymbol{\theta}_0)$, $\widehat{R}_i := \widehat{R}_i^{(n)}(\boldsymbol{\theta}_0)$, $\dot{R}_i := \dot{R}_i^{(n)}(\boldsymbol{\theta}_0)$, and $\underline{\dot{R}}_i := \underline{\dot{R}}_i^{(n)}(\boldsymbol{\theta}_0)$ will respectively denote the interdirections, the pseudo-Mahalanobis ranks, the ranks of non-symmetrized lift-interdirections, and the ranks of symmetrized lift-interdirections, computed within the sample of the residuals $\mathbf{Z}_i(\boldsymbol{\theta}_0) = \mathbf{X}_i - \boldsymbol{\theta}_0$, $i = 1, \dots, n$. We intend to base locally and asymptotically optimal testing procedures on hyperplane-based versions of the test statistics in (1), i.e., on

$$\dot{Q}_K^{(n)} := \frac{k}{n \mathbb{E}[K^2(U)]} \sum_{i,j=1}^n K\left(\frac{\dot{R}_i}{n+1}\right) K\left(\frac{\dot{R}_j}{n+1}\right) \cos(\pi p_{ij}),$$

or alternatively, on

$$\underline{\dot{Q}}_K^{(n)} := \frac{k}{n \mathbb{E}[K^2(U)]} \sum_{i,j=1}^n K\left(\frac{\underline{\dot{R}}_i}{n+1}\right) K\left(\frac{\underline{\dot{R}}_j}{n+1}\right) \cos(\pi p_{ij}),$$

where $p_{ij} := \binom{n}{k-1}^{-1} q_{ij}$ and U is still uniform over $(0, 1)$. The following proposition shows that the procedures based on $\widehat{Q}_K^{(n)}$ and $\dot{Q}_K^{(n)}$ are asymptotically equivalent and share the same invariance properties (the procedures based on the non-symmetrized lift-interdirections suffer from a lack of symmetry that does not allow to adapt the methodology used in Hallin and Paindaveine (2002a); see the proof of the following proposition).

Proposition 3 Denote by $\dot{Q}_g^{(n)}$ the statistic $\dot{Q}_{K(g)}^{(n)}$, associated with $K(g) := \varphi_g \circ \widetilde{G}_k^{-1}$, and consider the sequence of tests $\dot{\phi}_K^{(n)}$ (resp., $\dot{\phi}_g^{(n)}$) rejecting the null hypothesis $\mathcal{H}_0^{(n)}$ whenever $\dot{Q}_K^{(n)}$ (resp., $\dot{Q}_g^{(n)}$) exceeds the α -upper quantile $\chi_{k;1-\alpha}^2$ of a chi-square distribution with k degrees of freedom. Then,

- (i) $\dot{Q}_K^{(n)}$ is affine-invariant.
- (ii) $\dot{Q}_K^{(n)}$ is asymptotically invariant under the group of monotone continuous radial transformations.
- (iii) $\dot{Q}_K^{(n)}$ is asymptotically chi-square with k degrees of freedom under $\mathcal{H}_0^{(n)} = \bigcup_{\Sigma} \bigcup_f \mathcal{H}_{\theta_0}^{(n)}(\Sigma, f)$.
- (iv) The sequence of tests $\dot{\phi}_K^{(n)}$ has asymptotic level α .
- (v) $\dot{\phi}_g^{(n)}$ is locally asymptotically maximin, at asymptotic level α , for $\bigcup_{\Sigma} \bigcup_f \mathcal{H}_{\theta_0}^{(n)}(\Sigma, f)$ against alternatives of the form $\bigcup_{\theta \neq \theta_0} \bigcup_{\Sigma} \mathcal{H}_{\theta}^{(n)}(\Sigma, g)$.

For the concept of *locally asymptotically maximin* tests, we refer to Section 11.9 of Le Cam (1986), which defines and develops *locally asymptotically most stringent* tests—of which locally asymptotically maximin tests are just a particular case.

Proof. First note that (i) directly results from the affine-invariance of interdirections and symmetrized lift-interdirections. Now, the methodology used by Hallin and Paindaveine (2002a) to prove that the test statistics $\widehat{Q}_K^{(n)}$ and the corresponding tests $\widehat{\phi}_K^{(n)}$ fulfill (ii), (iii), (iv), and (v) above consists in establishing the asymptotic representation result

$$\widehat{Q}_K^{(n)} = \widetilde{Q}_{K;\Sigma,f}^{(n)} + o_P(1), \quad \text{as } n \rightarrow \infty, \quad (11)$$

under $\mathcal{H}_{\theta_0}^{(n)}(\Sigma, f)$ (i.e., under $\mathcal{H}^{(n)}(\Sigma, f)$ for the residuals $\mathbf{Z}_i(\theta_0)$), where

$$\widetilde{Q}_{K;\Sigma,f}^{(n)} := \frac{k}{n \mathbb{E}[K^2(U)]} \sum_{i,j=1}^n K(\widetilde{F}_k(\|\Sigma^{-1/2} \mathbf{Z}_i(\theta_0)\|)) K(\widetilde{F}_k(\|\Sigma^{-1/2} \mathbf{Z}_j(\theta_0)\|)) \mathbf{U}_i^T(\Sigma) \mathbf{U}_j(\Sigma),$$

with $\mathbf{U}_i(\Sigma)$ defined as after (2).

Consequently, to show that the test statistics $\dot{Q}_K^{(n)}$ and the tests $\dot{\phi}_K^{(n)}$ enjoy respectively the same asymptotic behavior as $\widehat{Q}_K^{(n)}$ and $\widehat{\phi}_K^{(n)}$ (and in particular fulfill (ii), (iii), (iv), and (v)), it is sufficient to prove that (11) also holds for these hyperplane-based statistics, i.e., that

$$\dot{Q}_K^{(n)} = \widetilde{Q}_{K;\Sigma,f}^{(n)} + o_P(1), \quad \text{as } n \rightarrow \infty, \quad (12)$$

under $\mathcal{H}_{\theta_0}^{(n)}(\Sigma, f)$. Now, a close look at the proof of (11) (see Appendix B in Hallin and Paindaveine 2002a) shows that this asymptotic equivalence result only requires the estimated ranks $\widehat{R}_i^{(n)}$ to be

- (a) invariant under permutations of the residuals,
- (b) invariant under reflections of the residuals with respect to the origin in \mathbb{R}^k , and
- (c) asymptotically equivalent to the “true” ranks, meaning that, for all i , $\widehat{R}_i^{(n)}/(n+1) = [R_i^{(n)}(\Sigma)/(n+1)] + o_{\mathbb{P}}(1)$, under $\bigcup_f \mathcal{H}_{\theta_0}^{(n)}(\Sigma, f)$, as $n \rightarrow \infty$.

The results from the previous sections ensure that (a), (b), and (c) are fulfilled by the estimated ranks $\underline{\dot{R}}_i^{(n)}$ of symmetrized lift-interdirections (note that (b) is not fulfilled by the ranks of the non-symmetrized lift-interdirections). This implies that (12) holds, and consequently that $\widehat{Q}_K^{(n)}$ and $\underline{\dot{Q}}_K^{(n)}$ are asymptotically equivalent under the sequence of null hypotheses $\mathcal{H}_0^{(n)}$ (as well as under sequences of contiguous alternatives). The result follows. \square

Due to their asymptotic equivalence, the hyperplane-based tests $\underline{\dot{\phi}}_K$ and hybrid tests $\widehat{\phi}_K$ share the same asymptotic relative efficiencies with respect to Hotelling’s test (see Proposition 5 in Hallin and Paindaveine (2002a) for an expression of those AREs, as well as numerical values under t -distributions). This implies that the multivariate extension of the Chernoff-Savage (1958) result, which shows that the Gaussian version of the hybrid procedure dominates uniformly Hotelling’s procedure in the Pitman sense, is also valid for the proposed hyperplane-based procedures. More precisely, we have the following :

Corollary 2 *Denote by $\underline{\dot{\phi}}_{vDW}$ the hyperplane-based Gaussian signed-rank procedure $\underline{\dot{\phi}}_g$, $g(r) = \phi(r) = \exp(-r^2/2)$ (van der Waerden scores), and write $\text{ARE}_{k,f}(\underline{\dot{\phi}}_{vDW}/\phi_{T^2})$ for the asymptotic relative efficiency of that procedure w.r.t. Hotelling’s T^2 test under radial density f , in the k -dimensional case. Then, for any f satisfying some mild regularity conditions, $\text{ARE}_{k,f}(\underline{\dot{\phi}}_{vDW}/\phi_{T^2}) \geq 1$, and equality holds if and only if f is Gaussian.*

Another important corollary of the asymptotic equivalence between hyperplane-based and hybrid procedures is that the multivariate extension (see Hallin and Paindaveine 2002a) of the Hodges-Lehmann (1956) result also holds for the proposed Wilcoxon (linear scores) hyperplane-based procedure:

Corollary 3 *Denote by $\underline{\dot{\phi}}_W$ the hyperplane-based Wilcoxon signed-rank procedure $\underline{\dot{\phi}}_K$, $K(u) = u$ for all u . Then, if the infimum is taken over all radial densities f satisfying some mild regularity conditions, $\inf_f \text{ARE}_{k,f}(\underline{\dot{\phi}}_W/\phi_{T^2}) = \frac{81}{500} \frac{(\sqrt{2k-1}+1)^5}{k^2(\sqrt{2k-1}+5)}$.*

The numerical values of this lower bound are 0.864, 0.916, 0.883, 0.853, 0.811, and 0.765, for dimensions 1, 2, 3, 4, 6, and 10 of the observations, respectively. This sequence of lower bounds is monotonically decreasing for $k \geq 2$, and goes to 0.648 as $n \rightarrow \infty$ (see Hallin and Paindaveine (2002a) for the proofs of Corollaries 2 and 3, as well as for the densities in which the above infima are reached; see also Paindaveine (2004) for an elementary proof of Corollary 2).

4.2 Other location and serial problems.

It should be stressed that the same methodology can be used in a broad class of location and serial problems. For instance, consider the problem of testing for multivariate (elliptically symmetric) randomness against VARMA dependence. This problem is the special case

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{\theta} = \mathbf{0} \\ \mathcal{H}_1 : \boldsymbol{\theta} \neq \mathbf{0}, \end{cases}$$

in the model $\mathbf{A}(L)\mathbf{X}_t = \mathbf{B}(L)\boldsymbol{\varepsilon}_t$, $t = 1, \dots, n$, $\mathbf{A}(L) := \mathbf{I}_k - \sum_{r=1}^p \mathbf{A}_r L^r$, $\mathbf{B}(L) := \mathbf{I}_k + \sum_{s=1}^q \mathbf{B}_s L^s$, $\boldsymbol{\theta} := ((\text{vec } \mathbf{A}_1)^T, \dots, (\text{vec } \mathbf{A}_p)^T, (\text{vec } \mathbf{B}_1)^T, \dots, (\text{vec } \mathbf{B}_q)^T)^T$. In this case, the (null) residuals $\mathbf{Z}_t = \mathbf{Z}_t(\mathbf{0})$ coincide with the observations \mathbf{X}_t .

Hallin and Paindaveine (2002c) developed locally and asymptotically optimal signed-rank procedures for this problem, which consist (at asymptotic level α) in rejecting $\mathcal{H}_0 : \boldsymbol{\theta} = \mathbf{0}$ as soon as

$$\begin{aligned} \widehat{Q}_K^{(n)} := & \frac{k^2}{\mathbb{E}[K_1^2(U)]\mathbb{E}[K_2^2(U)]} \sum_{i=1}^{\max(p,q)} (n-i)^{-1} \sum_{s,t=i+1}^n K_1\left(\frac{\widehat{R}_s}{n+1}\right) K_1\left(\frac{\widehat{R}_t}{n+1}\right) \\ & \times K_2\left(\frac{\widehat{R}_{s-i}}{n+1}\right) K_2\left(\frac{\widehat{R}_{t-i}}{n+1}\right) \cos(\pi p_{s,t}) \cos(\pi p_{s-i,t-i}) > \chi_{k^2 \max(p,q); 1-\alpha}^2, \end{aligned} \quad (13)$$

where U is uniform over $(0, 1)$, $K_1, K_2 : (0, 1) \rightarrow \mathbb{R}$ are square-integrable score functions, and the \widehat{R}_t 's and $p_{t,i}$'s are respectively the pseudo-Mahalanobis ranks and (normalized) interdirections of the $\mathbf{Z}_t = \mathbf{X}_t$'s.

Again, it can be checked that (a), (b), and (c) in the proof of Proposition 3 are sufficient conditions to replace safely pseudo-Mahalanobis ranks by the ranks of symmetrized lift-interdirections. Consequently, one can base purely hyperplane-based procedures for testing for multivariate randomness on statistics of the form

$$\begin{aligned} \dot{Q}_K^{(n)} := & \frac{k^2}{\mathbb{E}[K_1^2(U)]\mathbb{E}[K_2^2(U)]} \sum_{i=1}^{\max(p,q)} (n-i)^{-1} \sum_{s,t=i+1}^n K_1\left(\frac{\dot{R}_s}{n+1}\right) K_1\left(\frac{\dot{R}_t}{n+1}\right) \\ & \times K_2\left(\frac{\dot{R}_{s-i}}{n+1}\right) K_2\left(\frac{\dot{R}_{t-i}}{n+1}\right) \cos(\pi p_{s,t}) \cos(\pi p_{s-i,t-i}). \end{aligned} \quad (14)$$

The same methodology can also be used in more general models (such as the multivariate GLM with VARMA error terms), both for testing simple null hypotheses—for which the value of the parameter of interest $\boldsymbol{\theta}$ is specified under the null—and for testing linear hypotheses on the parameter $\boldsymbol{\theta}$ (for the latter, one ends up with *locally asymptotically most stringent*—rather than maximin—tests; see Section 11.9 of Le Cam 1986). This covers a lot of testing problems of practical interest, such as the two-sample problem, c -sample problem (MANOVA problem), Durbin-Watson problem, the problem of testing the order of a VARMA model, etc. (see Hallin and Paindaveine 2004b). Note that, for the c -sample problem, the resulting procedures are the (fully) hyperplane-based analogs to the hybrid procedures developed by Randles and Um (1998).

5 Finite-sample performances.

We performed two Monte-Carlo experiments in order to compare the small-sample behaviors and robustness features of the proposed hyperplane-based procedures with those of hybrid procedures

(based on pseudo-Mahalanobis ranks and interdirections) and those of shape-based procedures (based on pseudo-Mahalanobis ranks and standardized spatial signs).

5.1 Small-sample powers.

We considered the bivariate one-sample problem (with $\boldsymbol{\theta}_0 = \mathbf{0} \in \mathbb{R}^2$), and generated $N = 2,000$ independent samples $\mathbf{X}_i = \boldsymbol{\theta} + \boldsymbol{\varepsilon}_i$, $i = 1, \dots, 30$ of size $n = 30$. The noise $\boldsymbol{\varepsilon}_i$, $i = 1, \dots, 30$ was drawn from the following distributions :

- \mathcal{N} (elliptical symmetry; light tails) bivariate standard normal $\mathcal{N}_2(\mathbf{0}, \mathbf{I}_2)$,
- (t_3) (elliptical symmetry; heavy tails) bivariate standard Student with 3 degrees of freedom,
- (Eh) a mixture of $\mathcal{N}_2(\boldsymbol{\mu}, \mathbf{I}_2)$ and $\mathcal{N}_2(-\boldsymbol{\mu}, \mathbf{I}_2)$ with mixing probabilities 1/2, 1/2 and $\boldsymbol{\mu} = (3, 0)'$,
- (Ec) (*radial* outliers) a mixture of $\mathcal{N}_2(\mathbf{0}, \mathbf{I}_2)$ and $\mathcal{N}_2(\mathbf{0}, \sigma^2 \mathbf{I}_2)$ with $\sigma = 20$ and mixing probabilities 0.8 and 0.2, respectively,
- (O20) (*angular* outliers) a mixture in which a proportion $(1 - c)$ of the sample is drawn from a $\mathcal{N}_2\left(\mathbf{0}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & 1 \end{pmatrix}\right)$ distribution, a proportion $c/2$ is drawn from a uniform distribution over the unit ball centered at $\boldsymbol{\mu}$, and a proportion $c/2$ is drawn from a uniform distribution over the unit ball centered at $-\boldsymbol{\mu}$, with $\sigma = 20$, $c = 0.2$ and $\boldsymbol{\mu} = (0, 2)'$,
- (O40) (*angular* outliers) same as (O20), but $c = 0.4$,
- ($\tilde{\text{O}}40$) (*angular* outliers) same as (O20), but $c = 0.4$ and $\boldsymbol{\mu} = (0, 4)'$.

For each replication, the following tests were performed at nominal probability level $\alpha = 5\%$: the exact and asymptotic Hotelling T^2 tests, and the shape-based, hybrid, and hyperplane-based signed-rank tests with sign scores, Wilcoxon scores, van der Waerden scores, and t_3 -scores. Note that both the hybrid sign test and the hyperplane-based sign test do coincide with Randles (1989)'s test, and that the shape-based sign test is nothing but Randles (2000)'s test. Similarly, the hybrid Wilcoxon procedure is Peters and Randles (1990)'s test.

Rejection frequencies, estimating the corresponding sizes and powers, were recorded at four values of $\boldsymbol{\theta}$, of the form $m\boldsymbol{\Delta}$ ($m = 0, 1, 2, 3$) (the value of $\boldsymbol{\Delta}$ was adjusted for the various populations to examine somewhat similar points on the power curves); they are reported in Table 1 (Tyler's estimator of scatter was used to compute pseudo-Mahalanobis distances; it was obtained via the iterative scheme described in Randles (2000), and iterations were stopped when the Frobenius distance between both sides of the estimating equation defining Tyler's estimator fell below 10^{-8}).

This simulation shows that the three classes of procedures behave quite similarly under elliptically symmetric distributions ((\mathcal{N}) , (t_3)) and under the mixtures (Eh) and (Ec) (however, note that the relative behaviors under the mixture (Eh) pretty much depend on m , that is, on the points considered on the power curves). Finally, the most important differences occur in the presence of *angular* outliers: the hyperplane-based procedures seem to be significantly more robust than the hybrid ones, which, in turn, appear to be more robust than the shape-based procedures.

pop.	test	Shift			
		0	Δ	2Δ	3Δ
\mathcal{N} $\Delta = \begin{pmatrix} .25 \\ 0 \end{pmatrix}$	T^2	.0505 / .0735	.1845 / .2340	.6220 / .6910	.9500 / .9620
	S	.0470 (.0495) [.0495]	.1550 (.1560) [.1560]	.5080 (.5040) [.5040]	.8760 (.8695) [.8695]
	W	.0445 (.0425) [.0415]	.1620 (.1630) [.1615]	.5795 (.5755) [.5865]	.9245 (.9200) [.9245]
	vdW	.0355 (.0355) [.0360]	.1480 (.1475) [.1505]	.5555 (.5530) [.5625]	.9180 (.9105) [.9180]
	t_3	.0535 (.0550) [.0555]	.1740 (.1795) [.1755]	.5740 (.5680) [.5655]	.9130 (.9055) [.9045]
t_3 $\Delta = \begin{pmatrix} .33 \\ 0 \end{pmatrix}$	T^2	.0405 / .0705	.1735 / .2205	.5410 / .6025	.8525 / .8870
	S	.0470 (.0495) [.0495]	.2170 (.2115) [.2115]	.6645 (.6585) [.6585]	.9480 (.9410) [.9410]
	W	.0455 (.0465) [.0475]	.1815 (.1780) [.1760]	.5585 (.5515) [.5590]	.8615 (.8625) [.8815]
	vdW	.0410 (.0410) [.0385]	.1625 (.1630) [.1610]	.5435 (.5375) [.5450]	.8660 (.8625) [.8805]
	t_3	.0505 (.0530) [.0530]	.2385 (.2320) [.2270]	.7110 (.7050) [.7020]	.9575 (.9545) [.9555]
Eh $\Delta = \begin{pmatrix} .70 \\ 0 \end{pmatrix}$	T^2	.0565 / .0720	.1520 / .2020	.4920 / .5690	.8770 / .9175
	S	.0525 (.0500) [.0500]	.0660 (.0730) [.0730]	.1605 (.1655) [.1655]	.4185 (.3920) [.3920]
	W	.0510 (.0540) [.0445]	.1265 (.1370) [.2515]	.4835 (.5110) [.6420]	.8960 (.8680) [.8400]
	vdW	.0475 (.0480) [.0365]	.0960 (.1005) [.1755]	.3535 (.3750) [.4785]	.8030 (.7695) [.7355]
	t_3	.0545 (.0485) [.0540]	.1215 (.1365) [.1410]	.3390 (.3535) [.3575]	.6540 (.6220) [.6265]
Ec $\Delta = \begin{pmatrix} .37 \\ 0 \end{pmatrix}$	T^2	.0160 / .0290	.0300 / .0545	.0820 / .1200	.1540 / .1840
	S	.0470 (.0495) [.0495]	.2075 (.2050) [.2050]	.6585 (.6560) [.6560]	.9400 (.9395) [.9395]
	W	.0415 (.0450) [.0420]	.1645 (.1610) [.1605]	.4790 (.4740) [.4750]	.7200 (.7115) [.7340]
	vdW	.0395 (.0375) [.0380]	.1465 (.1465) [.1480]	.4550 (.4530) [.4515]	.7225 (.7140) [.7245]
	t_3	.0505 (.0530) [.0525]	.2450 (.2420) [.2355]	.7325 (.7285) [.7250]	.9570 (.9540) [.9565]
O20 $\Delta = \begin{pmatrix} .32 \\ .32 \end{pmatrix}$	T^2	.0525 / .0745	.1765 / .2300	.6105 / .6760	.9465 / .9660
	S	.0465 (.0510) [.0510]	.1185 (.1210) [.1210]	.3710 (.3955) [.3955]	.6705 (.6940) [.6940]
	W	.0460 (.0460) [.0400]	.1430 (.1460) [.1705]	.4970 (.5315) [.5890]	.8615 (.8895) [.9155]
	vdW	.0405 (.0395) [.0345]	.1275 (.1315) [.1530]	.4450 (.4730) [.5530]	.8085 (.8370) [.8910]
	t_3	.0540 (.0585) [.0550]	.1380 (.1435) [.1415]	.4560 (.4820) [.4625]	.8095 (.8280) [.8155]
O40 $\Delta = \begin{pmatrix} .36 \\ .36 \end{pmatrix}$	T^2	.0545 / .0770	.1645 / .2190	.5670 / .6375	.9215 / .9520
	S	.0505 (.0445) [.0445]	.0755 (.1270) [.1270]	.1515 (.2900) [.2900]	.3545 (.5070) [.5070]
	W	.0665 (.0575) [.0430]	.1035 (.1490) [.2245]	.2815 (.4675) [.6195]	.6310 (.7910) [.8630]
	vdW	.0555 (.0475) [.0375]	.0825 (.1255) [.1890]	.2090 (.3740) [.5255]	.4985 (.6665) [.7830]
	t_3	.0595 (.0490) [.0530]	.1175 (.1765) [.1630]	.3020 (.4440) [.4100]	.5900 (.7020) [.6675]
$\tilde{O}40$ $\Delta = \begin{pmatrix} .60 \\ .60 \end{pmatrix}$	T^2	.0450 / .0665	.1505 / .2020	.5155 / .5880	.8955 / .9295
	S	.0485 (.0510) [.0510]	.1040 (.2195) [.2195]	.2380 (.4570) [.4570]	.4275 (.5885) [.5885]
	W	.0405 (.0445) [.0470]	.1290 (.2665) [.3545]	.3330 (.5995) [.6590]	.6560 (.8285) [.8205]
	vdW	.0325 (.0390) [.0385]	.1050 (.2265) [.3105]	.2615 (.5125) [.6110]	.5150 (.7265) [.7775]
	t_3	.0525 (.0565) [.0570]	.1175 (.2600) [.2625]	.3210 (.5545) [.5020]	.6090 (.7485) [.6720]

Table 1 Estimated sizes and powers of the exact/asymptotic Hotelling T^2 tests, and of the shape-based, hybrid (in parentheses), and hyperplane-based (in brackets) signed-rank tests with sign scores (S), Wilcoxon scores (W), van der Waerden scores (vdW) and t_3 -scores (t_3), under various values of the shift and various densities; simulations are based on 2,000 bivariate samples of size 30.

5.2 A robustness study.

We also conducted the following simulation in order to investigate further the robustness properties of the different types of procedures considered in the paper. For each value of $m \in \{0, 1, 2\}$, we generated $N = 1,000$ samples of size $n = 50$ from the bivariate normal distribution with mean vector $m\Delta = m(0, .25)'$ and covariance matrix \mathbf{I}_2 . In all samples, we replaced the first two observations by $m\Delta - r\mathbf{c}$ and $m\Delta + r\mathbf{c}$ (according to the value of m), with $\mathbf{c} = (1, 1)'/\sqrt{2}$.

In Figure 3, we report the graphs of the mean p -values of Hotelling's test and of the different versions of the sign, Wilcoxon, and van der Waerden tests, for each value of m and for a broad range of values for r . Clearly, the Hotelling test crashes when the outlying observations are taken away from the center of the distribution. On the contrary, the p -values of the different signed-rank tests remain fairly stable in the same situation. There does not seem however to be any significant differences between the shape-based, hybrid, and hyperplane-based procedures.

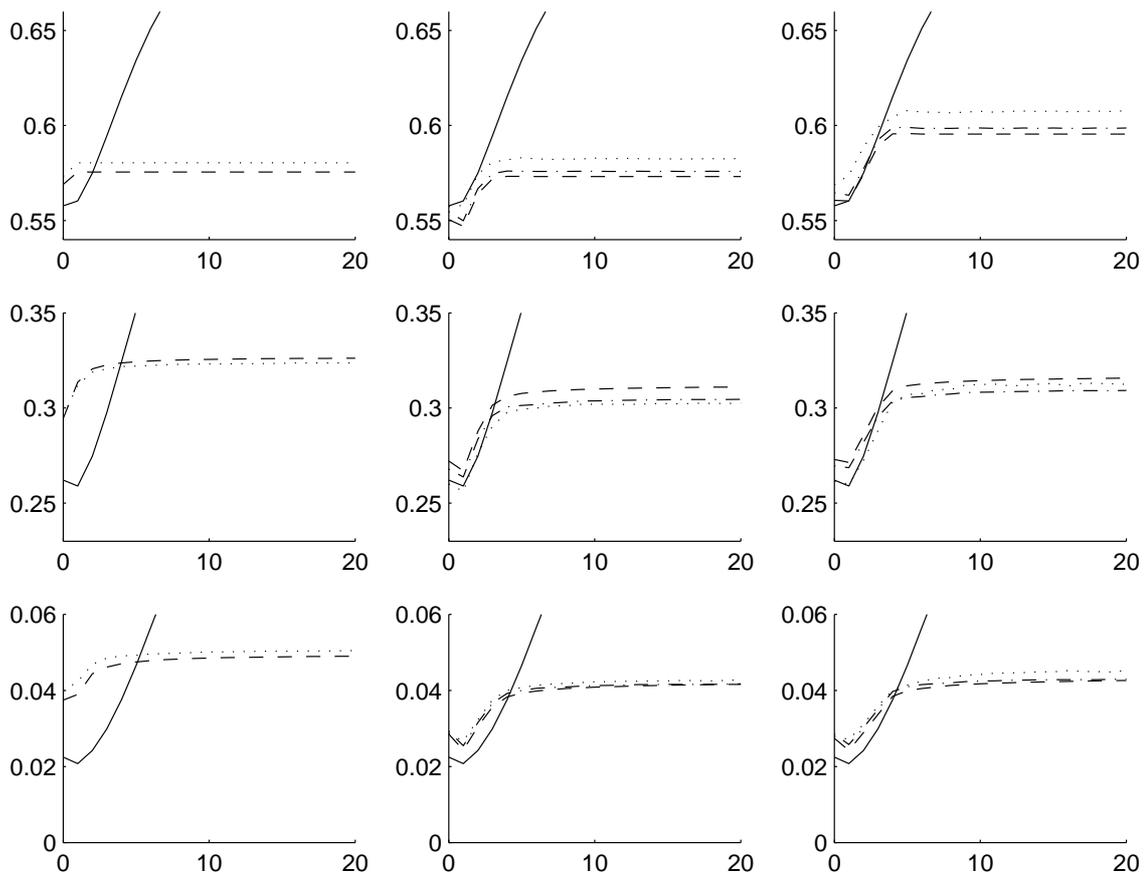


Figure 3: Mean p -values, for $m = 0, 1, 2$ (in the 1st, 2nd, and 3rd row, respectively) of Hotelling's test and of the shape-based, hybrid, and hyperplane-based signed-rank tests with sign scores (1st column), Wilcoxon scores (2nd column), and van der Waerden scores (3rd column), for $r \in [0, 20]$. The solid line refers to Hotelling's test, the dashed to the shape-based tests, the dashdotted line to the hybrid tests, and the dotted line to the hyperplane-based tests.

6 Conclusions and final comments.

In this paper, we introduced a class of hyperplane-based signed-rank procedures for the multivariate one-sample problem, as for the problem of testing for multivariate randomness against VARMA serial dependence (the same methodology can be used for a much broader family of location and serial models, culminating in the multivariate general linear models with VARMA errors). The proposed hyperplane-based procedures are asymptotically equivalent to their (optimal) hybrid and shape-based counterparts based on pseudo-Mahalanobis ranks (introduced in Hallin and Paindaveine 2002a, b, and c). Therefore the three classes of procedures share the same optimality properties (in particular, the van der Waerden versions of these procedures dominate uniformly the classical Gaussian procedures). They also share the same invariance properties.

Interestingly enough, these hyperplane-based procedures show that (locally and asymptotically) optimal inference for the problems under consideration can be based only on indicator functions associated with the separations of pairs of observations by data-based hyperplanes. In particular, these procedures allow to *avoid the estimation of the shape matrix*.

The small-sample performances of the different procedures are quite similar under elliptical distributions or mixtures of elliptical distributions. However, hyperplane-based procedures are significantly more powerful in the presence of (radial) outliers. The main disadvantage of the hyperplane-based procedures lies in computability issues. Interdirections and lift-interdirections are *counts*, and their computation is heavy, even for relatively small space dimensions. Of course, one can also resort to subsampling (i.e., not consider *all* hyperplanes, and end up with *estimated* interdirections and lift-interdirections).

7 Appendix.

Proof of Proposition 1. With $J = (1, 2, \dots, k)$, we have

$$d_0(J) + \mathbf{d}^T(J)\mathbf{v} = \det \begin{pmatrix} \mathbf{1}^T & \mathbf{1} \\ \mathbf{Z} & \mathbf{v} \end{pmatrix} = \det \begin{pmatrix} \mathbf{1}^T & 0 \\ \mathbf{Z} & \mathbf{v} - \mathbf{Z}\mathbf{1} \end{pmatrix} = \mathbf{d}^T(J)(\mathbf{v} - \mathbf{Z}\mathbf{1}),$$

so that $l_f(\mathbf{v}) = \mathbb{P}[\text{sign}(d_0(J) + \mathbf{d}^T(J)\mathbf{v}) \text{sign}(d_0(J) - \mathbf{d}^T(J)\mathbf{v}) = -1] = \mathbb{P}[|\mathbf{d}^T(J)\mathbf{Z}\mathbf{1}| \leq |\mathbf{d}^T(J)\mathbf{v}|]$. Since $\mathbf{Z}\mathbf{1}$ belongs to hyperplane J (so that $d_0(J) + \mathbf{d}^T(J)\mathbf{Z}\mathbf{1} = 0$), we obtain

$$l_f(\mathbf{v}) = \mathbb{P} \left[\frac{|d_0(J)|}{\|\mathbf{d}(J)\|} \leq |\mathbf{u}^T(J)\mathbf{v}| \right], \quad (15)$$

where $\mathbf{u}(J) := \mathbf{d}(J)/\|\mathbf{d}(J)\|$ is uniformly distributed over \mathcal{S}^{k-1} (if the \mathbf{Z}_i 's are spherically symmetric, it trivially follows from (8) that $\mathbf{d}(J)$ is also spherically symmetric). Now, write $[\mathbf{Z}]_i$ (resp., $[\mathbf{Z}]_{i,j}$) for the array obtained by deleting the i th row in \mathbf{Z} (resp., by deleting the i th row and the j th column in \mathbf{Z}). Then, denoting by \mathbf{e}_i the i th vector of the canonical basis in \mathbb{R}^k , we obtain

$$|\mathbf{d}^T(J)\mathbf{e}_i| = \left| \det \begin{pmatrix} \mathbf{1}^T & 0 \\ \mathbf{Z} & \mathbf{e}_i \end{pmatrix} \right| = \left| \det \begin{pmatrix} \mathbf{1}^T \\ [\mathbf{Z}]_i \end{pmatrix} \right| = \left| \sum_{j=1}^k (-1)^{i+j} \det[\mathbf{Z}]_{i,j} \right|.$$

Writing $\text{Adj } \mathbf{Z}$ for the adjoint matrix of \mathbf{Z} (defined by $(\text{Adj } \mathbf{Z})_{ji} := (-1)^{i+j} \det[\mathbf{Z}]_{i,j}$), this yields

$$|\mathbf{d}^T(J)\mathbf{e}_i| = \left| ((\text{Adj } \mathbf{Z})^T \mathbf{1})^T \mathbf{e}_i \right| = |\det \mathbf{Z}| \left| ((\mathbf{Z}^T)^{-1} \mathbf{1})^T \mathbf{e}_i \right|.$$

Consequently, $\|\mathbf{d}(J)\|^2 = (\det \mathbf{Z})^2 \mathbf{1}^T \mathbf{Z}^{-1} (\mathbf{Z}^T)^{-1} \mathbf{1}$, which yields the result since $d_0(J) = (-1)^k \det(\mathbf{Z})$. \square

Proof of Proposition 2. Thanks to affine-invariance, we may restrict without loss of generality to the spherical case $\Sigma = \mathbf{I}_k$. Letting $R^{(n)}(\mathbf{v}) := R^{(n)}(\mathbf{v}; \mathbf{I}_k)$, we have

$$\begin{aligned} \dot{R}^{(n)}(\mathbf{v}) - R^{(n)}(\mathbf{v}) &= \sum_{i=1}^n \left(\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}]} - \mathbb{I}_{[\|\mathbf{v}\| \geq \|\mathbf{Z}_i\|]} \right) \\ &= \sum_{i=1}^n \left(\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}]} \mathbb{I}_{[\|\mathbf{v}\| < \|\mathbf{Z}_i\|]} - \mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_i^{(n)}]} \mathbb{I}_{[\|\mathbf{v}\| \geq \|\mathbf{Z}_i\|]} \right), \end{aligned}$$

so that

$$n^{-2} \left(\dot{R}^{(n)}(\mathbf{v}) - R^{(n)}(\mathbf{v}) \right)^2 = n^{-2} \sum_{i,j=1}^n \left(\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}, \|\mathbf{v}\| < \|\mathbf{Z}_i\|]} - \mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_i^{(n)}, \|\mathbf{v}\| \geq \|\mathbf{Z}_i\|]} \right) \left(\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_j^{(n)}, \|\mathbf{v}\| < \|\mathbf{Z}_j\|]} - \mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_j^{(n)}, \|\mathbf{v}\| \geq \|\mathbf{Z}_j\|]} \right).$$

Consequently, it is sufficient to prove that

$$\mathbb{E} \left[\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}, \|\mathbf{v}\| < \|\mathbf{Z}_i\|]} \mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_j^{(n)}, \|\mathbf{v}\| < \|\mathbf{Z}_j\|]} \right], \mathbb{E} \left[\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}, \|\mathbf{v}\| < \|\mathbf{Z}_i\|]} \mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_j^{(n)}, \|\mathbf{v}\| \geq \|\mathbf{Z}_j\|]} \right],$$

and

$$\mathbb{E} \left[\mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_i^{(n)}, \|\mathbf{v}\| \geq \|\mathbf{Z}_i\|]} \mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_j^{(n)}, \|\mathbf{v}\| \geq \|\mathbf{Z}_j\|]} \right]$$

are $o(1)$, for all i, j , as $n \rightarrow \infty$. This in turn will follow if we can show that

$$\mathbb{E} \left[\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l_i^{(n)}, \|\mathbf{v}\| < \|\mathbf{Z}_i\|]} \right] \tag{16}$$

and

$$\mathbb{E} \left[\mathbb{I}_{[l^{(n)}(\mathbf{v}) < l_i^{(n)}, \|\mathbf{v}\| \geq \|\mathbf{Z}_i\|]} \right] \tag{17}$$

are $o(1)$, for all i , as $n \rightarrow \infty$. Conditioning with respect to \mathbf{Z}_i , (16) becomes

$$\int_{\mathbb{R}^k} g_{\mathbf{v}}^{(n)}(\mathbf{z}) \underline{f}(\mathbf{z}) d\mathbf{z}, \tag{18}$$

where

$$g_{\mathbf{v}}^{(n)}(\mathbf{z}) := \mathbb{E} \left[\mathbb{I}_{[l^{(n)}(\mathbf{v}) \geq l^{(n)}(\mathbf{z}), \|\mathbf{v}\| < \|\mathbf{z}\|]} \right].$$

Of course, if $\|\mathbf{z}\| \leq \|\mathbf{v}\|$, $g_{\mathbf{v}}^{(n)}(\mathbf{z}) = 0$ for all n . Now, if $\|\mathbf{z}\| > \|\mathbf{v}\|$, we also have $l_f(\mathbf{v}) < l_f(\mathbf{z})$, so that

$$\begin{aligned} g_{\mathbf{v}}^{(n)}(\mathbf{z}) &= P \left[l^{(n)}(\mathbf{v}) \geq l^{(n)}(\mathbf{z}) \right] \\ &\leq P \left[\left| (l^{(n)}(\mathbf{v})/\binom{n}{k} - l^{(n)}(\mathbf{z})/\binom{n}{k}) - (l_f(\mathbf{v}) - l_f(\mathbf{z})) \right| > |l_f(\mathbf{v}) - l_f(\mathbf{z})| \right] \\ &\leq \frac{\text{Var} \left[l^{(n)}(\mathbf{v})/\binom{n}{k} - l^{(n)}(\mathbf{z})/\binom{n}{k} \right]}{(l_f(\mathbf{v}) - l_f(\mathbf{z}))^2}. \end{aligned}$$

Since the U-statistic $l^{(n)}(\mathbf{v})/\binom{n}{k}$ (resp., $l^{(n)}(\mathbf{z})/\binom{n}{k}$) has a bounded kernel, $\text{Var} [l^{(n)}(\mathbf{v})/\binom{n}{k}]$ (resp., $\text{Var} [l^{(n)}(\mathbf{z})/\binom{n}{k}]$) is $O(1/n)$, as $n \rightarrow \infty$. Therefore, $g_{\mathbf{v}}^{(n)}(\mathbf{z}) = o(1)$ as $n \rightarrow \infty$, also when $\|\mathbf{z}\| > \|\mathbf{v}\|$. Consequently, we showed that $g_{\mathbf{v}}^{(n)}(\mathbf{z}) = o(1)$ as $n \rightarrow \infty$, for all \mathbf{z} . Since $g_{\mathbf{v}}^{(n)}(\mathbf{z}) \leq 1$ for all $\mathbf{z} \in \mathbb{R}^k$, Lebesgue's dominated convergence theorem allows to conclude that (18) is $o(1)$, as $n \rightarrow \infty$, i.e., that (16) is $o(1)$ as $n \rightarrow \infty$. One can similarly prove that (17) is $o(1)$ as $n \rightarrow \infty$. In that case, however, the associated pointwise convergence $g_{\mathbf{v}}^{(n)}(\mathbf{z}) = o(1)$ as $n \rightarrow \infty$, only holds \mathbf{z} -a.e. (which is enough for Lebesgue's dominated convergence theorem).

The result on the ranks of the symmetrized lift-interdirections can be proved in the same way, since $\underline{l}^{(n)}(\mathbf{v})/[2^k \binom{n}{k}] = l^{(n)}(\mathbf{v})/\binom{n}{k} + o_{L^2}(1) = l_f(\mathbf{\Sigma}^{-1/2}\mathbf{v}) + o_{L^2}(1)$, as $n \rightarrow \infty$, under $\mathcal{H}^{(n)}(\mathbf{\Sigma}, f)$. \square

References

- [1] Bilodeau, M., and D. Brenner (1999). *Theory of multivariate statistics*. Springer-Verlag, New York.
- [2] Chernoff, H., and I. R. Savage (1958). Asymptotic normality and efficiency of certain non-parametric tests, *Ann. Math. Statist.* **29**, 972-994.
- [3] Hallin, M., and D. Paindaveine (2002a). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks, *Ann. Statist.* **30**, 1103-1133.
- [4] Hallin, M., and D. Paindaveine (2002b). Multivariate signed ranks : Randles' interdirections or Tyler's angles? In Y. Dodge, Ed., *Statistical Data Analysis Based on the L_1 -Norm and Related Procedures*, Birkhäuser, Basel, 271-282.
- [5] Hallin, M., and D. Paindaveine (2002c). Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence, *Bernoulli* **8**, 787-815.
- [6] Hallin, M., and D. Paindaveine (2004a). Rank-based optimal tests of the adequacy of an elliptic VARMA model. *Ann. Statist.*, to appear.
- [7] Hallin, M., and D. Paindaveine (2004b). Affine invariant aligned rank tests for the multivariate general linear model with ARMA errors. *J. Multivariate Anal.*, to appear.
- [8] Hettmansperger, T. P., J. Nyblom, and H. Oja (1994). Affine invariant multivariate one-sample sign tests, *J. Roy. Statist. Soc. Ser. B* **56**, 221-234.
- [9] Hettmansperger, T. P., J. Möttönen, and H. Oja (1997). Affine invariant multivariate one-sample signed-rank tests, *J. Amer. Statist. Assoc.* **92**, 1591-1600.
- [10] Hettmansperger, T.P., Möttönen, J. & Oja, H. (1998). The geometry of the affine invariant multivariate sign and ranks methods, *J. Nonparam. Statist.* **11**, 271-285.
- [11] Hodges, J. L. and E. L. Lehmann (1956). The efficiency of some nonparametric competitors of the t -test, *Ann. Math. Statist.* **27**, 324-335.
- [12] Koshevoy, G., J. Möttönen, and H. Oja (2003). A scatter matrix estimate based on the zonotope, *Ann. Statist.* **31**, 1439-1459.
- [13] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer-Verlag, New York.
- [14] Möttönen, J., and H. Oja (1995). Multivariate spatial sign and rank methods, *J. Nonparam. Statist.* **5**, 201-213.
- [15] Möttönen, J., H. Oja, and J. Tienari (1997). On the efficiency of multivariate spatial sign and rank methods, *Ann. Statist.* **25**, 542-552.
- [16] Möttönen, J., T. P. Hettmansperger, H. Oja, and J. Tienari (1998). On the efficiency of the multivariate affine invariant rank methods, *J. Multivariate Anal.* **66**, 118-132.

- [17] Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates : a review, *Scand. J. Statist.* **26**, 319–343.
- [18] Paindaveine, D. (2004). A unified and elementary proof of serial and nonserial, univariate and multivariate, Chernoff-Savage results. Submitted.
- [19] Peters, D., and R. H. Randles (1990). A multivariate signed-rank test for the one-sample location problem, *J. Amer. Statist. Assoc.* **85**, 552–557.
- [20] Puri, M. L., and P. K. Sen (1971). *Nonparametric methods in multivariate analysis*. J. Wiley, New York.
- [21] Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections, *J. Amer. Statist. Assoc.* **84**, 1045-1050.
- [22] Randles, R. H. (2000). A simpler, affine-invariant, multivariate, distribution-free sign test, *J. Amer. Statist. Assoc.* **95**, 1263–1268.
- [23] Randles, R. H. and Y. Um (1998). Nonparametric tests for the multivariate multi-sample location problem, *Statistica Sinica* **8**, 801–812.
- [24] Singh, K. (1991). A notion of majority depth. Preprint.
- [25] Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter, *Ann. Statist.* **15**, 234–251.
- [26] Zuo, Y. and R. Serfling (2000). General notions of statistical depth function, *Ann. Statist.* **28**, 461-482.