OPTIMAL *R*-ESTIMATION OF A SPHERICAL LOCATION

Christophe Ley[1], Yvik Swan[2], Baba Thiam[3] and Thomas Verdebout[3]

[1]*Université libre de Bruxelles,* [2]*Université du Luxembourg,*
and [3]*Université Lille Nord de France*

Abstract: In this paper, we provide *R*-estimators of the location of a rotationally symmetric distribution on the unit sphere of $\mathbb{R}^k$. In order to do so we first prove the local asymptotic normality property of a sequence of rotationally symmetric models; this is a non-standard result due to the curved nature of the unit sphere. We then construct our estimators by adapting the Le Cam one-step methodology to spherical statistics and ranks. We show that they are asymptotically normal under any rotationally symmetric distribution and achieve the efficiency bound under a specific density. Their small sample behavior is studied via a Monte Carlo simulation and our methodology is illustrated on geological data.

Key words and phrases: Local asymptotic normality, rank-based methods, *R*-estimation, spherical statistics.

1. Introduction

Spherical data arise naturally in such natural sciences as geology and astrophysics (see, e.g., Watson (1983) or Mardia and Jupp (2000)), as well as in studies of animal behavior (see Fisher, Lewis, and Embleton (1987)) and in neuroscience (see Leong and Carlile (1998)). It is common practice to view such data as realizations of random vectors $\boldsymbol{X}$ taking values on the surface of the unit sphere $\mathcal{S}^{k-1} := \{\mathbf{v} \in \mathbb{R}^k \,|\, \|\mathbf{v}\| = 1\}$, the distribution of $\boldsymbol{X}$ depending only on its distance – in a sense to be made precise – from a fixed point $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$. This parameter $\boldsymbol{\theta}$, which can be viewed as a "north pole" (or "mean direction") for the problem under study, is then a location parameter for the distribution.

The first distribution tailored to the specificities of spherical data is the *Fisher-Von Mises-Langevin* (FVML) distribution introduced in Fisher (1953). To this date it remains the distribution which is the most widely used in practice; it plays, in the spherical context, the central role enjoyed by the Gaussian distribution for linear data. The FVML is not the only "spherical distribution", and there exists a wide variety of families of such distributions possessing different advantages and drawbacks (see Mardia and Jupp (2000) for an overview).

In this work we concentrate our attention on the family of *rotationally symmetric distributions* introduced by Saw (1978) (see Section 2 below for a definition). Aside from the fact that it encompasses many well-known spherical

distributions (including the FVML), this family satisfies a natural requirement : it is invariant to the actual choice of "north pole". This entails that the family falls within the much more general class of *statistical group models* (see for instance Chang (2004)) and thus enjoys all the advantages of this class. Moreover, it satisfies a fundamental lemma, due to Watson (1983).

Lemma 1 (Watson (1983)). *If the distribution of $\boldsymbol{X}$ is rotationally symmetric on $\mathcal{S}^{k-1}$ and if the true location is $\boldsymbol{\theta}$, then* (i) $\boldsymbol{X}'\boldsymbol{\theta}$ *and* $\boldsymbol{S_\theta}(\boldsymbol{X}) := (\boldsymbol{X} - (\boldsymbol{X}'\boldsymbol{\theta})\boldsymbol{\theta})/\|\boldsymbol{X} - (\boldsymbol{X}'\boldsymbol{\theta})\boldsymbol{\theta}\|$ *are stochastically independent and* (ii) *the multivariate sign vector* $\boldsymbol{S_\theta}(\boldsymbol{X})$ *is uniformly distributed on* $\mathcal{S}^{k-1}(\boldsymbol{\theta}^\perp) := \{\mathbf{v} \in \mathbb{R}^k \,|\, \|\mathbf{v}\| = 1, \mathbf{v}'\boldsymbol{\theta} = 0\}$.

Estimation and testing procedures for the spherical location parameter $\boldsymbol{\theta}$ have been extensively studied in the literature, with much of the focus in the past years being put on the class of M-estimators. An M-estimator $\hat{\boldsymbol{\theta}}$ associated with a given function $\rho_0(\mathbf{x}; \boldsymbol{\theta})$ is defined as the value of $\boldsymbol{\theta}$ that minimizes the objective function

$$\boldsymbol{\theta} \mapsto \rho(\boldsymbol{\theta}) := \sum_{i=1}^{n} \rho_0(\boldsymbol{X}_i; \boldsymbol{\theta}),$$

where $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are spherical observations. These M-estimators are robust to outliers (see Ko and Chang (1993)) and enjoy nice asymptotic properties (see Chang and Rivest (2001), Chang and Tsai (2003), or Chang (2004)). In particular, the choice $\rho_0(\mathbf{x}; \boldsymbol{\theta}) = \arccos(\mathbf{x}'\boldsymbol{\theta})$ yields the so-called *spherical median* introduced by Fisher (1985); taking $\rho_0(\mathbf{x}; \boldsymbol{\theta}) = \|\mathbf{x} - \boldsymbol{\theta}\|^2$ yields $\hat{\boldsymbol{\theta}} = \bar{\boldsymbol{X}}/\|\bar{\boldsymbol{X}}\|$, the *spherical mean*.

The first to have studied ranks and rank-based methods within the spherical framework were Neeman and Chang (2001), who constructed rank score statistics of the form $\mathbf{T}_{\varphi;\boldsymbol{\theta}}^{(n)} := \sum_{i=1}^{n} \varphi\left(R_i^+/(n+1)\right) \boldsymbol{S_\theta}(\boldsymbol{X}_i)$, where φ is a score generating function and where R_i^+ $(i = 1, \ldots, n)$ stands for the rank of $\|\boldsymbol{X}_i - (\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta}\|$ among the scalars $\|\boldsymbol{X}_1 - (\boldsymbol{X}_1'\boldsymbol{\theta})\boldsymbol{\theta}\|, \ldots, \|\boldsymbol{X}_n - (\boldsymbol{X}_n'\boldsymbol{\theta})\boldsymbol{\theta}\|$. Some years later, Tsai and Sen (2007) developed rank tests (using a similar definition for the ranks R_i^+) for the location problem $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$. Assuming that the observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are independent and identically distributed (i.i.d.) with a rotationally symmetric distribution on the unit sphere, they considered statistics of the form $(\mathbf{T}_{\varphi;\boldsymbol{\theta}_0}^{(n)})'\mathbf{\Gamma_T^{-1}}\mathbf{T}_{\varphi;\boldsymbol{\theta}_0}^{(n)}$, where $\mathbf{\Gamma_T}$ is the asymptotic variance of $\mathbf{T}_{\varphi;\boldsymbol{\theta}_0}^{(n)}$ under the null. They obtained the asymptotic properties of their procedures via permutational central limit theorems.

The purpose of the present work is to propose optimal rank-based estimators (R-estimators) for the spherical location $\boldsymbol{\theta}$. The backbone of our approach is the "Le Cam methodology" (see Le Cam (1986)), which has, to the best of our knowledge, never been used in the framework of spherical statistics. This is

perhaps explained by the curved nature of the parameter space: the unit sphere $\mathcal{S}^{k-1}$ being a non-linear manifold, it typically generates non-traditional Gaussian shift experiments and, as a consequence, the usual arguments behind Le Cam's theory break down in this context.

The first step consists in establishing the *uniform local asymptotic normality* (ULAN) of a sequence of rotationally symmetric models. This we achieve by rewriting $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ in terms of the usual spherical $\boldsymbol{\eta}$-coordinates and showing that ULAN holds for this more common re-parameterization. Although the latter parameterization is not valid uniformly on the whole sphere (its Jacobian is not full-rank everywhere), we then use a recent result of Hallin, Paindaveine, and Verdebout (2010) to show how the ULAN result in the $\boldsymbol{\eta}$-parameterization carries through to the original $\boldsymbol{\theta}$-parameterization.

The second step consists in adapting to the spherical context the method of *one-step R-estimation "à la Le Cam"*, introduced in Hallin, Oja and Paindaveine (2006) in the context of the estimation of a shape matrix of a multivariate elliptical distribution. The ULAN property mentioned above guarantees that the resulting rank-based estimators enjoy the desired optimality properties.

Our estimators constitute attractive alternatives to the traditional Hodges-Lehmann rank-based estimators which have, in general, two major drawbacks. First, they are defined through minimization of a rank-based function which is therefore non-continuous; this complicates greatly the study of their asymptotic properties. Secondly, as shown via a Monte Carlo study in Hallin, Swan, Verdebout, and Veredas (2011) in a regression context, the performances of Hodges-Lehmann-type estimators tend to deteriorate as (the dimension) k increases. As is shown in this paper, one-step R-estimators suffer from neither of these flaws and are even good competitors against the M-estimators studied in the literature.

The outline of the paper is as follows. In Section 2 we recall the definition of the family of rotationally symmetric distributions and prove uniform local asymptotic normality of this family. We devote Section 3 to a description of our rank-based estimators and to the study of their asymptotic properties. In Section 4 we compare our R-estimators with the M-estimators from the literature in terms of Asymptotic Relative Efficiency as well as by means of a Monte Carlo study. In Section 5 we apply our R-estimators to geological data. Finally an Appendix collects the technical proofs.

2. Spherical Model and ULAN

Throughout, the data points $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are assumed to belong to the unit sphere $\mathcal{S}^{k-1}$ of $\mathbb{R}^k$, and to satisfy the following.

Assumption A. $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ *are i.i.d. with common distribution* $P_{\boldsymbol{\theta}; f_1}$ *characterized by a density function (with respect to the usual surface area measure on*

spheres)

$$\mathbf{x} \mapsto f_{\boldsymbol{\theta}}(\mathbf{x}) = c_{k,f_1}\, f_1(\mathbf{x}'\boldsymbol{\theta}), \quad \mathbf{x} \in \mathcal{S}^{k-1}, \tag{2.1}$$

where $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ *is a location parameter and* $f_1 : [-1, 1] \to \mathbb{R}_0^+$ *is absolutely continuous and* (*strictly*) *monotone increasing.*

A function f_1 satisfying Assumption A is called an *angular function*, and we denote by $\mathcal{F}$ the set of angular functions. This choice of terminology reflects the fact that, under Assumption A, the distribution of $\boldsymbol{X}$ depends only on the angle between it and some location $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$. If $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are i.i.d. with density (2.1), then $\boldsymbol{X}_1'\boldsymbol{\theta}, \ldots, \boldsymbol{X}_n'\boldsymbol{\theta}$ are i.i.d. with density

$$t \mapsto \tilde{f}_1(t) := \frac{\omega_k\, c_{k,f_1}}{B(1/2, (k-1)/2)} f_1(t)(1 - t^2)^{(k-3)/2}, \quad -1 \leq t \leq 1,$$

where $\omega_k = 2\pi^{k/2}/\Gamma(k/2)$ is the surface area of $\mathcal{S}^{k-1}$ and $B(\cdot, \cdot)$ is the beta function. The corresponding cdf is denoted by $\tilde{F}_1(t)$.

A special instance of (2.1) is the FVML distribution, obtained by taking angular functions of the form $f_1(t) = \exp(\kappa t) =: f_{1;\exp}(\kappa t)$ for some concentration parameter $\kappa > 0$. Aside from the FVML distribution we also consider spherical distributions with angular functions

$$f_{1;\mathrm{Lin(a)}}(t) := t + a, \quad f_{1;\mathrm{Log(a)}}(t) := \log(t + a), \tag{2.2}$$

and

$$f_{1;\mathrm{Logis}(a,b)}(t) := \frac{a \exp(-b \arccos(t))}{(1 + a \exp(-b \arccos(t)))^2}; \tag{2.3}$$

we refer to these as the *linear*, *logarithmic* and *logistic* spherical distributions, respectively (a and b are constants chosen so that all the above angular functions belong to $\mathcal{F}$).

The rest of this section is devoted to the establishment of the ULAN property of the family $\{\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)} \mid \boldsymbol{\theta} \in \mathcal{S}^{k-1}\}$, where $\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}$ stands for the joint distribution of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$; see the comment just below Proposition 1 for a definition of ULAN. As mentioned in the Introduction, in order to obtain this ULAN property we first need to circumvent a difficulty inherited from the curved nature of the experiment we are considering.

Among the first to have considered such "curved experiments", Chang and Rivest (2001) and Chang (2004) suggest bypassing the problem by reformulating the notion of Fisher information in terms of inner products on the tangent space to $\mathcal{S}^{k-1}$. In this paper, we rather adopt an approach based on recent results from Hallin, Paindaveine, and Verdebout (2010) and, in particular, on the following lemma.

Lemma 2 (Hallin, Paindaveine, and Verdebout (2010)). *Consider a family of probability distributions* $\mathcal{P}^{(n)} = \{\mathrm{P}_{\boldsymbol{\omega}}^{(n)} \mid \boldsymbol{\omega} \in \boldsymbol{\Omega}\}$ *with* $\boldsymbol{\Omega}$ *an open subset of* $\mathbb{R}^{k_1}$ *($k_1 \in \mathbb{N}_0$). Suppose that the parameterization* $\boldsymbol{\omega} \mapsto \mathrm{P}_{\boldsymbol{\omega}}^{(n)}$ *is ULAN for* $\mathcal{P}^{(n)}$ *at some point* $\boldsymbol{\omega}_0 \in \Omega$, *with central sequence* $\boldsymbol{\Delta}_{\boldsymbol{\omega}_0}^{(n)}$ *and Fisher information matrix* $\boldsymbol{\Gamma}_{\boldsymbol{\omega}_0}$. *Let* $\bar{d} : \boldsymbol{\omega} \mapsto \boldsymbol{\vartheta} := \bar{d}(\boldsymbol{\omega})$ *be a continuously differentiable mapping from* $\mathbb{R}^{k_1}$ *to* $\mathbb{R}^{k_2}$ *($k_1 \leq k_2 \in \mathbb{N}_0$) with full column rank Jacobian matrix* $D\bar{d}(\boldsymbol{\omega})$ *at every* $\boldsymbol{\omega}$ *in some neighborhood of* $\boldsymbol{\omega}_0$. *Write* $\boldsymbol{\Theta} := \bar{d}(\boldsymbol{\Omega})$, *and assume that* $\boldsymbol{\vartheta} \mapsto \mathrm{P}_{\boldsymbol{\vartheta}}^{(n);\bar{d}}$, $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$, *provides another parameterization of* $\mathcal{P}^{(n)}$. *Then* $\boldsymbol{\vartheta} \mapsto \mathrm{P}_{\boldsymbol{\vartheta}}^{(n);\bar{d}}$ *is also ULAN for* $\mathcal{P}^{(n)}$ *at* $\boldsymbol{\vartheta}_0 = \bar{d}(\boldsymbol{\omega}_0)$, *with central sequence* $\boldsymbol{\Delta}_{\boldsymbol{\vartheta}_0}^{(n);\bar{d}} = (D^-\bar{d}(\boldsymbol{\omega}_0))'\boldsymbol{\Delta}_{\boldsymbol{\omega}_0}^{(n)}$ *and Fisher information matrix* $\boldsymbol{\Gamma}_{\boldsymbol{\vartheta}_0}^{\bar{d}} = (D^-\bar{d}(\boldsymbol{\omega}_0))'\boldsymbol{\Gamma}_{\boldsymbol{\omega}_0}D^-\bar{d}(\boldsymbol{\omega}_0)$, *where* $D^-\bar{d}(\boldsymbol{\omega}_0) := ((D\bar{d}(\boldsymbol{\omega}_0))'D\bar{d}(\boldsymbol{\omega}_0))^{-1}(D\bar{d}(\boldsymbol{\omega}_0))'$ *is the Moore-Penrose inverse of* $D\bar{d}(\boldsymbol{\omega}_0)$.

We start by establishing ULAN for a re-parameterization of the problem in terms of spherical coordinates. Any vector $\boldsymbol{\theta}$ on the unit sphere of $\mathbb{R}^k$ can be represented via the chart

$$\hbar : \boldsymbol{\eta} := (\eta_1, \ldots, \eta_{k-1})' \in \mathbb{R}^{k-1} \mapsto \hbar(\boldsymbol{\eta}) = \boldsymbol{\theta} = (\cos\eta_1, \sin\eta_1 \, \cos\eta_2, \qquad (2.4)$$
$$\ldots, \sin\eta_1 \cdots \sin\eta_{k-2} \, \cos\eta_{k-1}, \sin\eta_1 \cdots \sin\eta_{k-2} \, \sin\eta_{k-1})',$$

whose Jacobian matrix $D\hbar(\boldsymbol{\eta})$ is given by

$$\begin{pmatrix} -\sin\eta_1 & 0 & \ldots & 0 \\ \cos\eta_1\cos\eta_2 & -\sin\eta_1\sin\eta_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \cos\eta_1\prod_{j=2}^{k-3}\sin\eta_j\cos\eta_{k-2} & \sin\eta_1\cos\eta_2\prod_{j=3}^{k-3}\sin\eta_j\cos\eta_{k-2}\ldots & & 0 \\ \cos\eta_1\prod_{j=2}^{k-2}\sin\eta_j\cos\eta_{k-1} & \sin\eta_1\cos\eta_2\prod_{j=3}^{k-2}\sin\eta_j\cos\eta_{k-1}\ldots & & -\prod_{j=1}^{k-1}\sin\eta_j \\ \cos\eta_1\prod_{j=2}^{k-1}\sin\eta_j & \sin\eta_1\cos\eta_2\prod_{j=3}^{k-1}\sin\eta_j & \ldots\prod_{j=1}^{k-2}\sin\eta_j\cos\eta_{k-1} \end{pmatrix}.$$

The corresponding $\boldsymbol{\eta}$-parameterization $\{\mathrm{P}_{\boldsymbol{\eta};f_1}^{(n);\hbar} \mid \boldsymbol{\eta} \in \mathbb{R}^{k-1}\}$ is simple to construct and enjoys the advantage of a classical parameter space, namely $\mathbb{R}^{k-1}$. This is very convenient for asymptotic calculations; in particular, it is helpful when quadratic mean differentiability and consequently ULAN have to be proved.

Before proceeding to the statement of our results we need the following (essentially technical) assumption.

Assumption B. *Letting* $\varphi_{f_1} := \dot{f}_1/f_1$ *($\dot{f}_1$ is the a.e.-derivative of f_1), the quantity* $\mathcal{J}_k(f_1) := \int_{-1}^{1} \varphi_{f_1}^2(t)(1-t^2)\tilde{f}_1(t)dt < +\infty$.

Assumption B entails that the Fisher information matrix for spherical location is finite (in both the $\boldsymbol{\eta}$- and the original $\boldsymbol{\theta}$-parameterization). More precisely, we show that the information matrix for the $\boldsymbol{\eta}$-parameterization with chart $\hbar$ is of the form

$$\frac{\mathcal{J}_k(f_1)}{k-1} D\hbar(\boldsymbol{\eta})'(\mathbf{I}_k - \hbar(\boldsymbol{\eta})\hbar(\boldsymbol{\eta})')D\hbar(\boldsymbol{\eta})$$

with $\mathbf{I}_k$ standing for the $k \times k$ identity matrix. This matrix can be rewritten as $[\mathcal{J}_k(f_1)/(k-1)]\boldsymbol{\Omega}_{k,\boldsymbol{\eta}}^{\hbar}$, where

$$\begin{aligned}
\boldsymbol{\Omega}_{k,\boldsymbol{\eta}}^{\hbar} &:= \operatorname{diag}(1, \sin^2\eta_1, \sin^2\eta_1\sin^2\eta_2, \ldots, \sin^2\eta_1\cdots\sin^2\eta_{k-2}) \\
&= D\hbar(\boldsymbol{\eta})'D\hbar(\boldsymbol{\eta}) \\
&= D\hbar(\boldsymbol{\eta})'(\mathbf{I}_k - \hbar(\boldsymbol{\eta})\hbar(\boldsymbol{\eta})')D\hbar(\boldsymbol{\eta}),
\end{aligned}$$

with $\operatorname{diag}(a_1, \ldots, a_l)$ denoting a diagonal matrix with diagonal elements $a_1, \ldots,$ a_l. Hence, if $D\hbar(\boldsymbol{\eta})$ is full column rank, then the information matrix $\boldsymbol{\Omega}_{k,\boldsymbol{\eta}}^{\hbar}$ is also full-rank.

With this in hand we are able to establish the ULAN property of the family $\{\mathrm{P}_{\boldsymbol{\eta};f_1}^{(n);\hbar} \mid \boldsymbol{\eta} \in \mathbb{R}^{k-1}\}$ at any point $\boldsymbol{\eta}_0 \in \mathbb{R}^{k-1}$ (note that, clearly, at some points the information matrix is singular due to the rank deficiency of the Jacobian matrix). In order to avoid heavy notation, we drop the index 0 in what follows and write out the LAN property at $\boldsymbol{\eta} \in \mathbb{R}^{k-1}$ in the following proposition (see the Appendix for a proof).

Proposition 1. *Let Assumptions A and B hold. Then the family of probability distributions* $\left\{\mathrm{P}_{\boldsymbol{\eta};f_1}^{(n);\hbar} \mid \boldsymbol{\eta} \in \mathbb{R}^{k-1}\right\}$ *is LAN with central sequence*

$$\boldsymbol{\Delta}_{\boldsymbol{\eta};f_1}^{(n);\hbar} := n^{-1/2}\sum_{i=1}^{n}\varphi_{f_1}(\boldsymbol{X}_i'\hbar(\boldsymbol{\eta}))(1 - (\boldsymbol{X}_i'\hbar(\boldsymbol{\eta}))^2)^{1/2}D\hbar(\boldsymbol{\eta})'\boldsymbol{S}_{\hbar(\boldsymbol{\eta})}(\boldsymbol{X}_i) \qquad (2.5)$$

and Fisher information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\eta};f_1}^{\hbar} := \frac{\mathcal{J}_k(f_1)}{k-1}\boldsymbol{\Omega}_{k,\boldsymbol{\eta}}^{\hbar}. \qquad (2.6)$$

More precisely, for any bounded sequence $\mathbf{e}^{(n)} \in \mathbb{R}^{k-1}$,

$$\log\left(\frac{d\mathrm{P}_{\boldsymbol{\eta}+n^{-1/2}\mathbf{e}^{(n)};f_1}^{(n);\hbar}}{d\mathrm{P}_{\boldsymbol{\eta};f_1}^{(n);\hbar}}\right) = (\mathbf{e}^{(n)})'\boldsymbol{\Delta}_{\boldsymbol{\eta};f_1}^{(n);\hbar} - \frac{1}{2}(\mathbf{e}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\eta};f_1}^{\hbar}\mathbf{e}^{(n)} + o_{\mathrm{P}}(1) \qquad (2.7)$$

and $\boldsymbol{\Delta}_{\boldsymbol{\eta};f_1}^{(n);\hbar} \xrightarrow{\mathcal{L}} \mathcal{N}_{k-1}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\eta};f_1}^{\hbar})$, *both under* $\mathrm{P}_{\boldsymbol{\eta};f_1}^{(n);\hbar}$, *as* $n \to \infty$.

In the sequel, we make use of a slightly reinforced version of Proposition 1, namely the ULAN property of our sequence of models. A sequence of rotationally symmetric models is called ULAN if, for any $\boldsymbol{\eta}^{(n)}$ such that $\boldsymbol{\eta}^{(n)} - \boldsymbol{\eta} = O(n^{-1/2})$ and any bounded sequence $\mathbf{e}^{(n)} \in \mathbb{R}^{k-1}$,

$$\log\left(\frac{d\mathrm{P}^{(n);\hbar}_{\boldsymbol{\eta}^{(n)}+n^{-1/2}\mathbf{e}^{(n)};f_1}}{d\mathrm{P}^{(n);\hbar}_{\boldsymbol{\eta}^{(n)};f_1}}\right) = (\mathbf{e}^{(n)})'\boldsymbol{\Delta}^{(n);\hbar}_{\boldsymbol{\eta}^{(n)};f_1} - \frac{1}{2}(\mathbf{e}^{(n)})'\boldsymbol{\Gamma}^{\hbar}_{\boldsymbol{\eta};f_1}\mathbf{e}^{(n)} + o_{\mathrm{P}}(1)$$

and $\boldsymbol{\Delta}^{(n);\hbar}_{\boldsymbol{\eta}^{(n)};f_1} \xrightarrow{\mathcal{L}} \mathcal{N}_{k-1}(\mathbf{0}, \boldsymbol{\Gamma}^{\hbar}_{\boldsymbol{\eta};f_1})$, both under $\mathrm{P}^{(n);\hbar}_{\boldsymbol{\eta}^{(n)};f_1}$, as $n \to \infty$. Note that the information matrix $\boldsymbol{\Gamma}^{\hbar}_{\boldsymbol{\eta};f_1}$ is uniformly (in $\boldsymbol{\eta}$) the same in the neighborhood $\{\boldsymbol{\eta}^{(n)}, \boldsymbol{\eta}^{(n)} - \boldsymbol{\eta} = O(n^{-1/2})\}$ of $\boldsymbol{\eta}$. In the present setup, it can easily be verified that ULAN holds by using the same arguments as in the proof of Proposition 1.

Note that the diagonality of the Fisher information (2.6) is convenient when working in the $\boldsymbol{\eta}$-parameterization. Indeed it is the structural reason why, when performing asymptotic inference on η_1 for example, the non-specification of the parameters $\eta_2, \ldots, \eta_{k-1}$ will not be responsible for any loss of efficiency. Our focus being on the $\boldsymbol{\theta}$-parameterization, we will not investigate this issue any further.

The $\boldsymbol{\eta}$-parameterization obtained via the chart $\hbar$ suffers from the drawback that the information matrix is singular at several points on the sphere. More precisely, if $\sin \eta_l = 0$ for some $\eta_l \in \{\eta_1, \ldots, \eta_{k-2}\}$, then $\boldsymbol{\Omega}^{\hbar}_{k,\boldsymbol{\eta}}$ is singular. For example, in the 3-dimensional case, the parameter values $\boldsymbol{\theta} = (1, \mathbf{0}')'$ or $\boldsymbol{\theta} = (-1, \mathbf{0}')'$ are very particular; for those values, $\boldsymbol{\Omega}^{\hbar}_{3,\boldsymbol{\eta}} = \mathrm{diag}(1, 0)$. This phenomenon is due to an identification problem: in the general k-dimensional case, when $\eta_1 = 0$ for example, then $\mathrm{P}_{\eta_1, \eta_2, \ldots, \eta_{k-1}; f_1} = \mathrm{P}_{\eta_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_{k-1}; f_1}$ for any $(k-2)$-uples $(\eta_2, \ldots, \eta_{k-1}) \neq (\tilde{\eta}_2, \ldots, \tilde{\eta}_{k-1})$. In the same way, whenever $\eta_j = 0$ for some $j \in \{1, \ldots, k-2\}$, the Jacobian matrix $D\hbar(\boldsymbol{\eta})$ is not full-rank and as a consequence $\boldsymbol{\Gamma}^{\hbar}_{\boldsymbol{\eta};f_1}$ is singular. This singularity – and the identification problem from which it results – is however not structural and only originates in the choice of the chart.

Since every point of a m-dimensional manifold has a neighborhood homeomorphic to an open subset of the m-dimensional space $\mathbb{R}^m$, we see that, for all $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$, one can find a chart $\bar{t} : \mathbb{R}^{k-1} \mapsto \mathcal{S}^{k-1} : \boldsymbol{\eta} \mapsto \boldsymbol{\theta} = \bar{t}(\boldsymbol{\eta})$ with a full column rank Jacobian matrix $D\bar{t}(\boldsymbol{\eta})$ in the vicinity of $\boldsymbol{\eta}$. Making use of Proposition 1, whose proof does not involve the particular form of $\hbar$, we obtain that the family $\{\mathrm{P}^{(n);\bar{t}}_{\boldsymbol{\eta};f_1} \mid \boldsymbol{\eta} \in \mathbb{R}^{k-1}\}$ is ULAN with central sequence

$$\boldsymbol{\Delta}^{(n);\bar{t}}_{\boldsymbol{\eta};f_1} := n^{-1/2} \sum_{i=1}^{n} \varphi_{f_1}(\boldsymbol{X}_i'\bar{t}(\boldsymbol{\eta}))(1 - (\boldsymbol{X}_i'\bar{t}(\boldsymbol{\eta}))^2)^{1/2} D\bar{t}(\boldsymbol{\eta})' \boldsymbol{S}_{\bar{t}(\boldsymbol{\eta})}(\boldsymbol{X}_i)$$

and full-rank Fisher information matrix

$$\mathbf{\Gamma}_{\boldsymbol{\eta};f_1}^{\ell} = \frac{\mathcal{J}_k(f_1)}{k-1}\mathbf{\Omega}_{k,\boldsymbol{\eta}}^{\ell} = \frac{\mathcal{J}_k(f_1)}{k-1}D\ell(\boldsymbol{\eta})'D\ell(\boldsymbol{\eta}).$$

This shows that, for any $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$, it is possible to find a chart ℓ such that, in a neighborhood of $\boldsymbol{\eta} = \ell^{-1}(\boldsymbol{\theta})$, the Jacobian matrix is non-singular and the ULAN property for the related family $\left\{\mathrm{P}_{\boldsymbol{\eta};f_1}^{(n);\ell} \mid \boldsymbol{\eta} \in \mathbb{R}^{k-1}\right\}$ holds. This observation, combined with Lemma 2, finally yields the desired ULAN property for the family $\left\{\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)} \mid \boldsymbol{\theta} \in \mathcal{S}^{k-1}\right\}$ (see the Appendix for a proof).

Let $\boldsymbol{\theta}^{(n)} \in \mathcal{S}^{k-1}$ be such that $\boldsymbol{\theta}^{(n)} - \boldsymbol{\theta} = O(n^{-1/2})$ and consider local alternatives on the sphere of the form $\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}$. For $\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}$ to remain in $\mathcal{S}^{k-1}$, it is necessary that the sequence $\mathbf{t}^{(n)}$ satisfies

$$\begin{aligned} 0 &= (\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)})'(\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}) - 1 \\ &= 2n^{-1/2}(\boldsymbol{\theta}^{(n)})'\mathbf{t}^{(n)} + n^{-1}(\mathbf{t}^{(n)})'\mathbf{t}^{(n)}. \end{aligned} \qquad (2.8)$$

Consequently, $\mathbf{t}^{(n)}$ must be such that $2n^{-1/2}(\boldsymbol{\theta}^{(n)})'\mathbf{t}^{(n)} + n^{-1}(\mathbf{t}^{(n)})'\mathbf{t}^{(n)} = 0$ or, equivalently, such that $2n^{-1/2}(\boldsymbol{\theta}^{(n)})'\mathbf{t}^{(n)} + o(n^{-1/2}) = 0$. Therefore, for $\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}$ to remain in $\mathcal{S}^{k-1}$, $\mathbf{t}^{(n)}$ must belong, up to a $o(n^{-1/2})$ quantity, to the tangent space to $\mathcal{S}^{k-1}$ at $\boldsymbol{\theta}^{(n)}$.

Now, Lemma 2 provides a link between the ULAN properties of two different parameterizations of the same model. This link is directly related to the link between the local alternatives $\boldsymbol{\eta}^{(n)} + n^{-1/2}\mathbf{e}^{(n)}$ in the full-rank parameterization and the local alternatives $\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}$ described above. As shown in (2.8), in order for $\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}$ to belong to $\mathcal{S}^{k-1}$, it is necessary that $\mathbf{t}^{(n)}$ be of the form $\mathbf{t}_*^{(n)} + o(1)$, with $\mathbf{t}_*^{(n)}$ in the tangent space to $\mathcal{S}^{k-1}$ at $\boldsymbol{\theta}^{(n)}$, hence of the form $\mathbf{t}_{**}^{(n)} + o(1)$ with $\mathbf{t}_{**}^{(n)}$ in the tangent space to $\mathcal{S}^{k-1}$ at $\boldsymbol{\theta}$ that is, $\mathbf{t}_{**}^{(n)} = D\ell(\boldsymbol{\eta})\mathbf{e}^{(n)}$ for some bounded sequence $\mathbf{e}^{(n)} \in \mathbb{R}^{k-1}$. It follows from differentiability that, letting $\boldsymbol{\eta}^{(n)} = \ell^{-1}(\boldsymbol{\theta}^{(n)})$,

$$\begin{aligned} \boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)} &= \boldsymbol{\theta}^{(n)} + n^{-1/2}D\ell(\boldsymbol{\eta})\mathbf{e}^{(n)} + o(n^{-1/2}) \\ &= \ell(\boldsymbol{\eta}^{(n)}) + n^{-1/2}D\ell(\boldsymbol{\eta})\mathbf{e}^{(n)} + o(n^{-1/2}) \\ &= \ell(\boldsymbol{\eta}^{(n)} + n^{-1/2}\mathbf{e}^{(n)} + o(n^{-1/2})), \end{aligned} \qquad (2.9)$$

hence there is a clear correspondence between linear perturbations in the $\boldsymbol{\eta}$-parameterization and perturbations on the sphere in the $\boldsymbol{\theta}$-parameterization (through the chart ℓ). As a direct consequence, the local alternatives $\boldsymbol{\theta}^{(n)} + n^{-1/2}\mathbf{t}^{(n)}$ are equivalent to those considered in Tsai (2009). Now, turning to

local log-likelihood ratios, in view of ULAN for the $\boldsymbol{\eta}$-parameterization,

$$\log\left(d\mathrm{P}^{(n)}_{\boldsymbol{\theta}^{(n)}+n^{-1/2}\mathbf{t}^{(n)};f_1}/d\mathrm{P}^{(n)}_{\boldsymbol{\theta}^{(n)};f_1}\right) = \log\left(\frac{d\mathrm{P}^{(n)}_{\boldsymbol{\eta}^{(n)}+n^{-1/2}\mathbf{e}^{(n)}+o(n^{-1/2});f_1}}{d\mathrm{P}^{(n)}_{\boldsymbol{\eta}^{(n)};f_1}}\right)$$

$$= \mathbf{e}^{(n)\prime}\boldsymbol{\Delta}^{(n);l}_{\boldsymbol{\eta}^{(n)};f_1} - \frac{1}{2}\,\mathbf{e}^{(n)\prime}\boldsymbol{\Gamma}^l_{\boldsymbol{\eta};f_1}\mathbf{e}^{(n)} + o_\mathrm{P}(1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\eta}^{(n)};f_1} = \mathrm{P}^{(n)}_{\boldsymbol{\vartheta}^{(n)};f_1}$-probability, as $n \to \infty$. Summing up, we have the following result.

Proposition 2. *Let Assumptions A and B hold. Then the family of probability distributions* $\left\{\mathrm{P}^{(n)}_{\boldsymbol{\theta};f_1} \mid \boldsymbol{\theta} \in \mathcal{S}^{k-1}\right\}$ *is ULAN with central sequence*

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};f_1} := n^{-1/2}\sum_{i=1}^{n}\varphi_{f_1}(\boldsymbol{X}'_i\boldsymbol{\theta})(1-(\boldsymbol{X}'_i\boldsymbol{\theta})^2)^{1/2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$$

and Fisher information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1} := \frac{\mathcal{J}_k(f_1)}{k-1}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}').$$

More precisely, for any $\boldsymbol{\theta}^{(n)} \in \mathcal{S}^{k-1}$ *such that* $\boldsymbol{\theta}^{(n)} - \boldsymbol{\theta} = O(n^{-1/2})$ *and any bounded sequence* $\mathbf{t}^{(n)}$ *as in (2.8), we have*

$$\log\left(\frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}^{(n)}+n^{-1/2}\mathbf{t}^{(n)};f_1}}{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}^{(n)};f_1}}\right) = (\mathbf{t}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};f_1} - \frac{1}{2}(\mathbf{t}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1}\mathbf{t}^{(n)} + o_\mathrm{P}(1)$$

and $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};f_1} \overset{\mathcal{L}}{\to} \mathcal{N}_{k-1}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1})$, *both under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta}^{(n)};f_1}$, *as* $n \to \infty$.

See the proof in the appendix in which the expressions of $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};f_1}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1}$, are obtained.

3. Rank-based Estimation: Optimal *R*-estimators

In this section, we make use of the ULAN property of Proposition 2 to construct semi-parametrically efficient *R*-estimators of $\boldsymbol{\theta}$. To this end, we adapt the Le Cam technique of one-step *R*-estimation, introduced in Hallin, Oja and Paindaveine (2006), to the present context.

Le Cam's one-step *R*-estimation method assumes the existence of a preliminary estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ satisfying some conditions summarized in the following assumption.

Assumption C. *The preliminary estimator $\hat{\boldsymbol{\theta}} \in \mathcal{S}^{k-1}$ is such that*

(i) *$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = O_{\mathrm{p}}(n^{-1/2})$ under $\bigcup_{g_1 \in \mathcal{F}} \mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$.*

(ii) *$\hat{\boldsymbol{\theta}}$ is locally and asymptotically discrete; that is, it only takes a bounded number of distinct values in $\boldsymbol{\theta}$-centered balls with $O(n^{-1/2})$ radius.*

Assumption C(i) requires that the preliminary estimator is root-n consistent under the whole set $\mathcal{F}$ of angular functions. This condition is fulfilled by, e.g., the spherical mean or the spherical median. This uniformity in $g_1 \in \mathcal{F}$ plays an important role for our R-estimation procedures, as we shall see in the sequel. Regarding Assumption C(ii), it should be noted that this discretization condition is a purely technical requirement (see for instance Hallin, Swan, Verdebout, and Veredas (2011)), with little practical implications (in fixed-n practice, such discretizations are irrelevant as the discretization radius can be taken arbitrarily large). Therefore, for the sake of simplicity, we tacitly assume in the sequel that $\hat{\boldsymbol{\theta}}$ satisfies Assumption C(ii).

3.1. Rank-based central sequence and its asymptotic properties

The main idea behind one-step R-estimation consists in adding to the preliminary estimator $\hat{\boldsymbol{\theta}}$ a rank-based quantity which provides optimality under a fixed density. This rank-based quantity appears through a rank-based version $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$ (rigorously defined in (3.2) below) of the parametric central sequence obtained in Proposition 2. A natural requirement for our rank-based central sequence $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$ is its distribution-freeness under the broadest possible family of distributions, say $\mathrm{P}_{\mathrm{broad}}^{(n)}$. A classical way to achieve this goal consists in having recourse to the *invariance principle* which, when $\mathrm{P}_{\mathrm{broad}}^{(n)}$ is invariant under a group of transformations, recommends expressing $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$ in terms of the corresponding maximal invariant. This entails that if, furthermore, this group of transformations is a generating group for $\mathrm{P}_{\mathrm{broad}}^{(n)}$, then $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$ is distribution-free under $\mathrm{P}_{\mathrm{broad}}^{(n)}$.

Clearly, in this rotationally symmetric context, invariance with respect to rotations is crucial, and therefore it seems at first sight natural to study $\mathrm{P}_{\mathrm{broad}}^{(n)} = \bigcup_{\boldsymbol{\theta} \in \mathcal{S}^{k-1}} \mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$ for fixed $g_1 \in \mathcal{F}$ under the effect of the group of rotations. This group is generating for $\bigcup_{\boldsymbol{\theta} \in \mathcal{S}^{k-1}} \mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$, and the ranks R_i^+ defined in the Introduction are rotationally invariant because the scalar products $\boldsymbol{X}_i'\boldsymbol{\theta}$ are invariant with respect to rotations. This is not the case of the multivariate signs $\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$, which are only rotationally equivariant in the sense that $\boldsymbol{S}_{\boldsymbol{\theta}}(\mathbf{O}\boldsymbol{X}_i) = \mathbf{O}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$ for any $\mathbf{O} \in \mathcal{SO}_k$, the class of all $k \times k$ orthogonal matrices. However, since our aim consists in estimating $\boldsymbol{\theta}$, we rather work under fixed-$\boldsymbol{\theta}$ assumptions, hence the family $\mathrm{P}_{\mathrm{broad}}^{(n)}$ makes little sense in the present context. Furthermore, when inference on $\boldsymbol{\theta}$ is considered, the angular function g_1 remains an infinite-dimensional nuisance,

which indicates that statistics that are invariant, and therefore distribution-free under $\bigcup_{g_1 \in \mathcal{F}} P_{\boldsymbol{\theta};g_1}^{(n)}$ ($\boldsymbol{\theta}$ fixed), should in fact be taken into account, if possible without losing invariance properties with respect to the group of rotations.

Fix $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ and consider $P_{\text{broad}}^{(n)} = \bigcup_{g_1 \in \mathcal{F}} P_{\boldsymbol{\theta};g_1}^{(n)}$. We obviously have that $\boldsymbol{X}_i = (\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta} + \sqrt{1 - (\boldsymbol{X}_i'\boldsymbol{\theta})^2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$ by definition of the multivariate signs. Now, let $\mathcal{G}_h^{(n)}$ be the group of transformations of the form $g_h^{(n)} : (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \mapsto (g_h(\boldsymbol{X}_1), \ldots, g_h(\boldsymbol{X}_n))$ with

$$g_h(\boldsymbol{X}_i) := h(\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta} + \sqrt{1 - h(\boldsymbol{X}_i'\boldsymbol{\theta})^2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i), \quad i = 1, \ldots, n, \tag{3.1}$$

where $h : [-1, 1] \to [-1, 1]$ is a monotone continuous nondecreasing function such that $h(1) = 1$ and $h(-1) = -1$. For any $g_h^{(n)} \in \mathcal{G}_h^{(n)}$, it is easy to verify that $\|g_h(\boldsymbol{X}_i)\| = 1$; this means that $g_h^{(n)} \in \mathcal{G}_h^{(n)}$ is a monotone transformation from $(\mathcal{S}^{k-1})^n$ to $(\mathcal{S}^{k-1})^n$. It is quite straightforward to see that the group $\mathcal{G}_h^{(n)}$ is a generating group of the family of distributions $\bigcup_{g_1 \in \mathcal{F}} P_{\boldsymbol{\theta};g_1}^{(n)}$.

These considerations naturally raise the question of finding the maximal invariant associated with $\mathcal{G}_h^{(n)}$. Note that the definition of the mapping g_h in (3.1), together with the fact that $\boldsymbol{\theta}'\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i) = 0$, entails that

$$g_h(\boldsymbol{X}_i) - (g_h(\boldsymbol{X}_i)'\boldsymbol{\theta})\boldsymbol{\theta} = \sqrt{1 - h(\boldsymbol{X}_i'\boldsymbol{\theta})^2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i).$$

Then, we readily obtain that, for $i = 1, \ldots, n$,

$$\begin{aligned}
\boldsymbol{S}_{\boldsymbol{\theta}}(g_h(\boldsymbol{X}_i)) &= \frac{g_h(\boldsymbol{X}_i) - (g_h(\boldsymbol{X}_i)'\boldsymbol{\theta})\boldsymbol{\theta}}{\|g_h(\boldsymbol{X}_i) - (g_h(\boldsymbol{X}_i)'\boldsymbol{\theta})\boldsymbol{\theta}\|} \\
&= \frac{\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)}{\|\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)\|} = \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i).
\end{aligned}$$

This shows that the vector of signs $\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_1), \ldots, \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_n)$ is invariant under the action of $\mathcal{G}_h^{(n)}$. However, it is not a maximal invariant; indeed, the latter is in most semiparametric setups composed of signs *and* ranks (see for instance Hallin and Paindaveine (2006)). Now, the ranks R_i^+ are not invariant under the action of $\mathcal{G}_h^{(n)}$: one can easily find a monotone transformation $g_{\tilde{h}}$ as in (3.1) such that the vector of ranks R_i^+ computed from $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ differs from the vector of ranks R_i^+ computed from $g_{\tilde{h}}(\boldsymbol{X}_1), \ldots, g_{\tilde{h}}(\boldsymbol{X}_n)$. Thus, we need a different concept of ranks in order to build a rank-based version of our central sequence. For this purpose, define, for all $i = 1, \ldots, n$, R_i as the rank of $\boldsymbol{X}_i'\boldsymbol{\theta}$ among $\boldsymbol{X}_1'\boldsymbol{\theta}, \ldots, \boldsymbol{X}_n'\boldsymbol{\theta}$. Since in general, the ranks are invariant with respect to monotone transformations, noting that $g_h(\boldsymbol{X}_i)'\boldsymbol{\theta} = h(\boldsymbol{X}_i'\boldsymbol{\theta})$ directly entails the invariance of these new ranks under the action of the group $\mathcal{G}_h^{(n)}$. The maximal invariant associated with $\mathcal{G}_h^{(n)}$ is, therefore, the vector of signs $\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_1), \ldots, \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_n)$ and ranks $R_1, \ldots, R_n$.

While it is preferable to use the ranks R_i when invariance with respect to $\mathcal{G}_h^{(n)}$ is required, there exist situations in which the ranks R_i^+ are appealing, e.g. when the angular functions are of the form $f_1(t) = \exp(g(t))$ with $g(-t) = -g(t)$ (since then the score function $\varphi_{f_1}(t) = \dot{g}(t)$ in (2.5) is symmetric), see Tsai and Sen (2007).

As explained above, it follows that any statistic measurable with respect to the signs $\mathbf{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$ and ranks R_i is distribution-free under $\bigcup_{g_1 \in \mathcal{F}} \mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$. In accordance with these findings, we choose to base our inference procedures on the following sign- and rank-based version of the parametric central sequence obtained in Proposition 2:

$$\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)} := n^{-1/2} \sum_{i=1}^{n} K\left(\frac{R_i}{n+1}\right) \mathbf{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i), \tag{3.2}$$

where K is a score function satisfying

Assumption D. *The score function K is a continuous function from $[0,1]$ to $\mathbb{R}$.*

Note that all the score functions associated with the densities in (2.3) satisfy this assumption.

In the following result (the proof is given in the Appendix), we derive the asymptotic properties of $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$ under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$ and under contiguous alternatives $\mathrm{P}_{\boldsymbol{\theta}+n^{-1/2}\mathbf{t}^{(n)};g_1}^{(n)}$ (where $\mathbf{t}^{(n)}$ is a bounded sequence as described in (2.8)) for some angular function $g_1 \in \mathcal{F}$.

Proposition 3. *Let Assumptions A, B and D hold. Then the rank-based central sequence $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$*

(i) *is such that $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)} - \boldsymbol{\Delta}_{\boldsymbol{\theta};K,g_1}^{(n)} = o_{\mathrm{P}}(1)$ under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$ as $n \to \infty$, where ($\tilde{G}_1$ stands for the common cdf of the $\boldsymbol{X}_i'\boldsymbol{\theta}$'s under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$)*

$$\boldsymbol{\Delta}_{\boldsymbol{\theta};K,g_1}^{(n)} := n^{-1/2} \sum_{i=1}^{n} K\left(\tilde{G}_1(\boldsymbol{X}_i'\boldsymbol{\theta})\right) \mathbf{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i).$$

Hence, for $K(u) = K_{f_1}(u) = \varphi_{f_1}(\tilde{F}_1^{-1}(u))(1-(\tilde{F}_1^{-1}(u))^2)^{1/2}$, $\underset{\sim}{\boldsymbol{\Delta}}_{\boldsymbol{\theta};K}^{(n)}$ is asymptotically equivalent to the efficient central sequence $\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)}$ for spherical location under $\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}$.

(ii) *is asymptotically normal under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$ with mean zero and covariance matrix $\boldsymbol{\Gamma}_{\boldsymbol{\theta};K} := [(\mathcal{J}_k(K))/(k-1)](\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')$, where $\mathcal{J}_k(K) := \int_0^1 K^2(u)du$.*

(iii) *is asymptotically normal under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\mathbf{t}^{(n)};g_1}$ *with mean* $\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}$ *(where*
$\mathbf{t} := \lim_{n\to\infty}\mathbf{t}^{(n)}$*) and covariance matrix* $\boldsymbol{\Gamma}_{\boldsymbol{\theta};K}$*, where, for* $\mathcal{J}_k(K,g_1) :=$
$\int_0^1 K(u)K_{g_1}(u)du$,

$$\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1} := \frac{\mathcal{J}_k(K,g_1)}{k-1}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}').$$

(iv) *satisfies, under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ *as* $n \to \infty$*, the asymptotic linearity property*

$$\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\mathbf{t}^{(n)};K} - \underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta};K} = -\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}^{(n)} + o_\mathrm{P}(1)$$

for any bounded sequence $\mathbf{t}^{(n)}$ *as described in* (2.8).

The main point of this proposition is part (iv); the preceding three results are necessary for proving that last part. However, they have an interest *per se* as they give the asymptotic properties of the rank-based central sequence $\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta};K}$. Moreover, these results happen to be important when the focus lies on testing procedures for the location parameter $\boldsymbol{\theta}$. This is part of ongoing research. For the present paper, we are mainly interested in the asymptotic linearity property of $\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta};K}$, more precisely on the asymptotic linearity property of

$$\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\hat{\boldsymbol{\theta}};K} := n^{-1/2}\sum_{i=1}^n K\left(\frac{\hat{R}_i}{n+1}\right)\boldsymbol{S}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{X}_i),$$

where $\hat{R}_i$, $i = 1,\ldots,n$, stands for the rank of $\boldsymbol{X}_i'\hat{\boldsymbol{\theta}}$ among $\boldsymbol{X}_1'\hat{\boldsymbol{\theta}},\ldots,\boldsymbol{X}_n'\hat{\boldsymbol{\theta}}$. It is precisely here that Assumption C comes into play: it ensures that the asymptotic linearity property of Proposition 3(iv) holds after replacement of $\mathbf{t}^{(n)}$ by the random quantity $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ (this can be seen via Lemma 4.4 of Kreiss (1987)), which eventually entails that

$$\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\hat{\boldsymbol{\theta}};K} - \underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta};K} = -\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + o_\mathrm{P}(1) \tag{3.3}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ for $g_1 \in \mathcal{F}$. We draw the reader's attention to the fact that (2.8) is satisfied when $\mathbf{t}^{(n)}$ is replaced by $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$.

3.2. Properties of our *R*-estimators

Keeping the notation $\mathbf{A}^-$ for the Moore-Penrose inverse of some matrix $\mathbf{A}$, let

$$\begin{aligned}
\tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)} &:= \hat{\boldsymbol{\theta}} + n^{-1/2}\boldsymbol{\Gamma}^-_{\hat{\boldsymbol{\theta}};K,g_1}\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\hat{\boldsymbol{\theta}};K} \\
&= \hat{\boldsymbol{\theta}} + n^{-1/2}\frac{(k-1)}{\mathcal{J}_k(K,g_1)}(\mathbf{I}_k - \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}')\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\hat{\boldsymbol{\theta}};K} \\
&= \hat{\boldsymbol{\theta}} + n^{-1/2}\frac{(k-1)}{\mathcal{J}_k(K,g_1)}\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\hat{\boldsymbol{\theta}};K},
\end{aligned}$$

where the last inequality holds since $\hat{\boldsymbol{\theta}}'\mathbf{S}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{X}_i) = 0$ for all $i = 1, \ldots, n$. The expression of $\tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}$ is quite traditional when one-step estimation is considered. However, using $\tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}$ itself to estimate $\boldsymbol{\theta}$ is clearly unnatural since $\tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}$ does not belong to $\mathcal{S}^{k-1}$ in general. This is why we propose the one-step R-estimator

$$\hat{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)} := \tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}/\|\tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}\| \in \mathcal{S}^{k-1},$$

which is a normalized version of $\tilde{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}$. As is shown in the sequel, the normalization has no asymptotic cost on the efficiency of the one-step method.

Nevertheless, $\hat{\boldsymbol{\theta}}_{K;\mathcal{J}_k(K,g_1)}$ is not a genuine estimator because it is still a function of the unknown scalar $\mathcal{J}_k(K, g_1)$. This cross-information quantity requires consistent estimation in order to ensure asymptotic normality of our R-estimators. To tackle this problem, we adopt here the idea developed in Hallin, Oja and Paindaveine (2006) based on the ULAN property of the model. Let $\tilde{\boldsymbol{\theta}}(\beta) := \hat{\boldsymbol{\theta}} + n^{-1/2}\beta(k-1)\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)}$, $\hat{\boldsymbol{\theta}}(\beta) := \tilde{\boldsymbol{\theta}}(\beta)/\|\tilde{\boldsymbol{\theta}}(\beta)\|$ and consider the quadratic form

$$\beta \mapsto h^{(n)}(\beta) := (\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)})'\boldsymbol{\Gamma}_{\hat{\boldsymbol{\theta}};K}^{-}\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}}(\beta);K}^{(n)}$$

$$= \frac{(k-1)}{\mathcal{J}_k(K)}(\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)})'\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}}(\beta);K}^{(n)}. \qquad (3.4)$$

Then, part (iv) of Proposition 3 and the root-n consistency of $\hat{\boldsymbol{\theta}}(\beta)$ (which follows from the root-n consistency of $\tilde{\boldsymbol{\theta}}(\beta)$ and the Delta method applied to the mapping $\mathbf{x} \mapsto \mathbf{x}/\|\mathbf{x}\|$) imply that, after some direct computations involving (3.3),

$$\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}}(\beta);K}^{(n)} - \underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)} = -\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}n^{1/2}(\hat{\boldsymbol{\theta}}(\beta) - \hat{\boldsymbol{\theta}}) + o_{\mathrm{P}}(1)$$

as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$. Moreover, it is clear that $\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}n^{1/2}(\hat{\boldsymbol{\theta}}(\beta) - \hat{\boldsymbol{\theta}}) = \boldsymbol{\Gamma}_{\hat{\boldsymbol{\theta}};K,g_1}n^{1/2}(\hat{\boldsymbol{\theta}}(\beta) - \hat{\boldsymbol{\theta}}) + o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$. These facts combined with the definition of $\hat{\boldsymbol{\theta}}(\beta)$ entail that, under $\mathrm{P}_{\boldsymbol{\theta};g_1}^{(n)}$ and for $n \to \infty$,

$$h^{(n)}(\beta) = \frac{(k-1)}{\mathcal{J}_k(K)}(\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)})'\left(\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)} - \boldsymbol{\Gamma}_{\hat{\boldsymbol{\theta}};K,g_1}n^{1/2}(\hat{\boldsymbol{\theta}}(\beta) - \hat{\boldsymbol{\theta}})\right) + o_{\mathrm{P}}(1)$$

$$= \frac{(k-1)}{\mathcal{J}_k(K)}(\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)})'\left(\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)} - \mathcal{J}_k(K, g_1)\beta\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)}\right) + o_{\mathrm{P}}(1)$$

$$= \frac{(k-1)(1 - \mathcal{J}_k(K, g_1)\beta)}{\mathcal{J}_k(K)}(\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)})'\underset{\sim}{\boldsymbol{\Delta}}_{\hat{\boldsymbol{\theta}};K}^{(n)} + o_{\mathrm{P}}(1),$$

where the passage from the first to the second line requires some computations involving the Delta method; for the sake of readability we dispense with these

details here, as they follow along the same lines as those achieved below. In view of (3.4), $h^{(n)}(\beta)$ can be rewritten as

$$h^{(n)}(\beta) = (1 - \mathcal{J}_k(K, g_1)\beta)h^{(n)}(0) + o_{\mathrm{P}}(1)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ as $n \to \infty$. Since $h^{(n)}(0) > 0$, one obtains (the proof is along the same lines as in Hallin, Oja and Paindaveine (2006)) a consistent estimator of $(\mathcal{J}_k(K, g_1))^{-1}$ given by $\hat{\beta} := \inf\{\beta > 0 : h^{(n)}(\beta) < 0\}$. Therefore, $\hat{\mathcal{J}}_k(K, g_1) := \hat{\beta}^{-1}$ provides a consistent estimator of the cross-information quantity, and a genuine estimator of $\boldsymbol{\theta}$ is provided by $\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}$. Now, the Delta method and some easy computations show that under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ and for $n \to \infty$

$$n^{1/2}(\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)} - \boldsymbol{\theta}) = n^{1/2}\left(\frac{\tilde{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}}{\|\tilde{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}\|} - \frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}\right)$$

$$= n^{1/2}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\left(\tilde{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)} - \boldsymbol{\theta}\right) + o_{\mathrm{P}}(1), \quad (3.5)$$

where $\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}'$ is the Jacobian matrix of the mapping $\mathbf{x} \mapsto \mathbf{x}/\|\mathbf{x}\|$ evaluated at $\boldsymbol{\theta}$. Then, (3.5), the definition of $\tilde{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}$, part (iv) of Proposition 3, the consistency of $\hat{\mathcal{J}}_k(K, g_1)$, and Assumption C entail that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ and as $n \to \infty$,

$$n^{1/2}(\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)} - \boldsymbol{\theta})$$
$$= n^{1/2}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\left(\tilde{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)} - \boldsymbol{\theta}\right) + o_{\mathrm{P}}(1)$$
$$= n^{1/2}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\left(\hat{\boldsymbol{\theta}} + n^{-1/2}\boldsymbol{\Gamma}^{-}_{\hat{\boldsymbol{\theta}};K,g_1}\boldsymbol{\Delta}^{(n)}_{\hat{\boldsymbol{\theta}};K} - \boldsymbol{\theta}\right) + o_{\mathrm{P}}(1)$$
$$= (\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\left(n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \boldsymbol{\Gamma}^{-}_{\boldsymbol{\theta};K,g_1}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K} - n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right) + o_{\mathrm{P}}(1)$$
$$= (\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\boldsymbol{\Gamma}^{-}_{\boldsymbol{\theta};K,g_1}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K} + o_{\mathrm{P}}(1)$$
$$= \frac{(k-1)}{\mathcal{J}_k(K,g_1)}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K} + o_{\mathrm{P}}(1), \quad (3.6)$$

where the passage from the second to the third line draws upon (3.3). We obtain the following result which summarizes the asymptotic properties of $\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}$.

Proposition 4. *Let Assumptions A, B, C and D hold. Then,*

(i) *under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$, $n^{1/2}(\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)} - \boldsymbol{\theta})$ *is asymptotically normal with mean* $\mathbf{0}$ *and covariance matrix* $(k-1)(\mathcal{J}_k(K)/\mathcal{J}_k^2(K,g_1))(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')$;

(ii) *for* $K(u) = K_{f_1}(u) = \varphi_{f_1}(\tilde{F}_1^{-1}(u))(1 - (\tilde{F}_1^{-1}(u))^2)^{1/2}$, *the estimator* $\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,f_1)}$ *is semi-parametrically efficient under* $\mathrm{P}^{(n)}_{\boldsymbol{\theta};f_1}$.

While part (i) of Proposition 4 is a direct consequence of (3.6), part (ii) requires explanation. From (3.6), we obviously have that under $\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}$,

$$n^{1/2}(\hat{\boldsymbol{\theta}}_{K_{f_1};\hat{\mathcal{J}}_k(K_{f_1},f_1)} - \boldsymbol{\theta}) = \frac{(k-1)}{\mathcal{J}_k(K_{f_1},f_1)}\boldsymbol{\Delta}_{\boldsymbol{\theta};K_{f_1}}^{(n)} + o_{\mathrm{P}}(1) \qquad (3.7)$$

as $n \to \infty$. Now, using the identities $\mathcal{J}_k(K_{f_1},f_1) = \mathcal{J}_k(f_1)$ and $\boldsymbol{\Delta}_{\boldsymbol{\theta};K_{f_1},f_1}^{(n)} = \boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)}$, we obtain via part (i) of Proposition 3 that

$$\frac{(k-1)}{\mathcal{J}_k(K_{f_1},f_1)}\boldsymbol{\Delta}_{\boldsymbol{\theta};K_{f_1}}^{(n)} = \frac{(k-1)}{\mathcal{J}_k(f_1)}\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1,}^{(n)} + o_{\mathrm{P}}(1) \qquad (3.8)$$

under $\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}$ as $n \to \infty$. Combining (3.7) and (3.8) (and using again the fact that $\boldsymbol{\theta}'\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}) = 0$), it follows that

$$n^{1/2}(\hat{\boldsymbol{\theta}}_{K_{f_1};\hat{\mathcal{J}}_k(K_{f_1},f_1)} - \boldsymbol{\theta}) = \frac{(k-1)}{\mathcal{J}_k(f_1)}\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)} + o_{\mathrm{P}}(1) = \boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1}^{-}\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)} + o_{\mathrm{P}}(1)$$

still under $\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}$ as $n \to \infty$. This entails that $\hat{\boldsymbol{\theta}}_{K_{f_1};\hat{\mathcal{J}}_k(K_{f_1},f_1)}$ is asymptotically efficient under $\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}$, which then yields the desired optimality properties.

Finally note that if the preliminary estimator $\hat{\boldsymbol{\theta}}$ is rotation-equivariant, meaning that $\hat{\boldsymbol{\theta}}(\mathbf{O}\boldsymbol{X}_i) = \mathbf{O}\hat{\boldsymbol{\theta}}(\boldsymbol{X}_i)$ for any matrix $\mathbf{O} \in \mathcal{SO}_k$, then $\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}(\mathbf{O}\boldsymbol{X}_i) = \mathbf{O}\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}(\boldsymbol{X}_i)$. This implies that rotation-equivariance of our R-estimators is inherited from the preliminary estimator.

4. Asymptotic Relative Efficiencies and Simulation Results

In this section, we compare the asymptotic and finite-sample performances of the proposed R-estimators with those of the spherical mean and spherical median. The asymptotic performances are analyzed in Section 4.1 on the basis of asymptotic relative efficiencies (ARE), while the finite-sample behavior is investigated in Section 4.2 by means of a Monte Carlo study.

4.1. Asymptotic relative efficiencies

Traditional estimators of $\boldsymbol{\theta}$ in the literature are the M-estimators defined, for a given function $\rho_0(\mathbf{x};\boldsymbol{\theta})$, as the value $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ which minimizes the objective function

$$\boldsymbol{\theta} \mapsto \rho(\boldsymbol{\theta}) := \sum_{i=1}^{n} \rho_0(\boldsymbol{X}_i;\boldsymbol{\theta}),$$

where $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n$ are spherical observations. Letting $\rho_0(\mathbf{x};\boldsymbol{\theta}) =: \tilde{\rho}(\mathbf{x}'\boldsymbol{\theta})$, special instances are the spherical mean $\hat{\boldsymbol{\theta}}_{\mathrm{Mean}}$ (maximum likelihood estimator under

FVML distributions, obtained by taking $\psi(t) := -\dot{\tilde{\rho}}(t) = 2)$ and the spherical median $\hat{\boldsymbol{\theta}}_{\text{Median}}$ (Fisher (1985), obtained by taking $\psi(t) = (1 - t^2)^{-1/2}$). By Theorem 3.2 in Chang (2004), an M-estimator $\hat{\boldsymbol{\theta}}_M$ associated with the objective function $\tilde{\rho}(\mathbf{x}'\boldsymbol{\theta})$ is such that $n^{1/2}(\hat{\boldsymbol{\theta}}_M - \boldsymbol{\theta})$ is asymptotically normal with mean zero and covariance matrix

$$(k-1)\frac{\text{E}[\psi^2(\boldsymbol{X}'\boldsymbol{\theta})(1-(\boldsymbol{X}'\boldsymbol{\theta})^2)]}{\text{E}^2[\psi(\boldsymbol{X}'\boldsymbol{\theta})\varphi_{g_1}(\boldsymbol{X}'\boldsymbol{\theta})(1-(\boldsymbol{X}'\boldsymbol{\theta})^2)]}\left(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}'\right)$$

under $\text{P}_{\boldsymbol{\theta};g_1}^{(n)}$ (expectations are taken under $\text{P}_{\boldsymbol{\theta};g_1}^{(n)}$). Now, let $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ be two estimators of the spherical location such that $n^{1/2}(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta})$ and $n^{1/2}(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta})$ are asymptotically normal with mean zero and covariance matrices $\rho_1(\mathbf{I} - \boldsymbol{\theta}\boldsymbol{\theta}')$ and $\rho_2(\mathbf{I} - \boldsymbol{\theta}\boldsymbol{\theta}')$, respectively, under $\text{P}_{\boldsymbol{\theta};g_1}^{(n)}$. Then, a natural way to compare the asymptotic efficiencies of $\hat{\boldsymbol{\theta}}_1$ and $\hat{\boldsymbol{\theta}}_2$ (still under $\text{P}_{\boldsymbol{\theta};g_1}^{(n)}$) is through the ratio $\text{ARE}_{\boldsymbol{\theta};g_1}(\hat{\boldsymbol{\theta}}_1/\hat{\boldsymbol{\theta}}_2) = \rho_1/\rho_2$ which we refer to as an ARE. The following result provides a general formula for ARE between an R-estimator and an M-estimator.

Proposition 5. *Let* $\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}$ *be the R-estimator associated with the score function K and let $\hat{\boldsymbol{\theta}}_M$ be the M-estimator associated with the objective function $\rho(\mathbf{x}'\boldsymbol{\theta})$. Then*

$$\text{ARE}_{\boldsymbol{\theta};g_1}\Big(\frac{\hat{\boldsymbol{\theta}}_{K;\hat{\mathcal{J}}_k(K,g_1)}}{\hat{\boldsymbol{\theta}}_M}\Big) = \frac{\text{E}[\psi^2(\boldsymbol{X}'\boldsymbol{\theta})(1-(\boldsymbol{X}'\boldsymbol{\theta})^2)]\mathcal{J}_k^2(K,g_1)}{\text{E}^2[\psi(\boldsymbol{X}'\boldsymbol{\theta})\varphi_{g_1}(\boldsymbol{X}'\boldsymbol{\theta})(1-(\boldsymbol{X}'\boldsymbol{\theta})^2)]\mathcal{J}_k(K)},$$

where $\text{ARE}_{\boldsymbol{\theta};g_1}$ *denotes the asymptotic relative efficiency under* $\text{P}_{\boldsymbol{\theta};g_1}^{(n)}$.

In Tables 1 and 2, we collect numerical values of $\text{ARE}_{\boldsymbol{\theta};g_1}$ for $k = 3$ under various underlying rotationally symmetric densities, namely those described below Assumption A in Section 2. Several one-step R-estimators are compared to both the spherical mean and the spherical median: $\hat{\boldsymbol{\theta}}_{\text{FVML}(2)}$ and $\hat{\boldsymbol{\theta}}_{\text{FVML}(6)}$, based on FVML scores (with $\kappa = 2$ and $\kappa = 6$, respectively), $\hat{\boldsymbol{\theta}}_{\text{Lin}(2)}$ and $\hat{\boldsymbol{\theta}}_{\text{Lin}(4)}$ based on linear scores (associated with linear angular densities with $a = 2$ and $a = 4$, respectively, see (2.2)), $\hat{\boldsymbol{\theta}}_{\text{Log}(2.5)}$ based on logarithmic scores (associated with a logarithmic angular density with $a = 2.5$, see (2.2)) and finally $\hat{\boldsymbol{\theta}}_{\text{Logis}(1,1)}$ and $\hat{\boldsymbol{\theta}}_{\text{Logis}(2,1)}$ based on logistic scores (associated with logistic angular densities with respectively $a = 1$, $b = 1$ and $a = 2$, $b = 1$, see (2.3)).

Inspection of Tables 1 and 2 confirms the theoretical results obtained previously. When based on the score function associated with the underlying density, R-estimators are optimal. For example, $\hat{\boldsymbol{\theta}}_{\text{Lin}(2)}$ is the most precise estimator under $\text{P}_{\boldsymbol{\theta};f_{1;\text{Lin}(2)}}^{(n)}$. As expected the spherical mean dominates the R-estimators under

FVML densities since it is the maximum likelihood estimator in this case. For example, $\hat{\boldsymbol{\theta}}_{\text{FVML}(2)}$ is only just less or equally (under the FVML(2) density) efficient than the spherical mean under FVML densities but performs nicely under other densities. In general, the proposed R-estimators outperform the spherical median.

Table 1. Asymptotic relative efficiencies of R-estimators with respect to the spherical mean under various 3-dimensional rotationally symmetric densities.

	AREs with respect to the spherical mean ($\text{ARE}(\hat{\boldsymbol{\theta}}_K/\hat{\boldsymbol{\theta}}_{\text{Mean}})$)							
Underlying density	$\hat{\boldsymbol{\theta}}_{\text{FVML}(2)}$	$\hat{\boldsymbol{\theta}}_{\text{FVML}(6)}$	$\hat{\boldsymbol{\theta}}_{\text{Lin}(2)}$	$\hat{\boldsymbol{\theta}}_{\text{Lin}(4)}$	$\hat{\boldsymbol{\theta}}_{\text{Log}(2.5)}$	$\hat{\boldsymbol{\theta}}_{\text{Logis}(1,1)}$	$\hat{\boldsymbol{\theta}}_{\text{Logis}(2,1)}$	$\hat{\boldsymbol{\theta}}_{\text{Sq}(1.1)}$
FVML(1)	0.9744	0.8787	0.9813	0.9979	0.9027	0.9321	0.7364	0.7804
FVML(2)	1	0.9556	0.9978	0.9586	0.9749	0.9823	0.8480	0.8932
FVML(6)	0.9555	1	0.9381	0.8517	0.9768	0.9911	0.9280	0.9771
Lin(2)	1.0539	0.9909	1.0562	1.0215	1.0212	1.0247	0.8796	0.9174
Lin(4)	0.9709	0.8627	0.9795	1.0128	0.8856	0.9231	0.7097	0.7083
Log(2.5)	1.1610	1.1633	1.1514	1.0413	1.1908	1.1625	1.0951	1.1376
Log(4)	1.0182	0.9216	1.0261	1.0347	0.9503	0.9741	0.7851	0.8226
Logis(1,1)	1.0768	1.0865	1.0635	0.9991	1.0701	1.0962	0.9778	1.0349
Logis(2,1)	1.3182	1.4426	1.2946	1.0893	1.4294	1.3865	1.5544	1.4680
Sq(1.1)	1.2303	1.3460	1.1964	1.0264	1.3158	1.3004	1.3009	1.3774

4.2. Monte Carlo study

We now discuss the finite-sample behavior of different R-estimators of the spherical location. For this purpose, we generated $M = 1,000$ samples from various 3-dimensional ($k = 3$) rotationally symmetric distributions: the FVML(2) and FVML(4) distributions, the linear distribution with $a = 2$ and $a = 4$ (see (2.2)) and the *square root* distribution Sq(1.1) associated with an angular density of the form

$$f_1(t) := \sqrt{t + a}, \quad \text{with } a = 1.1.$$

The true location parameter is $\boldsymbol{\theta} = (\sqrt{2}/2, \sqrt{2}/2, 0)'$. For each replication, the spherical median $\hat{\boldsymbol{\theta}}_{\text{Median}}$, the spherical mean $\hat{\boldsymbol{\theta}}_{\text{Mean}}$, the FVML-score based R-estimators $\hat{\boldsymbol{\theta}}_{\text{FVML}(2)}$ and $\hat{\boldsymbol{\theta}}_{\text{FVML}(4)}$, the linear-score based R-estimators $\hat{\boldsymbol{\theta}}_{\text{Lin}(2)}$ and $\hat{\boldsymbol{\theta}}_{\text{Lin}(4)}$, and the Sq(1.1)-score based R-estimator $\hat{\boldsymbol{\theta}}_{\text{Sq}(1.1)}$ were computed. The preliminary estimators used in the construction of all our one-step R-estimators were $\hat{\boldsymbol{\theta}}_{\text{Median}}$ and $\hat{\boldsymbol{\theta}}_{\text{Mean}}$.

In Table 3, we report the Euclidean norm of $\mathbf{m} = (m_1, m_2, m_3)'$ where, letting $\hat{\theta}_i^{(j)}$ stand for the ith component of an estimator computed from the jth replication and θ_i for the ith component of $\boldsymbol{\theta}$,

$$m_i := \frac{1}{M} \sum_{j=1}^{M} (\hat{\theta}_i^{(j)} - \theta_i)^2, \ i = 1, 2, 3,$$

Table 2. Asymptotic relative efficiencies of *R*-estimators with respect to the spherical median under various 3-dimensional rotationally symmetric densities.

Underlying density	AREs with respect to the spherical median ($\mathrm{ARE}(\hat{\boldsymbol{\theta}}_K/\hat{\boldsymbol{\theta}}_{\mathrm{Median}})$)							
	$\hat{\boldsymbol{\theta}}_{\mathrm{FVML}(2)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{FVML}(6)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{Lin}(2)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{Lin}(4)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{Log}(2.5)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{Logis}(1,1)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{Logis}(2,1)}$	$\hat{\boldsymbol{\theta}}_{\mathrm{Sq}(1.1)}$
FVML(1)	1.0691	0.9641	1.0766	1.0949	0.9904	1.0226	0.8079	0.8562
FVML(2)	1.1321	1.0819	1.1297	1.0853	1.1038	1.1121	0.9601	1.0112
FVML(6)	1.1632	1.2174	1.1421	1.0369	1.1891	1.2065	1.1297	1.1895
Lin(2)	1.1391	1.0709	1.1415	1.1041	1.1037	1.1075	0.9507	0.9915
Lin(4)	1.0493	0.9324	1.0587	1.0946	0.9571	0.9977	0.7671	0.8157
Log(2.5)	1.2171	1.2195	1.2071	1.0917	1.2484	1.2188	1.1481	1.1926
Log(4)	1.0900	0.9865	1.0984	1.1076	1.0173	1.0427	0.8403	1.0228
Logis(1,1)	1.1264	1.1365	1.1125	1.0451	1.1194	1.1467	1.0228	1.0826
Logis(2,1)	1.4497	1.5868	1.4238	1.1980	1.5721	1.5249	1.7095	1.6145
Sq(1.1)	1.2282	1.3436	1.1943	1.0246	1.3135	1.2981	1.2986	1.3750

computed still for each of the aforementioned estimators and sample sizes $n = 100$, $n = 500$ and $n = 1,000$. Simulation results mostly confirm the ARE rankings. Under the distributions considered, the optimality of the *R*-estimators based on correctly specified densities is verified. In the FVML case, the spherical mean is efficient as expected, but the *R*-estimators based on FVML scores are reasonable competitors. In general, the efficiency of our *R*-estimators is better with respect to the spherical mean under departures from the FVML case, especially under the Sq(1.1) density. The spherical median is clearly dominated by the other estimators. The results also show that, as n increases, the influence of the choice of the preliminary estimator wanes; this confirms the fact that the asymptotic behavior of the *R*-estimators does not depend on this choice (see Assumption C).

5. A Data Application

We have applied our *R*-estimators to a data example. The data consists of 26 measurements of magnetic remanence made on samples collected from Palaeozoic red-beds in Argentina, and has already been used in Embleton (1970) and Fisher, Lewis, and Embleton (1987). The purpose of the study was to determine the origin of natural remanent magnetization in red-beds. While it is reasonable to assume that the underlying distribution associated with this data is unimodal (because a single component of magnetization is present, see Fisher, Lewis, and Embleton (1987)), a complete specification of the underlying distribution can't be justified, and hence semi-parametric methods are required.

First, we computed the spherical mean $\hat{\boldsymbol{\theta}}_{\mathrm{Mean}} = (0.3187373, 0.4924234, -0.8098924)'$ and the spherical median $\hat{\boldsymbol{\theta}}_{\mathrm{Median}} = (0.3129147, 0.486036,$

Table 3. MSE of various R-estimators and of the spherical mean and spherical median. For the sake of clarity, the results have been multiplied by 10^3. R-estimators are computed by taking both the spherical median and the spherical mean as preliminary estimators.

Sample size		$n = 100$		$n = 500$		$n = 1,000$	
Preliminary estimator		$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	$\hat{\boldsymbol{\theta}}_{\text{Median}}$
Actual density	Estimators						
	$\hat{\boldsymbol{\theta}}_{\text{FVML(2)}}$	0.12846	0.12813	0.00528	0.00528	0.00135	0.00134
	$\hat{\boldsymbol{\theta}}_{\text{FVML(4)}}$	0.13724	0.13779	0.00550	0.00551	0.00143	0.00142
	$\hat{\boldsymbol{\theta}}_{\text{Lin(2)}}$	0.12818	0.12836	0.00531	0.00532	0.00135	0.00134
FVML (2)	$\hat{\boldsymbol{\theta}}_{\text{Lin(4)}}$	0.13540	0.13568	0.00576	0.00577	0.00147	0.00146
	$\hat{\boldsymbol{\theta}}_{\text{Sq(1.1)}}$	0.16916	0.17003	0.00647	0.00648	0.00170	0.00170
	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	0.12686	——	0.00528	——	0.00135	——
	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	——	0.16649	——	0.00687	——	0.00172
	$\hat{\boldsymbol{\theta}}_{\text{FVML(2)}}$	0.01815	0.01810	0.00072	0.00072	0.00019	0.00018
	$\hat{\boldsymbol{\theta}}_{\text{FVML(4)}}$	0.01756	0.01757	0.00069	0.00069	0.00018	0.00017
	$\hat{\boldsymbol{\theta}}_{\text{Lin(2)}}$	0.01852	0.01849	0.00075	0.00074	0.00019	0.00019
FVML (4)	$\hat{\boldsymbol{\theta}}_{\text{Lin(4)}}$	0.02158	0.02160	0.00087	0.00087	0.00022	0.00022
	$\hat{\boldsymbol{\theta}}_{\text{Sq(1.1)}}$	0.01898	0.01909	0.00074	0.00074	0.00019	0.00018
	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	0.01742	——	0.00069	——	0.00018	——
	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	——	0.02481	——	0.00098	——	0.00023
	$\hat{\boldsymbol{\theta}}_{\text{FVML(2)}}$	16.40147	15.88246	0.80516	0.80281	0.20099	0.20068
	$\hat{\boldsymbol{\theta}}_{\text{FVML(4)}}$	17.61341	17.41290	0.85555	0.85874	0.21485	0.21566
	$\hat{\boldsymbol{\theta}}_{\text{Lin(2)}}$	16.27424	16.73361	0.80201	0.80703	0.20081	0.20137
Lin (2)	$\hat{\boldsymbol{\theta}}_{\text{Lin(4)}}$	16.13449	16.64150	0.85463	0.85917	0.21771	0.21827
	$\hat{\boldsymbol{\theta}}_{\text{Sq(1.1)}}$	20.41046	20.14262	1.05501	1.06998	0.25629	0.25712
	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	16.62219	——	0.89528	——	0.22810	——
	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	——	19.16783	——	1.06145	——	0.26050
	$\hat{\boldsymbol{\theta}}_{\text{FVML(2)}}$	201.7671	186.5235	16.30444	15.85266	3.75891	3.73754
	$\hat{\boldsymbol{\theta}}_{\text{FVML(4)}}$	212.7127	212.5497	18.71233	18.35010	4.35718	4.46196
	$\hat{\boldsymbol{\theta}}_{\text{Lin(2)}}$	199.7272	203.3786	15.94977	16.21466	3.68828	3.71774
Lin (4)	$\hat{\boldsymbol{\theta}}_{\text{Lin(4)}}$	187.5686	195.9556	14.23143	14.57779	3.38679	3.42702
	$\hat{\boldsymbol{\theta}}_{\text{Sq(1.1)}}$	231.3489	229.2380	22.61492	23.40273	5.64923	5.92582
	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	186.9112	——	14.26441	——	3.43922	——
	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	——	195.7693	——	17.05978	——	3.94507
	$\hat{\boldsymbol{\theta}}_{\text{FVML(2)}}$	13.28350	13.13711	0.50958	0.50620	0.12659	0.12613
	$\hat{\boldsymbol{\theta}}_{\text{FVML(4)}}$	13.05349	13.54705	0.44770	0.44563	0.10990	0.11025
	$\hat{\boldsymbol{\theta}}_{\text{Lin(2)}}$	13.51480	13.93615	0.53481	0.53583	0.13304	0.13324
Sq (1.1)	$\hat{\boldsymbol{\theta}}_{\text{Lin(4)}}$	15.82969	16.31837	0.71788	0.72184	0.18107	0.18156
	$\hat{\boldsymbol{\theta}}_{\text{Sq(1.1)}}$	13.57166	14.35157	0.42719	0.42348	0.10430	0.10487
	$\hat{\boldsymbol{\theta}}_{\text{Mean}}$	16.92412	——	0.77354	——	0.19337	——
	$\hat{\boldsymbol{\theta}}_{\text{Median}}$	——	18.55458	——	0.78308	——	0.19223

Cosine of the angles with respect to the spherical mean
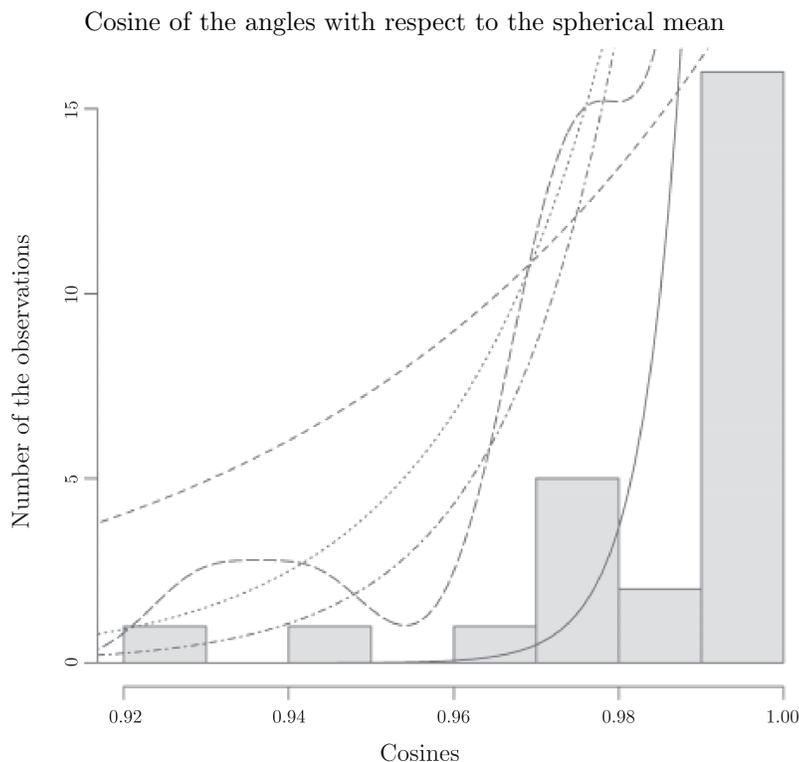
Figure 1. Histogram of the cosines $\boldsymbol{X}_1'\hat{\boldsymbol{\theta}}_{\text{Mean}}, \ldots, \boldsymbol{X}_{26}'\hat{\boldsymbol{\theta}}_{\text{Mean}}$ of the Palaeozoic data. The dashed line, the dotted line, the dotdashed line, and the solid line are plots of the FVML densities with $\kappa = 20$, $\kappa = 50$, $\kappa = 69.544$, and $\kappa = 200$, respectively. The longdashed line represents a Gaussian kernel density estimator.

$-0.8159984)'$. On the basis of this, we provide a histogram of the cosines $\boldsymbol{X}_1'\hat{\boldsymbol{\theta}}_{\text{mean}}, \ldots, \boldsymbol{X}_{26}'\hat{\boldsymbol{\theta}}_{\text{mean}}$; see Figure 1 (the histogram obtained from the cosines $\boldsymbol{X}_1'\hat{\boldsymbol{\theta}}_{\text{Median}}, \ldots, \boldsymbol{X}_{26}'\hat{\boldsymbol{\theta}}_{\text{Median}}$ has a very similar shape). A visual inspection of the histogram pleads in favor of a FVML score-based *R*-estimation (the black line represents a Gaussian kernel density estimator). Now, even if the FVML family is the target family of densities, the concentration parameter has to be chosen to perform our one-step *R*-estimation. The FVML maximum likelihood estimator of κ is given by $\hat{\kappa}_{\text{MLE}} = 69.544$ (see, e.g., Ko (1992)).

We computed FVML-score based *R*-estimators with $\kappa = 20$, $\kappa = 50$, $\kappa = 200$, and $\kappa = 69.544$ (we took $\hat{\boldsymbol{\theta}}_{\text{Median}}$ as the preliminary estimator). We obtained $\hat{\boldsymbol{\theta}}_{\kappa=20} = (0.3152545, 0.4947636, -0.8099027)'$, $\hat{\boldsymbol{\theta}}_{\kappa=50} = (0.3156331, 0.4949186, -0.8096652)'$, $\hat{\boldsymbol{\theta}}_{\kappa=200} = (0.3156907, 0.4948861, -0.8096624)'$ and $\hat{\boldsymbol{\theta}}_{\kappa=69.544} = (0.3156517, 0.4949093, -0.8096636)'$. The data are illustrated in Figure 2. Data
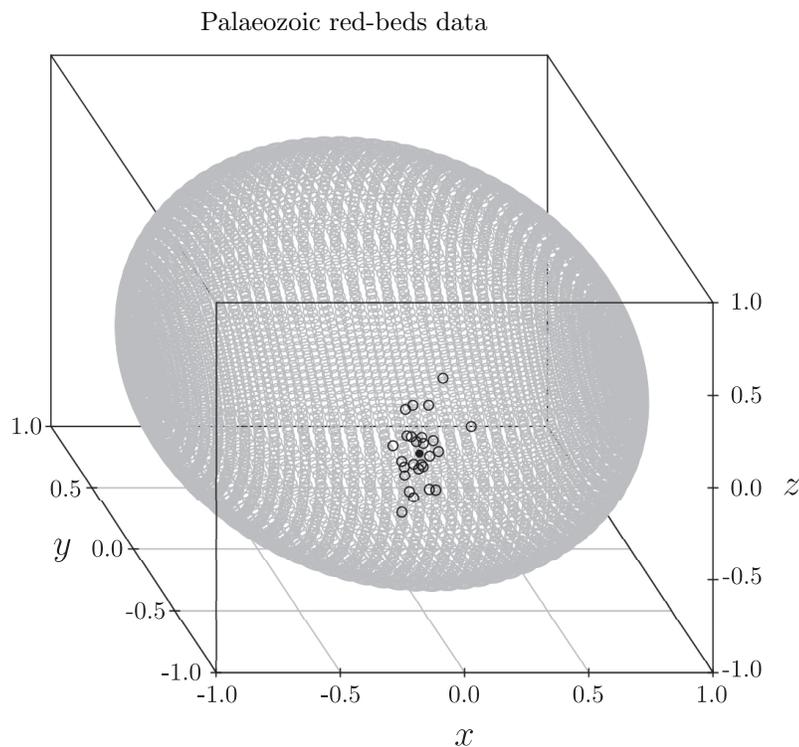
Figure 2. The Palaeozoic data (empty points) and $\hat{\boldsymbol{\theta}}_{\kappa=50}$ (full point). The data consists of 26 measurements of magnetic remanence.

points are represented by empty points, the full point is $\hat{\boldsymbol{\theta}}_{\kappa=50}$.

Of course, one could argue that if the FVML specification is chosen, the spherical mean $\hat{\boldsymbol{\theta}}_{\text{Mean}}$ is the efficient estimator and, as a consequence, has to be used for the estimation of $\boldsymbol{\theta}$. We insist here on the fact that we do not specify a FMVL distribution but rather choose the FVML family as a target. As shown in the ARE results (see Table 1), FVML-score based estimators are *robust-efficient* in the sense that if the underlying distribution is not a FVML one, they can be more efficient than the spherical mean.

The example provided in this section raises two questions that are beyond the scope of the present paper. First, is $\hat{\boldsymbol{\theta}}_{\hat{\kappa}}$ (the R-estimator based on a FVML score with an estimated concentration parameter) as efficient as $\hat{\boldsymbol{\theta}}_{\text{Mean}}$ under any FVML distribution, irrespective of the underlying concentration? Secondly, is it more efficient than the same spherical mean outside of the FVML case? The answer to these questions could be obtained by considering a location/scale model in which we quantify the substitution of the scale parameter by a root-n consistent estimate. This problem is currently under investigation.

Acknowledgements

Appendix: Proofs

Proof of Proposition 1. Since the LAN property is stated with respect to $\boldsymbol{\eta}$, the density function (2.1) is denoted by $f_{\boldsymbol{\eta}}$ in this proof but remains the same function. Our proof relies on Lemma 1 of Swensen (1985)–more precisely, on its extension in Garel and Hallin (1995). The sufficient conditions for LAN in those results readily follow from standard arguments, once it is shown that $\boldsymbol{\eta} \mapsto f_{\boldsymbol{\eta}}^{1/2}(\mathbf{x})$ is differentiable in quadratic mean.

In what follows, all $o(\|\cdot\|)$ or $O(\|\cdot\|)$ quantities are taken as $\|\cdot\| \to 0$. Denoting by

$$\operatorname{grad}_{\boldsymbol{\eta}} f_{\boldsymbol{\eta}}^{1/2}(\mathbf{x}) := \frac{1}{2} f_{\boldsymbol{\eta}}^{1/2}(\mathbf{x}) \varphi_{f_1}(\mathbf{x}'\hbar(\boldsymbol{\eta})) D\hbar(\boldsymbol{\eta})'\mathbf{x}$$

the gradient of the square root of the density $f_{\boldsymbol{\eta}}(\mathbf{x})$, quadratic mean differentiability holds if

$$\int_{\mathcal{S}^{k-1}} \left\{ f_{\boldsymbol{\eta}+\mathbf{e}}^{1/2}(\mathbf{x}) - f_{\boldsymbol{\eta}}^{1/2}(\mathbf{x}) - \mathbf{e}'\operatorname{grad}_{\boldsymbol{\eta}} f_{\boldsymbol{\eta}}^{1/2}(\mathbf{x}) \right\}^2 d\sigma(\mathbf{x}) \tag{A.1}$$

is $o(\|\mathbf{e}\|^2)$ for $\mathbf{e} \in \mathbb{R}^{k-1}$. Obviously, since $\boldsymbol{\eta} \mapsto \hbar(\boldsymbol{\eta})$ is differentiable, we have that $\mathbf{x}'\hbar(\boldsymbol{\eta} + \mathbf{e}) - \mathbf{x}'\hbar(\boldsymbol{\eta}) = \mathbf{x}'D\hbar(\boldsymbol{\eta})\mathbf{e} + o(\|\mathbf{e}\|)$ for all $\mathbf{x} \in \mathcal{S}^{k-1}$. This implies that the integral (A.1) takes the form

$$c_{k;f_1} \int_{\mathcal{S}^{k-1}} \left\{ f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta}) + \mathbf{x}'D\hbar(\boldsymbol{\eta})\mathbf{e} + o(\|\mathbf{e}\|)) - f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta})) \right.$$
$$\left. - \frac{1}{2} f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta})) \varphi_{f_1}(\mathbf{x}'\hbar(\boldsymbol{\eta}))\mathbf{x}'D\hbar(\boldsymbol{\eta})\mathbf{e} \right\}^2 d\sigma(\mathbf{x}). \tag{A.2}$$

Now, since $f_1^{1/2}$ inherits absolute continuity from f_1, $f_1^{1/2}$ is differentiable almost everywhere on $[-1, 1]$. Consequently, for almost all $\mathbf{x}$ and any perturbation $s \in \mathbb{R}$,

$$f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta}) + s) - f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta})) = \frac{1}{2} f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta})) \varphi_{f_1}(\mathbf{x}'\hbar(\boldsymbol{\eta}))s + o(s),$$

so that, using the fact that $\sup_{\mathbf{x}\in\mathcal{S}^{k-1}}|\mathbf{x}'D\hbar(\boldsymbol{\eta})\mathbf{e}|\leq C\|\mathbf{e}\|$ for some positive constant C, we have that

$$\sup_{\mathbf{x}\in\mathcal{S}^{k-1}}\Big|f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta})+\mathbf{x}'D\hbar(\boldsymbol{\eta})\mathbf{e}+o(\|\mathbf{e}\|))-f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta}))$$
$$-\frac{1}{2}f_1^{1/2}(\mathbf{x}'\hbar(\boldsymbol{\eta}))\varphi_{f_1}(\mathbf{x}'\hbar(\boldsymbol{\eta}))\mathbf{x}'D\hbar(\boldsymbol{\eta})\mathbf{e}\Big|$$

is $o(\|\mathbf{e}\|)$ uniformly in $\mathbf{x}$. Consequently, the integrand in (A.2) is $o(\|\mathbf{e}\|^2)$ uniformly in $\mathbf{x}$. The result follows since $\int_{\mathcal{S}^{k-1}}d\sigma(\mathbf{x})=2\pi^{k/2}/\Gamma(k/2)<\infty$.

Proof of Proposition 2. In this proof, we only provide the expressions of $\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)}$ and $\boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1}$ since ULAN directly follows from the combination of Lemma 2 and Proposition 1.

From Lemma 2, we obtain that the family $\left\{\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}\mid\boldsymbol{\theta}\in\mathcal{S}^{k-1}\right\}$ is also ULAN with central sequence

$$\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)}:=n^{-1/2}D\hbar(\boldsymbol{\eta})((D\hbar(\boldsymbol{\eta}))'D\hbar(\boldsymbol{\eta}))^{-1}(D\hbar(\boldsymbol{\eta}))'$$
$$\sum_{i=1}^{n}\varphi_{f_1}(\boldsymbol{X}_i'\boldsymbol{\theta})(1-(\boldsymbol{X}_i'\boldsymbol{\theta})^2)^{1/2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$$

and Fisher information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\theta};f_1}=\frac{\mathcal{J}_k(f_1)}{k-1}D\hbar(\boldsymbol{\eta})((D\hbar(\boldsymbol{\eta}))'D\hbar(\boldsymbol{\eta}))^{-1}(D\hbar(\boldsymbol{\eta}))'.$$

Next, note that if $\boldsymbol{\theta}=\hbar(\boldsymbol{\eta})$ for some $\boldsymbol{\eta}\in\mathbb{R}^{k-1}$ such that $D\hbar$ is full-rank in the vicinity of $\boldsymbol{\eta}$, then, clearly, $\boldsymbol{\theta}=\hbar(\boldsymbol{\eta})=\mathbf{O}\hbar(\boldsymbol{\eta})$, with $\hbar$ defined as in (2.4), for some rotation $\mathbf{O}\in\mathcal{SO}_k$. After some easy computations, we obtain that (putting $\tilde{\boldsymbol{\theta}}=\mathbf{O}'\boldsymbol{\theta}=\hbar(\boldsymbol{\eta})$),

$$D\hbar(\boldsymbol{\eta})((D\hbar(\boldsymbol{\eta}))'D\hbar(\boldsymbol{\eta}))^{-1}(D\hbar(\boldsymbol{\eta}))'=\mathbf{O}D\hbar(\boldsymbol{\eta})((D\hbar(\boldsymbol{\eta}))'D\hbar(\boldsymbol{\eta}))^{-1}(D\hbar(\boldsymbol{\eta}))'\mathbf{O}'$$
$$=\mathbf{O}(\mathbf{I}_k-\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}')\mathbf{O}'$$
$$=\mathbf{I}_k-\boldsymbol{\theta}\boldsymbol{\theta}'.$$

Then, combining Proposition 1 and Lemma 2, the family $\left\{\mathrm{P}_{\boldsymbol{\theta};f_1}^{(n)}\mid\boldsymbol{\theta}\in\mathcal{S}^{k-1}\right\}$ (in the $\boldsymbol{\theta}$-parameterization) is ULAN with central sequence

$$\boldsymbol{\Delta}_{\boldsymbol{\theta};f_1}^{(n)}=n^{-1/2}(\mathbf{I}_k-\boldsymbol{\theta}\boldsymbol{\theta}')\sum_{i=1}^{n}\varphi_{f_1}(\boldsymbol{X}_i'\boldsymbol{\theta})(1-(\boldsymbol{X}_i'\boldsymbol{\theta})^2)^{1/2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$$
$$=n^{-1/2}\sum_{i=1}^{n}\varphi_{f_1}(\boldsymbol{X}_i'\boldsymbol{\theta})(1-(\boldsymbol{X}_i'\boldsymbol{\theta})^2)^{1/2}\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)$$

and Fisher information matrix

$$\mathbf{\Gamma}_{\boldsymbol{\theta};f_1} = \frac{\mathcal{J}_k(f_1)}{k-1}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}').$$

Proof of Proposition 3. Part (i) follows easily from Hájek's classical result for linear signed-rank statistics (see Hájek and Šidák (1967)). Parts (ii) and (iii) are consequences of part (i), the Multivariate Central Limit theorem and Le Cam's third lemma. We therefore only prove in detail part (iv) of the Proposition.

For the sake of simplicity, we let $\boldsymbol{\theta}^{(n)} = \boldsymbol{\theta} + n^{-1/2}\mathbf{t}^{(n)}$ be the perturbed spherical location. In the sequel, we put $\mathbf{U}_i^{(n)} := \boldsymbol{X}_i - (\boldsymbol{X}_i'\boldsymbol{\theta}^{(n)})\boldsymbol{\theta}^{(n)}$, $\mathbf{U}_i^0 := \boldsymbol{X}_i - (\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta}$, $\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_i) := \mathbf{U}_i^{(n)}/\|\mathbf{U}_i^{(n)}\|$, and $\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i) := \mathbf{U}_i^0/\|\mathbf{U}_i^0\|$. We then have the following.

Lemma A.1. *For all $i \in \{1,\ldots,n\}$, we have that $\|\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_i) - \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)\|$ is $o_P(1)$ under $P_{\boldsymbol{\theta};f_1}^{(n)}$ with $f_1 \in \mathcal{F}$ as $n \to \infty$.*

Proof of Lemma A.1. First note that

$$
\begin{aligned}
\|\mathbf{U}_i^{(n)} - \mathbf{U}_i^0\| &= \|(\boldsymbol{X}_i - (\boldsymbol{X}_i'\boldsymbol{\theta}^{(n)})\boldsymbol{\theta}^{(n)}) - (\boldsymbol{X}_i - (\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta})\| \\
&= \|(\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta} - (\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta}^{(n)} + (\boldsymbol{X}_i'\boldsymbol{\theta})\boldsymbol{\theta}^{(n)} - (\boldsymbol{X}_i'\boldsymbol{\theta}^{(n)})\boldsymbol{\theta}^{(n)}\| \\
&\le \|(\boldsymbol{X}_i'\boldsymbol{\theta})(\boldsymbol{\theta} - \boldsymbol{\theta}^{(n)})\| + \|\boldsymbol{X}_i'(\boldsymbol{\theta} - \boldsymbol{\theta}^{(n)})\boldsymbol{\theta}^{(n)}\|.
\end{aligned}
\tag{A.3}
$$

Both terms in (A.3) are clearly $o_P(1)$ as $n \to \infty$. Consequently, we have that $\|\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_i) - \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{X}_i)\| \le 2\,\|\mathbf{U}_i^0\|^{-1}(\|\mathbf{U}_i^{(n)} - \mathbf{U}_i^0\|)$ is $o_P(1)$ under $P_{\boldsymbol{\theta};f_1}^{(n)}$ as $n \to \infty$.

Proof of part (iv). From part (i) we know that $\boldsymbol{\Delta}_{\boldsymbol{\theta};K}^{(n)} - \boldsymbol{\Delta}_{\boldsymbol{\theta};K,g_1}^{(n)} = o_P(1)$ under $P_{\boldsymbol{\theta};g_1}^{(n)}$ as $n \to \infty$. Similarly, $\boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K}^{(n)} - \boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K,g_1}^{(n)} = o_P(1)$ under $P_{\boldsymbol{\theta}^{(n)};g_1}^{(n)}$ as $n \to \infty$. Hence, from contiguity, $\boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K}^{(n)} - \boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K,g_1}^{(n)}$ is also $o_P(1)$ under $P_{\boldsymbol{\theta};g_1}^{(n)}$ as $n \to \infty$. This entails that the claim holds if

$$\boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K,g_1}^{(n)} - \boldsymbol{\Delta}_{\boldsymbol{\theta};K,g_1}^{(n)} + \mathbf{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}^{(n)}$$

is $o_P(1)$ under $P_{\boldsymbol{\theta};g_1}^{(n)}$ as $n \to \infty$. Consequently, the result follows if we can show that

(a) $\boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K,g_1}^{(n)} - \boldsymbol{\Delta}_{\boldsymbol{\theta};K,g_1}^{(n)} - E[\boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K,g_1}^{(n)}]$ is $o_{L^2}(1)$ under $P_{\boldsymbol{\theta};g_1}^{(n)}$ as $n \to \infty$, and that

(b) $E\left[\boldsymbol{\Delta}_{\boldsymbol{\theta}^{(n)};K,g_1}^{(n)}\right] + \mathbf{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}^{(n)}$ is $o(1)$ under $P_{\boldsymbol{\theta};g_1}^{(n)}$ as $n \to \infty$.

We first prove

(a). Using the fact that $\boldsymbol{X}_i'\boldsymbol{\theta}$ and $\boldsymbol{S_\theta}(\boldsymbol{X}_i)$ are independent (as mentioned in the Introduction), we have (the expectation is taken under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$)

$$\mathrm{E}\left[\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K,g_1}\right] = n^{-1/2}\sum_{i=1}^n \mathrm{E}\left[K(\tilde{G}_1(\boldsymbol{X}_i'\boldsymbol{\theta}))\right]\mathrm{E}\left[\boldsymbol{S_\theta}(\boldsymbol{X}_i)\right] = \boldsymbol{0},$$

since, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$, the sign $\boldsymbol{S_\theta}(\boldsymbol{X}_i)$ is uniformly distributed on $\mathcal{S}^{k-1}(\boldsymbol{\theta}^\perp)$. Now, let $\mathbf{D}^{(n)} := n^{-1/2}\sum_{i=1}^n(\mathbf{T}^{(n)}_i - \mathrm{E}[\mathbf{T}^{(n)}_i])$, where $\mathbf{T}^{(n)}_i := K(\tilde{G}_1(\boldsymbol{X}_i'\boldsymbol{\theta}^{(n)}))\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_i)$ $-K(\tilde{G}_1(\boldsymbol{X}_i'\boldsymbol{\theta}))\boldsymbol{S_\theta}(\boldsymbol{X}_i)$. Clearly, $\mathbf{D}^{(n)}$ corresponds exactly to the expression in (a), since $\mathrm{E}[\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K,g_1}] = 0$ as explained above. Using i.i.d.-ness and standard algebra arguments, we have that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$,

$$\mathrm{E}[\|\mathbf{D}^{(n)}\|^2] = n^{-1}\mathrm{E}\left[\left\|\sum_{i=1}^n \mathbf{T}^{(n)}_i - \mathrm{E}[\mathbf{T}^{(n)}_i]\right\|^2\right]$$

$$= n^{-1}\mathrm{tr}\left[\mathrm{Var}\left[\sum_{i=1}^n \mathbf{T}^{(n)}_i - \mathrm{E}[\mathbf{T}^{(n)}_i]\right]\right]$$

$$\leq \mathrm{E}[\|\mathbf{T}^{(n)}_1\|^2].$$

Therefore, it remains to show that $\mathrm{E}[\|\mathbf{T}^{(n)}_1\|^2]$ is $o(1)$ as $n \to \infty$. We have that

$$\|\mathbf{T}^{(n)}_1\|^2 = \left\|\left(K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}^{(n)})) - K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}))\right)\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_1)\right.$$

$$\left.+K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}))\left(\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_1) - \boldsymbol{S_\theta}(\boldsymbol{X}_1)\right)\right\|^2$$

$$\leq 2\left|K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}^{(n)})) - K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}))\right|^2\|\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_1)\|^2$$

$$+2\,K^2(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}))\|\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_1) - \boldsymbol{S_\theta}(\boldsymbol{X}_1)\|^2.$$

The continuity of $K\circ\tilde{G}_1$ together with the fact that $\boldsymbol{X}_1'\boldsymbol{\theta} - \boldsymbol{X}_1'\boldsymbol{\theta}^{(n)}$ is $o_{\mathrm{P}}(1)$ (under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ as $n \to \infty$) imply that $K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}^{(n)})) - K(\tilde{G}_1(\boldsymbol{X}_1'\boldsymbol{\theta}))$ is $o_{\mathrm{P}}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ as $n \to \infty$. Since K is continuous on a compact support (and is therefore bounded, see Assumption D), this convergence also holds in quadratic mean. Similarly, the boundedness of $\boldsymbol{S_\theta}(\boldsymbol{X}_1)$ and $\boldsymbol{S}_{\boldsymbol{\theta}^{(n)}}(\boldsymbol{X}_1)$, together with Lemma A.1, yields the result of (a).

We now turn to the proof of (b). Parts (i), (ii) and (iii) readily state that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ and for $n \to \infty$,

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K,g_1} \overset{\mathcal{L}}{\to} \mathcal{N}_k\left(\boldsymbol{0},\boldsymbol{\Gamma}_{\boldsymbol{\theta};K}\right), \tag{A.4}$$

and that under the sequence of local alternatives $\mathrm{P}^{(n)}_{\boldsymbol{\theta}^{(n)};g_1}$, as $n \to \infty$,

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K,g_1} - \boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}^{(n)} \overset{\mathcal{L}}{\to} \mathcal{N}_k\left(\boldsymbol{0},\boldsymbol{\Gamma}_{\boldsymbol{\theta};K}\right).$$

Now, it follows from ULAN that the model is locally and asymptotically linear (see, e.g., Bickel (1982)). Therefore, Assumption (D), the continuity of $\boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}$ and contiguity entail that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$,

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};K,g_1} + \boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}^{(n)} \overset{\mathcal{L}}{\to} \mathcal{N}_k\left(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\theta};K}\right) \tag{A.5}$$

as $n \to \infty$. We have shown in (a) that $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};K,g_1} - \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta};K,g_1} - \mathrm{E}[\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};K,g_1}]$ is $o_{L^2}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ as $n \to \infty$. Therefore, in view of (A.4), we have that

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};K,g_1} - \mathrm{E}[\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};K,g_1}] \overset{\mathcal{L}}{\to} \mathcal{N}_k\left(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\theta};K}\right). \tag{A.6}$$

Comparing (A.5) and (A.6), it follows that $\mathrm{E}[\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}^{(n)};K,g_1}] + \boldsymbol{\Gamma}_{\boldsymbol{\theta};K,g_1}\mathbf{t}^{(n)}$ is $o(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};g_1}$ which is the desired result.

Since we have proved both (a) and (b), the claim of part (iv) of the proposition holds, which concludes the proof.

References

Bickel, P. J. (1982). On adaptative estimation. *Ann. Statist.* **10**, 647-671.

Chang, T. (2004). Spatial statistics. *Statist. Sci.* **19**, 624-635.

Chang, T. and Rivest, L. P. (2001). *M*-estimation for location and regression parameters in group models: A case study using Stiefel manifolds. *Ann. Statist.* **29**, 784-814.

Chang, T. and Tsai, M. T. (2003). Asymptotic relative Pitman efficiency in group models. *J. Multivariate Anal.* **85**, 395-415.

Embleton, B. J. J. (1970). Palaeomagnetic results for the Premiam of South America and a comparison with the African and Australian data. *Geophys. J. R. Astr. Soc.* **21**, 105-118.

Fisher, R. A. (1953). Dispersion on a sphere. *Proc. R. Soc. Lond. Ser. A* **217**, 295–305.

Fisher, N. I. (1985). Spherical medians. *J. Roy. Statist. Soc. Ser. B* **47**, 342-348.

Fisher, N. I., Lewis, T. and Embleton, B. J. J. (1987). *Statistical Analysis of Spherical Data.* Cambridge University Press, UK.

Garel, B. and Hallin, M. (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend. *Ann. Inst. Statist. Math.* **47**, 551-579.

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

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

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

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

Hallin, M., Swan, Y., Verdebout, T. and Veredas, D. (2011). One-step *R*-estimation in linear models with stable errors. *Ann. J. Econometrics*, to appear.

Ko, D. (1992). Robust estimation of the concentration parameter of the Von Mises-Fisher distribution. *Ann. Statist.* **20**, 917-928.

Ko, D. and Chang, T. (1993). Robust *M*-estimators on spheres. *J. Multivariate Anal.* **45**, 104-136.

Kreiss, J. P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15**, 112-133.

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

Leong, P. and Carlile, S. (1998). Methods for spherical data analysis and visualization. *J. Neurosci. Meth.* **80**, 191-200.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics.* Wiley, New York.

Neeman, T. and Chang, T. (2001). Rank score statistics for spherical data. *Contemp. Math.* **287**, 241-254.

Saw, J. G. (1978). A family of distributions on the m-sphere and some hypothesis tests. *Biometrika* **65**, 69-73.

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

Tsai, M. T. and Sen, P. K. (2007). Locally best rotation-invariant rank tests for modal location. *J. Multivariate Anal.* **98**, 1160-1179.

Tsai, M. T. (2009). Asymptotically efficient two-sample rank tests for modal directions on spheres. *J. Multivariate Anal.* **100**, 445-458.

Watson, G. S. (1983). *Statistics on Spheres.* Wiley, New York.

Département de Mathématique and E.C.A.R.E.S., Université libre de Bruxelles, Boulevard du Triomphe, B-1050 Bruxelles, Belgium.

E-mail: chrisley@ulb.ac.be

Université du Luxembourg, Campus Kirchberg, UR en mathématiques, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Grand-Duché de Luxembourg.

E-mail: yvik.swan@gmail.com

EQUIPPE, Université Lille III, Domaine Universitaire du Pont de Bois, BP 60149, F-59653 Villeneuve d'Ascq Cedex, France.

E-mail: baba.thiam@univ-lille3.fr

EQUIPPE, Université Lille III, Domaine Universitaire du Pont de Bois, BP 60149, F-59653 Villeneuve d'Ascq Cedex, France.

E-mail: thomas.verdebout@univ-lille3.fr