Electronic Journal of Statistics
Vol. 8 (2014) 795–816
ISSN: 1935-7524
DOI: 10.1214/14-EJS904

A new concept of quantiles for directional data and the angular Mahalanobis depth

Christophe Ley

Département de Mathématiques and ECARES, Université Libre de Bruxelles
e-mail: chrisley@ulb.ac.be

Camille Sabbah

Laboratoire EQUIPPE, Université Lille Nord de France
e-mail: camille.sabbah@univ-lille3.fr

and

Thomas Verdebout

INRIA, Laboratoire EQUIPPE, Université Lille Nord de France
e-mail: thomas.verdebout@univ-lille3.fr

Abstract: In this paper, we introduce a new concept of quantiles and depth for directional (circular and spherical) data. In view of the similarities with the classical Mahalanobis depth for multivariate data, we call it the angular Mahalanobis depth. Our unique concept combines the advantages of both the depth and quantile settings: appealing depth-based geometric properties of the contours (convexity, nestedness, rotation-equivariance) and typical quantile-asymptotics, namely we establish a Bahadur-type representation and asymptotic normality (these results are corroborated by a Monte Carlo simulation study). We introduce new user-friendly statistical tools such as directional DD- and QQ-plots and a quantile-based goodness-of-fit test. We illustrate the power of our new procedures by analyzing a cosmic rays data set.

Keywords and phrases: Bahadur representation, directional statistics, DD- and QQ-plot, Mahalanobis depth, rotationally symmetric distributions.

Received September 2013.

Contents

1. Introduction

The notion of *quantile* of a probability distribution is extremely popular in the statistical world, be it for descriptive statistics, exploratory data analysis, inferential procedures or probabilistic aspects. In order to describe the main features of a univariate data set (location, dispersion, skewness, kurtosis), concepts such as the *median, interquartile range, Bowley coefficient of skewness* or *Moor coefficient of kurtosis* (see [4] and [26]), to cite but these, are very useful due to their simplicity and their robustness compared to moment-based measures. The famous *QQ-plot* constitutes a widely used graphical method allowing to assess the accuracy of a theoretical model for a given data set or to determine whether two samples are drawn from a same population. Further, less exploratory, uses of quantiles include the celebrated quantile regression (see [11] and [10]) as well as quantile-based goodness-of-fit tests (see [15] and references therein). Typical quantile-probabilistic results are a Bahadur representation and asymptotic normality.

The success of univariate quantiles has stimulated several researchers to try to extend this fundamental one-dimensional concept to higher dimensions and to circumvent the inherent difficulty of a lack of a natural order in higher dimensions. The complexity of the task can already be perceived through the various definitions of a multivariate/spatial median; see [31] for a survey. Early proposals of multivariate quantiles usually are either descriptive statistics that generalize univariate quantiles or order statistics to higher dimensions, either an extension of a given concept of spatial median, or defined through the coordinate variables. We refer to [5] for a discussion on those attempts (and for a more geometric notion of multivariate quantiles), and to the survey paper of [30] for a review of the distinct existing proposals. In recent years, one proposal has received particular attention, namely the discussion paper [8]. Their concept, based on a directional version of the [11] regression quantiles, enjoys several advantages: easy computation, typical quantile-asymptotics (Bahadur representation and asymptotic normality) and, quite notably, their quantile contours coincide with the classical half space (or Tukey) depth contours.

Depth functions associated with a probability distribution also enjoy a strong popularity among statisticians. They address precisely the above-mentioned lack of natural order in the multivariate setting from a slightly different angle: they provide a center-outward ordering for any multivariate data set by affecting each point $\mathbf{x} \in \mathbb{R}^k$ with a value, the depth of $\mathbf{x}$, determining its centrality

within the data cloud. [30] explains how depth functions provide multivariate notions of order statistics. Numerous distinct notions of data depth have been successfully introduced and studied in the literature, such as the half space depth [33], the simplicial depth [20], the Mahalanobis depth [21] or the zonoid depth [13], to cite but these. A systematic theoretical treatment of statistical depth functions is given in [35], who state four desirable properties that depth functions should satisfy, namely (1) affine-invariance (the depth of any $\mathbf{x}$ should not depend on the underlying coordinate system), (2) maximality at center (for symmetric distributions, the center of symmetry should possess the highest depth), (3) monotonicity relative to the deepest point (the depth of any point $\mathbf{x}$ should be decreasing as $\mathbf{x}$ moves away from the unique deepest point along any ray from that point), and (4) vanishing at infinity (the depth of $\mathbf{x}$ should converge to zero as $||\mathbf{x}||$ tends to infinity). Depth-based procedures are applied to several problems, including location estimation, regression (especially the regression depth of [29]), classification (e.g., via DD-plots, see [19]), trimming, testing procedures (e.g., testing for distinct notions of symmetry) and functional data analysis. For recent reviews on data depth, we refer to [23] and [27].

Deriving depth asymptotics is usually a highly complicated task in view of the difficult structures inherent to the distinct depth functions. This further underlines the merits of the [8] approach, which combines the geometric structure of depth contours with the asymptotic results from multivariate quantiles. This combination of the best properties of both worlds is also our aim in the present paper, where we develop a new concept of quantiles and data depth for directional statistics.

To the best of the authors' knowledge, so far there does not exist a proper concept of quantiles for directional (that is, circular and spherical) data, contrarily to the situation regarding depth. Indeed, three concepts of data depth for directional distributions are nicely presented in [24]: the angular simplicial, angular Tukey and arc distance depths. Those concepts of depth enjoy many attractive properties as it is summarized in [1]. The latter paper illustrates the fact that they are quite appealing for directional data since they provide natural nonparametric orderings. However, these notions of depth suffer from two major drawbacks: (i) they are computationally heavy and (ii) it is extremely difficult to base inference on empirical versions of those depth contours, since results such as asymptotic normality or asymptotic representations do not exist. Thus, the field of directional statistics can so far neither enjoy the above-cited advantages of a good notion of quantile nor does it possess a computationally simple notion of depth such as the Mahalanobis depth in the classical multivariate context. This is why we propose in the present paper a single novel concept serving both as quantile and depth for directional data and which (i) is simple to deal with and computationally light, (ii) provides new descriptive means for directional data, (iii) is canonical in the *rotationally symmetric* case (see the rest of the Introduction for more details on rotationally symmetric distributions), (iv) enjoys the geometric advantages of depth contours (convexity, nestedness, rotation-equivariance), (v) has an empirical version for which

the classical quantile-asymptotics can be proved (Bahadur representation and asymptotic normality), hence (vi) lends itself pretty well for inferential purposes. In view of these properties, we call our new depth *angular Mahalanobis depth*, see Section 2.2.

Directional data naturally arise from multivariate data for which the magnitude of the observed vector is irrelevant. Their domains of application are numerous and diverse: earth sciences, meteorology, neurosciences, astronomy, studies of animal behavior or the protein structure prediction problem (see [25], and [3] for a description of the latter problem). In general, there exist two different types of directional data: (i) the traditional circular/spherical data for which the observations are vectors on the unit sphere $\mathcal{S}^{k-1} := \{\mathbf{v} \in \mathbb{R}^k : \mathbf{v}'\mathbf{v} = 1\}$ and (ii) the axial data which are observed axes, that is, observed unit vectors up to a sign. Obviously, a natural assumption on the distribution of an axial data is the so-called *antipodal symmetry* under which the underlying density f is such that $f(-\mathbf{x}) = f(\mathbf{x})$ for any $\mathbf{x} \in \mathcal{S}^{k-1}$. As shown in [24], antipodally symmetric distributions have constant angular Tukey depth and, consequently, providing any notion of order within this class sounds unrealistic. On the contrary, circular/spherical data lend themselves very well for this purpose, especially under one of the most classical assumptions on the underlying distribution: rotational symmetry (reflective symmetry in the circular case). Rotationally symmetric distributions are characterized by densities of the form

$$\mathbf{x} \mapsto f_{\boldsymbol{\theta}}(\mathbf{x}) = c_{k,f_1}\ f_1(\mathbf{x}'\boldsymbol{\theta}), \quad \mathbf{x} \in \mathcal{S}^{k-1}, \tag{1.1}$$

where $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ is a location parameter (the modal direction), $f_1 : [-1, 1] \to \mathbb{R}_0^+$ an absolutely continuous and (strictly) monotone increasing function and c_{k,f_1} a normalizing constant. This class contains the most popular directional distributions, including the cardioid distribution, the wrapped-normal, the wrapped-Cauchy and, most importantly, the Fisher-von Mises-Langevin (hereafter FvML) distribution obtained by taking $f_1(u) = \exp(\kappa u)$ for some strictly positive concentration parameter κ. The latter, also known as von Mises distribution on the circle $\mathcal{S}^1$ and as Fisher distribution on the sphere $\mathcal{S}^2$, is the most studied and most used directional distribution and is therefore considered as the directional analogue of the classical Gaussian distribution. As announced previously, the concept of quantile/depth we propose can be seen as canonical in the rotationally symmetric case.

The paper is organized as follows. In Section 2, we first introduce both the population and the sample version of our new concept in the quantile setting, and then translate each notion to the depth setting. Then, in Section 3, we establish the typical quantile-asymptotics, namely a Bahadur-type representation and asymptotic normality, with particular emphasis on the case of rotational symmetry. Applications of our quantiles and depth in exploratory data analysis and statistical inference are described in Section 4. The new concept and the related new statistical tools are illustrated in the analysis of cosmic rays data, see Section 5. Monte Carlo simulation studies confirming our asymptotic results are conducted in Section 6. Finally, an appendix collects the technical proofs.

2. The new concept of quantiles and depth for directional data

In this section, we first define our novel concepts (both at population and empirical level) from a quantile point of view (Section 2.1), and then show how all these concepts translate to a depth-based terminology allowing us to define the angular Mahalanobis depth (Section 2.2).

2.1. The new concept of quantiles

We propose in this section our new concept of quantiles for circular and spherical data. The population version of the τ-quantile we provide can be seen as a vector of the form $c_\tau \boldsymbol{\theta}_m$, where c_τ is a real number taking values in $[-1, 1]$ and $\boldsymbol{\theta}_m \in \mathcal{S}^{k-1}$ is the median direction. Our quantiles shall be valid for all random vectors $\mathbf{X} \in \mathcal{S}^{k-1}$ satisfying the following Assumption:

Assumption A. The distribution of $\mathbf{X}$ belongs to the class $\mathcal{F}$ of probability laws on $\mathcal{S}^{k-1}$ with bounded density and which admit a unique median direction $\boldsymbol{\theta}_m$.

There exist distinct concepts of median on the unit sphere based on different concepts of depth (see [24]). We here choose the celebrated [7] spherical median for $\boldsymbol{\theta}_m$, which is related to the arc distance depth as stated in [24], but of course any other choice of directional median is possible. Note that Assumption A rules out antipodally symmetric distributions, for which quantiles, as explained in the Introduction, would anyway have no meaning. Most of the distributions generally used to fit non-axial directional data, and in particular the entire class $\mathcal{R}$ of rotationally symmetric distributions defined in (1.1), satisfy Assumption A, hence fall within the class $\mathcal{F}$. By construction, in the rotationally symmetric case, the Fisher (1985) median, the Tukey median and the angular simplicial median (see [24]) do coincide with the unique mode $\boldsymbol{\theta}$ (see (1.1)). Now, letting $\rho_\tau(u) := u(\tau - \mathbb{I}[u \leq 0]), u \in \mathbb{R}$, stand for the well-known *quantile check function* (see, e.g. [10]), we define

$$c_\tau := \arg \min_{c \in [-1,1]} \mathrm{E}\left[\rho_\tau(\mathbf{X}'\boldsymbol{\theta}_m - c)\right]. \tag{2.2}$$

In other words, the quantity c_τ is the univariate quantile obtained by projecting the vector $\mathbf{X}$ onto the median $\boldsymbol{\theta}_m$; we therefore call c_τ *projection quantile*. Each projection quantile c_τ leads to the subsets

$$\mathcal{C}_\tau^+ := \{\mathbf{x} \in \mathcal{S}^{k-1} \,|\, \mathbf{x}'\boldsymbol{\theta}_m \geq c_\tau\} \quad \text{and} \quad \mathcal{C}_\tau^- := \{\mathbf{x} \in \mathcal{S}^{k-1} \,|\, \mathbf{x}'\boldsymbol{\theta}_m < c_\tau\}$$

defining respectively an *upper quantile cap* and a *lower quantile cap* for $\mathbf{X}$. We denote by $H_{c_\tau \boldsymbol{\theta}_m}$ the hyperplane orthogonal to $\boldsymbol{\theta}_m$ that cuts $\mathcal{S}^{k-1}$ into the regions $\mathcal{C}_\tau^+$ and $\mathcal{C}_\tau^-$ and by $C_{c_\tau \boldsymbol{\theta}_m}$ the corresponding quantile contour, obtained as the intersection between $H_{c_\tau \boldsymbol{\theta}_m}$ and the sphere. This contour thus is formed by a $(k-2)$-dimensional sphere (a circle inside the sphere $\mathcal{S}^2$, two points on the circle $\mathcal{S}^1$) centered at $c_\tau \boldsymbol{\theta}_m$. Summing up, for a given $\tau \in [0, 1]$, the corresponding τ-quantile is thus constructed in two steps: first choose the median $\boldsymbol{\theta}_m$,

then determine the univariate projection quantile c_τ of the projected population $\mathbf{X}'\boldsymbol{\theta}_m$: the result is the τ-quantile $c_\tau\boldsymbol{\theta}_m$.

Our concept of quantiles clearly is of a directional nature, as we fix from the beginning the median direction $\boldsymbol{\theta}_m$ along which we order the data. Thus, the median quantile $c_{1/2}\boldsymbol{\theta}_m$ here is *not* associated with the most central point, but rather with that point on the diameter $c\boldsymbol{\theta}_m$, $c \in [-1,1]$, whose corresponding hyperplane $H_{c_{1/2}\boldsymbol{\theta}_m}$ leaves half of the probability mass above and below it. A value of $\tau = 1$ ($c_1 = 1$) is of course reached by $\boldsymbol{\theta}_m$ whereas $\tau = 0$ ($c_0 = -1$) characterizes its opposite $-\boldsymbol{\theta}_m$, provided that the neighborhood of $-\boldsymbol{\theta}_m$ has not probability mass zero. Would this be the case, then an entire cap centered at $-\boldsymbol{\theta}_m$ would be associated with $\tau = 0$.

While the population version of the projection quantile c_τ coincides with the proposal of [12] for multivariate data in the specific direction $\boldsymbol{\theta}_m$ (contrarily to their set-up, we here of course do not consider other directions), the empirical counterpart of c_τ we propose however strongly differs from theirs because, as explained above, we first have to consistently estimate $\boldsymbol{\theta}_m$ by an estimator $\hat{\boldsymbol{\theta}}_m$. More precisely, the empirical version of our quantiles $c_\tau\boldsymbol{\theta}_m$ and of the upper and lower quantile caps $\mathcal{C}_\tau^+$ and $\mathcal{C}_\tau^-$ can be constructed by following a simple scheme: (1) estimate the median $\boldsymbol{\theta}_m$ by an estimator $\hat{\boldsymbol{\theta}}_m$, (2) project all observations onto $\hat{\boldsymbol{\theta}}_m$ and (3) use a traditional definition of univariate quantiles for determining the $\hat{c}_\tau$. Now, let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ stand for an i.i.d. sequence of random vectors on $\mathcal{S}^{k-1}$ following a common distribution $F \in \mathcal{F}$ (see Assumption A). In view of what precedes, the natural choice of estimator $\hat{\boldsymbol{\theta}}_m$ is obviously to take the empirical spherical median introduced in [7]. Using this $\hat{\boldsymbol{\theta}}_m$, we define

$$\hat{c}_\tau := \arg\min_{c\in[-1,1]} n^{-1}\sum_{i=1}^{n} \rho_\tau(\mathbf{X}_i'\hat{\boldsymbol{\theta}}_m - c)$$

as the empirical version of (2.2). The resulting τ-quantile is then $\hat{c}_\tau\hat{\theta}_m$ and the empirical upper and lower quantile caps are respectively given by the subsets

$$\hat{\mathcal{C}}_\tau^+ := \{\mathbf{x} \in \mathcal{S}^{k-1} \,|\, \mathbf{x}'\hat{\boldsymbol{\theta}}_m \geq \hat{c}_\tau\} \quad \text{and} \quad \hat{\mathcal{C}}_\tau^- := \{\mathbf{x} \in \mathcal{S}^{k-1} \,|\, \mathbf{x}'\hat{\boldsymbol{\theta}}_m < \hat{c}_\tau\}.$$

In Section 3, we study the asymptotic properties of our quantiles, more precisely, of $\hat{c}_\tau$. The latter issue is not trivial since $\hat{c}_\tau$ is the quantile of the sequence $\mathbf{X}_1'\hat{\boldsymbol{\theta}}_m, \ldots, \mathbf{X}_n'\hat{\boldsymbol{\theta}}_m$ which is not an i.i.d. sequence.

We conclude this section by noting that the quantities c_τ and $\hat{c}_\tau$ are rotation-invariant which, combined to the rotation-equivariance of the theoretical and empirical median, entails that the quantiles as well as the quantile caps (but not their size which remains invariant!) are rotation-equivariant.

2.2. *The (related) new concept of depth: The angular Mahalanobis depth*

We now show how the quantities defined in the previous section translate into a depth setting and that our concept provides us with a new depth for directional data, namely the angular Mahalanobis depth.

Let $F \in \mathcal{F}$ with unique mode $\boldsymbol{\theta}_m$ denote the distribution associated with $\mathbf{X}$. Then, for every $\mathbf{x} \in \mathcal{S}^{k-1}$, $D_F(\mathbf{x}) := \arg\min_{\tau\in[0,1]}\{c_\tau \geq \mathbf{x}'\boldsymbol{\theta}_m\}$, the "quantile value" received by $\mathbf{x}$, measures the centrality of $\mathbf{x}$ w.r.t. F. More generally, the function D_F provides a center-outward ordering, with center given by the median direction; hence it actually defines a depth on the sphere. As described in the Introduction, a depth function on $\mathbb{R}^k$ should fulfill four requirements put forward by [35]. We shall now show that D_F does satisfy the $\mathcal{S}^{k-1}$-adapted versions of these conditions.

First, affine-invariance in the spherical case corresponds to rotation-invariance, which trivially follows from the observation at the end of the previous section. Second, D_F attains its maximal value at the center $\boldsymbol{\theta}_m$. Third, D_F decreases along each semi-great-circle from $\boldsymbol{\theta}_m$ to $-\boldsymbol{\theta}_m$, implying monotonicity relative to any deepest point. Four, D_F vanishes at $-\boldsymbol{\theta}_m$, which is the spherical equivalent of infinity in $\mathbb{R}^k$. The four conditions being fulfilled, D_F defines a correct notion of data depth on the sphere $\mathcal{S}^{k-1}$. Now, D_F varies between 0 and 1; the maximum value of the angular Tukey depth is $1/2$ irrespective of the dimension, whereas the angular simplicial depth is bounded by $1/2$ on the unit circle and by $1/4$ on $\mathcal{S}^2$ (see [24]). This naturally brings as to a normalization of our depth D_F into $D_F(\mathbf{x})/(1+D_F(\mathbf{x}))$, whose upper bound equals $1/2$. A simple rewriting of this normalized depth yields

$$\mathbf{x} \mapsto AMHD_F(\mathbf{x}) := \frac{1}{1+\frac{1}{D_F(\mathbf{x})}}, \quad \mathbf{x} \in \mathcal{S}^{k-1}, \tag{2.3}$$

and one readily sees that $AMHD_F$ satisfies the four conditions, too. The notation $AMHD$ stands for angular Mahalanobis depth, because of numerous similarities between the spherical depth defined in (2.3) and the classical Mahalanobis depth on $\mathbb{R}^k$ given by

$$\mathbf{x} \mapsto MHD_F(\mathbf{x}) = \frac{1}{1+(\mathbf{x}-\boldsymbol{\mu}(F))'(\boldsymbol{\Sigma}(F))^{-1}(\mathbf{x}-\boldsymbol{\mu}(F))}, \quad \mathbf{x} \in \mathbb{R}^k,$$

where $\boldsymbol{\mu}(F)$ and $\boldsymbol{\Sigma}(F)$ are location and scatter functionals under F. Clearly, the spherical center $\boldsymbol{\theta}_m$ plays the role of the center $\boldsymbol{\mu}(F)$. Denoting by $O_F(\mathbf{x})$ a measure of outlyingness of $\mathbf{x}$ w.r.t. the deepest point, both $AMHD_F$ and MHD_F are of the form $\frac{1}{1+O_F(\mathbf{x})}$, hence are Type C depth functions (following the terminology of [35]. Moreover, while MHD_F is particularly suited for elliptically symmetric distributions on $\mathbb{R}^k$, $AMHD_F$ is a canonical notion of depth for rotationally symmetric distributions on $\mathcal{S}^{k-1}$ (see Proposition 3.2). These similarities justify, in our opinion, the terminology of angular Mahalanobis depth.

The quantile contours $C_{c_\tau\boldsymbol{\theta}_m}$ of course are at the same time depth contours. They therefore inherit nice geometric properties: they are convex, nested and rotation-equivariant. In case of rotationally symmetric distributions, our depth quantiles coincide with angular Tukey depth contours on the hemisphere centered at $\boldsymbol{\theta}_m$, but our contours clearly improve on their competitor by the fact that they are not constant on the opposite hemisphere (all points lying on the hemisphere centered at $-\boldsymbol{\theta}_m$ possess the same angular Tukey depth).

Finally, letting $\mathbf{X}_1, \ldots, \mathbf{X}_n$ stand for an i.i.d. sequence of random vectors on $\mathcal{S}^{k-1}$ with common distribution $F \in \mathcal{F}$, the corresponding empirical angular Mahalanobis depth is defined as $AMHD(\mathbf{x}) = \frac{1}{1+\frac{1}{\hat{D}(\mathbf{x})}}$ where $\hat{D}(\mathbf{x}) := \arg\min_{\tau\in[0,1]}\{\hat{c}_\tau \geq \mathbf{x}'\hat{\boldsymbol{\theta}}_m\}$ is the empirical version of the depth D_F.

3. Asymptotic properties

In this section, we study the asymptotics of the empirical version of our quantiles and depth. Since these asymptotic results are typical quantile results, we will throughout restrict to the quantile language, whilst keeping in mind the equivalence quantiles-depth shown in the previous section. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ stand for an i.i.d. sequence of random vectors on $\mathcal{S}^{k-1}$ with common distribution $F \in \mathcal{F}$ (see Assumption A). The derivation of the asymptotic properties requires the use of a *locally and asymptotically discrete* version–$\hat{\boldsymbol{\theta}}^*_m$ say–of the empirical Fisher median $\hat{\boldsymbol{\theta}}_m$. This means that $\hat{\boldsymbol{\theta}}^*_m$ is such that, for all $c > 0$, there exists $M = M(c) > 0$ such that the number of possible values of $\hat{\boldsymbol{\theta}}^*_m$ in balls of the form $\{\mathbf{t} \in \mathbb{R}^k : n^{1/2}\|\mathbf{t} - \boldsymbol{\theta}_m\| \leq c\}$ is bounded by M, uniformly in n. Discretization allows to substitute, in $o_{\mathrm{P}}(1)$ sequences, $\hat{\boldsymbol{\theta}}^*_m$ by deterministic expressions of the form $\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}$ (with bounded $\mathbf{t}^{(n)}$ which can be taken such that $\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}$ remains on $\mathcal{S}^{k-1}$, see [18]); we refer to Lemma 4.4 in [14] for details. It is a theoretical assumption that has no impact in fixed-n practice (see pages 125 or 188 of [16] for a discussion). Therefore, for the sake of simplicity, we tacitly assume in what follows that $\hat{\boldsymbol{\theta}}_m$ is locally and asymptotically discrete.

Bearing this in mind and recalling that the empirical spherical median is root-n consistent, we are ready to provide, in the following result, an asymptotic representation of $\hat{c}_\tau$.

Proposition 3.1 (Bahadur-type representation)**.** *Let $F \in \mathcal{F}$ and f_{proj} stand for the common density of the projections $\mathbf{X}_i'\boldsymbol{\theta}_m$, $i = 1, \ldots, n$, and set $\Gamma_{c_\tau} := f_{\mathrm{proj}}(c_\tau)$. Then there exists a k-vector $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau}$ such that*

$$n^{1/2}(\hat{c}_\tau - c_\tau) = \frac{n^{-1/2}}{\Gamma_{c_\tau}} \sum_{i=1}^{n} (\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m \leq c_\tau]) - \frac{\boldsymbol{\Gamma}'_{\boldsymbol{\theta}_m, c_\tau}}{\Gamma_{c_\tau}} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) + o_{\mathrm{P}}(1) \tag{3.4}$$

as $n \to \infty$ under the joint distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_n$.

See the appendix for the proof, whose main difficulty of course lies in the fact that the median $\boldsymbol{\theta}_m$ needs to be estimated. The asymptotic representation in (3.4) directly entails that if the joint normality of

$$\left(n^{-1/2} \sum_{i=1}^{n} (\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \leq 0]), n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m)' \right)'$$

can be established, then the asymptotic normality of $n^{1/2}(\hat{c}_\tau - c_\tau)$ follows. Quite interestingly, the asymptotic Bahadur-type representation of Proposition 3.1 has

a nice form in the rotationally symmetric case, as summarized in the following proposition.

Proposition 3.2. *Let $F \in \mathcal{R}$. Then*

$$n^{1/2}(\hat{c}_\tau - c_\tau) = \frac{n^{-1/2}}{\Gamma_{c_\tau}} \sum_{i=1}^{n} (\tau - \mathbb{I}[\mathbf{X}_i' \boldsymbol{\theta}_m \leq c_\tau]) + o_{\mathrm{P}}(1) \tag{3.5}$$

as $n \to \infty$ under the joint distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_n$. Therefore, letting f_{proj} stand for the density of $\mathbf{X}_i' \boldsymbol{\theta}_m$, we have that $n^{1/2}(\hat{c}_\tau - c_\tau)$ is asymptotically normal with mean zero and variance $(1-\tau)\tau / f^2_{\mathrm{proj}}(c_\tau)$.

See the appendix for the proof. The asymptotic representation (3.5) is exactly equivalent to the Bahadur-type representation for univariate sample quantiles, see e.g. [2]. The absence, in the rotationally symmetric case, of the estimator $\hat{\boldsymbol{\theta}}_m$ in the representation (3.5) is a welcomed feature: it means that the substitution of $\boldsymbol{\theta}_m$ by any root-n consistent estimator (spherical median, *spherical mean* $\sum_{i=1}^n \mathbf{X}_i / \| \sum_{i=1}^n \mathbf{X}_i \|$, etc.) has no asymptotic effect, irrespective of the dimension k. In that sense, our quantiles for directional data are canonical in the rotationally symmetric case.

4. Applications: From exploratory data analysis to statistical inference

Quantiles as robust measures of concentration. Our easy-to-compute quantiles convey some interesting information on how distributions are spread around the median direction $\boldsymbol{\theta}_m$. For the sake of illustration, we provide in Table 1 the projection deciles $c_{i/10}$ for $i = 1, 2, \ldots, 9$ of various circular and spherical distributions: circular cardioid distributions with concentrations $\rho = .1, .2, .5$ ($f_{\mathrm{card}(\rho)}$), circular von Mises distributions with concentrations $\kappa = 1, 5, 10$ ($f_{\mathrm{vm}(\kappa)}$), circular wrapped Cauchy distributions with concentrations $\rho = .1, .5, .9$ ($f_{\mathrm{wrcau}(\rho)}$), circular wrapped normal distributions with concentrations $\rho = .1, .5, .9$ ($f_{\mathrm{wrno}(\rho)}$), spherical FvML distributions with concentrations $\kappa = 1, 2, 5, 10$ ($f_{\mathrm{FvML}(\kappa)}$), spherical linear distributions with parameters $a = 2, 5, 10$ ($f_{\mathrm{lin(a)}}$, see [18]) and spherical Purkayastha distributions (see [28]) with concentrations $\kappa = 1, 5$ ($f_{\mathrm{Pur}(\kappa)}$). It becomes clear from this table that our quantiles not only complement well but provide much more insight into the concentration of the distribution than the single value provided by the *mean resultant length* $R := (\mathrm{E}[\mathbf{X}]' \mathrm{E}[\mathbf{X}])^{1/2}$ or by any single concentration parameter. Moreover, by their nature our quantiles are robust concentration indicators.

QQ-plots. It is clear that our quantiles also allow us to propose QQ-plots for directional distributions. In Figure 1, we provide a few examples of QQ-plots (theoretical FvML(1) against a sample of size 1000 from an FvML(1), a theoretical FvML(1) against a sample of size 1000 from an FvML(3), a theoretical Pur(1) against a sample of size 1000 from a Pur(1) and finally a theoretical

TABLE 1

The projection deciles $c_{i/10}$ for $i = 1, 2, \ldots, 9$ of various circular and spherical distributions; circular cardioid distributions with concentration ρ ($f_{\mathrm{card}(\rho)}$), circular von Mises distributions with concentration κ ($f_{\mathrm{vm}(\kappa)}$), circular wrapped Cauchy distributions with concentration ρ ($f_{\mathrm{wrcau}(\rho)}$), circular wrapped normal distributions with concentration ρ ($f_{\mathrm{wrno}(\rho)}$), spherical FvML distributions with concentration κ ($f_{\mathrm{FvML}(\kappa)}$), spherical linear distributions with parameter a ($f_{\mathrm{lin(a)}}$) and spherical Purkayastha distributions with concentration κ ($f_{\mathrm{Pur}(\kappa)}$). The circular distributions are associated with $\boldsymbol{\theta}_m = (1, 0)'$, the spherical distributions with $\boldsymbol{\theta}_m = (1, 0, 0)'$

Density	$c_{1/10}$	$c_{2/10}$	$c_{3/10}$	$c_{4/10}$	$c_{5/10}$	$c_{6/10}$	$c_{7/10}$	$c_{8/10}$	$c_{9/10}$
$f_{\mathrm{card}(.1)}$	-.9250	-.7199	-.4333	-.1153	.1946	.4710	.6973	.8639	.9656
$f_{\mathrm{card}(.2)}$	-.8732	-.5780	-.2376	.0843	.3640	.5939	.7721	.8990	.9748
$f_{\mathrm{card}(.5)}$	-.2971	.0559	.3152	.5164	.6736	.7955	.8868	.9502	.9876
$f_{\mathrm{vm}(1)}$	-.5952	-.0807	.2744	.5178	.6896	.8123	.8987	.9562	.9892
$f_{\mathrm{vm}(5)}$	.7120	.8256	.8860	.9249	.9518	.9708	.9842	.9932	.9983
$f_{\mathrm{vm}(10)}$	.8609	.9156	.9448	.9636	.9766	.9858	.9923	.9967	.9991
$f_{\mathrm{wrcau}(.1)}$	-.9278	-.7276	-.4411	-.1183	.1979	.4776	.7043	.8679	.9669
$f_{\mathrm{wrcau}(.5)}$	-.6317	-.0256	.4005	.6525	.7999	.8891	.9439	.9768	.9944
$f_{\mathrm{wrcau}(.9)}$	.8010	.9488	.9788	.9895	.9944	.9970	.9985	.9994	.9998
$f_{\mathrm{wrno}(.1)}$	-.9250	-.7200	-.4334	-.1153	.1946	.4711	.6973	.8640	.9656
$f_{\mathrm{wrno}(.5)}$	-.3566	.0619	.3434	.5479	.7009	.8153	.8988	.9558	.9890
$f_{\mathrm{wrno}(.9)}$	.7282	.8318	.8889	.9262	.9524	.9711	.9843	.9932	.9983
$f_{\mathrm{FvML}(1)}$	-.5059	-.1767	.0705	.2686	.4338	.5756	.6998	.8102	.9096
$f_{\mathrm{FvML}(2)}$	-.0750	.2307	.4190	.5555	.6626	.7507	.8256	.8908	.9484
$f_{\mathrm{FvML}(5)}$	.5396	.6782	.7593	.8168	.8614	.8979	.9287	.9554	.9790
$f_{\mathrm{FvML}(10)}$	.7698	.8391	.8797	.9084	.9307	.9490	.9644	.9777	.9895
$f_{\mathrm{lin}(2)}$	-.6583	-.3875	-.1560	.0494	.2361	.4084	.5691	.7203	.8636
$f_{\mathrm{lin}(5)}$	-.7573	-.5278	-.3095	-.1010	.0991	.2916	.4773	.6569	.8310
$f_{\mathrm{lin}(10)}$	-.7804	-.5660	-.3563	-.1511	.0499	.2470	.4404	.6302	.8167
$f_{\mathrm{Pur}(1)}$	-.4373	-.1078	.1386	.3358	.4986	.6356	.7519	.8507	.9337
$f_{\mathrm{Pur}(5)}$	.7359	.8386	.8911	.9243	.9474	.9644	.9772	.9870	.9946

Pur(1) against a sample of size 1000 from a Pur(3)) . We clearly see from these plots that one can easily deduce via this graphical tool whether the theoretical distribution under investigation fits the sample or not. Of course, since our quantiles only take into account the projections onto the median direction $\boldsymbol{\theta}_m$, this QQ-plot does not permit us to distinguish probability laws which only vary on the orthogonal hyperplane to $\boldsymbol{\theta}_m$ but whose projections are similar. But recall that also the original QQ-plot is only a tool in exploratory data analysis to give an approximate idea of the underlying distribution; moreover, it definitely allows us to tell when two distributions are different. Finally note that our QQ-plots can be seen as a generalization of the *colatitude plot* introduced in [17] for FvML distributions. In that paper, the authors show how these plots can serve as graphical goodness-of-fit tests and so provide estimations for the concentration parameter of FvML distributions; the same evidently is also true for our QQ-plots, which are not restricted to the FvML case but can deal with several circular/spherical distributions.

DD-plots. Both the angular Mahalanobis depth $AMHD_F$ and the depth D_F allow us to build DD-plots for directional distributions. The concept of DD-plot

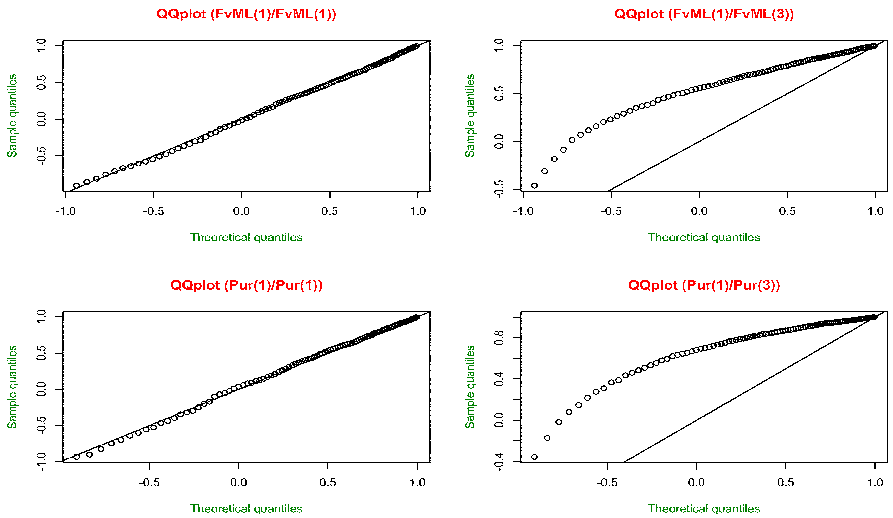

FIG 1. *QQ-plots (theoretical quantiles versus sample quantiles) using theoretical FvML(1) quantiles in the upper plots and theoretical Pur(1) quantiles in the lower plots. In each case, we generated a sample of* 1000 *observations from various distributions: for the upper left QQ-plot from an FvML(1) distribution, for the upper right QQ-plot from an FvML(3) distribution, for the lower left QQ-plot from a Pur(1) distribution, and for the lower right QQ-plot from a Pur(3) distribution.*

was first defined by [22] to compare two multivariate distributions. It has been used in a recent paper by [19] to propose classification methods. Directional DD-plots can be defined as in [22]: letting $\mathbf{X}_{11}, \ldots, \mathbf{X}_{1n_1}$ and $\mathbf{X}_{21}, \ldots, \mathbf{X}_{2n_2}$ be two directional samples with distributions F and G, respectively, the angular Mahalanobis DD-plot is defined as

$$\mathcal{DD}(F,G) := \{(AMHD_F(\mathbf{X}_{ij}), AMHD_G(\mathbf{X}_{ij})), i = 1, 2,\ j = 1, \ldots, n_i\}.$$

A comparison between F and G can then be based on the fact that, if $F = G$, the empirical version of $\mathcal{DD}(F,G)$ should be concentrated around the 45-degree line.

In this paper, we illustrate the usefulness of our depth concept with a "one-sample" version of the DD-plot described above. More precisely, let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be a directional sample with cdf F and define

$$\mathcal{DD}(G) := \{(AMHD(\mathbf{X}_i), AMHD_G(\mathbf{X}_i)), i = 1, \ldots, n\}.$$

The resulting graph $\mathcal{DD}(F)$ should look like a homoskedastic white noise around the same 45-degree line while the graph $\mathcal{DD}(G)$ with $G \neq F$ should show a clear departure from the homoskedastic white noise situation. This is confirmed in Figure 2. Similar DD-plots are obtained by replacing the angular Mahalanobis depth with D_F.

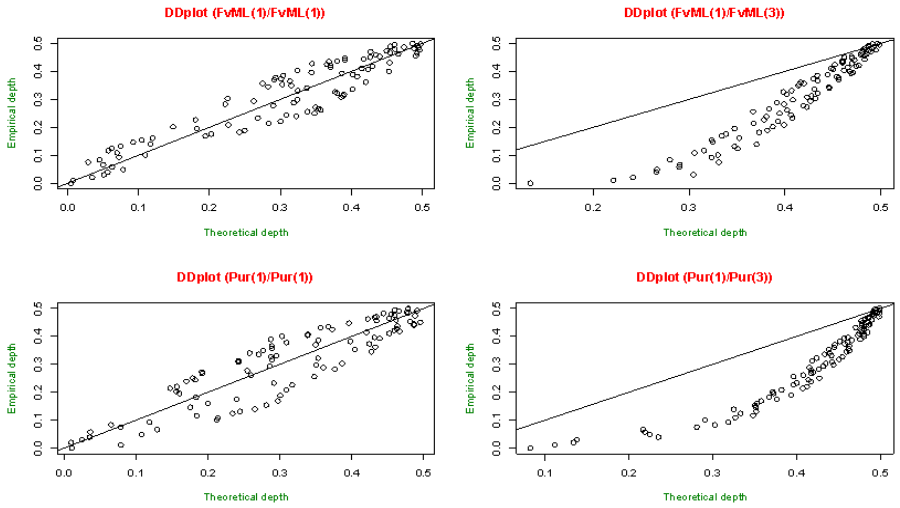

FIG 2. *DD-plots (theoretical depth versus empirical depth) using the theoretical FvML(1) depth in the upper plots and the theoretical Pur(1) depth in the lower plots. In each case, we generated a sample of* 100 *observations from various distributions: for the upper left DD-plot from an FvML(1) distribution, for the upper right DD-plot from an FvML(3) distribution, for the lower left DD-plot from a Pur(1) distribution, and for the lower right DD-plot from a Pur(3) distribution.*

Classification. The spherical classification problem, as considered in [24], consists in determining whether a new data point $\mathbf{Z}$ belongs to either the sample $\mathbf{X}_{11},\ldots,\mathbf{X}_{1n_1}$ or to the sample $\mathbf{X}_{21},\ldots,\mathbf{X}_{2n_2}$, with $n_1,n_2\in\mathbb{N}$. This question can be answered by first computing, within each sample, the depth of $\mathbf{Z}$ and then attributing $\mathbf{Z}$ to the sample where it has highest depth (which coincides with the DD-classification method of Li *et al.* 2012).

Goodness-of-fit tests. Consider the testing problem $\mathcal{H}_0: F=F_0$ for some specific $F_0\in\mathcal{R}$ against $\mathcal{H}_1: F\neq F_0$. Let $\boldsymbol{\tau}:=(\tau_1,\ldots,\tau_m)\in(0,1)^m$ and $\mathbf{T}_{\boldsymbol{\tau}}^{(n)}:=n^{1/2}((\hat{c}_{\tau_1}-c^0_{\tau_1}),\ldots,(\hat{c}_{\tau_m}-c^0_{\tau_m}))'$, where $c^0_{\tau_i}$ is the projection quantile of order τ_i $(i=1,\ldots,m)$ of F_0. It directly follows from Proposition 3.2 that, under $\mathcal{H}_0$, $\mathbf{T}_{\boldsymbol{\tau}}^{(n)}$ is asymptotically normal with mean zero and $(m\times m)$ covariance matrix $\boldsymbol{\Sigma}=(\boldsymbol{\Sigma}_{ij})$, with

$$\boldsymbol{\Sigma}_{ij}=\frac{\min(\tau_i,\tau_j)-\tau_i\tau_j}{f_{0;\mathrm{proj}}(c^0_{\tau_i})f_{0;\mathrm{proj}}(c^0_{\tau_j})},$$

where $f_{0;\mathrm{proj}}(c^0_{\tau_i})$ stands for the density of the projections under F_0 evaluated at $c^0_{\tau_i}$ $(i=1,\ldots,m)$. Note that the covariance matrix $\boldsymbol{\Sigma}$ does not need to be estimated under the null hypothesis. Based on this joint asymptotic normality result which directly follows from the multivariate central limit theorem, a goodness-of-fit test ϕ_{F_0} is obtained by rejecting the null (at the nominal asymptotic level α)

when $Q_{\boldsymbol{\tau}}^{(n)} := (\mathbf{T}_{\boldsymbol{\tau}}^{(n)})'\boldsymbol{\Sigma}^{-1}\mathbf{T}_{\boldsymbol{\tau}}^{(n)}$ exceeds the α-upper quantile of the chi-square distribution with m degrees of freedom.

Compared to other goodness-of-fit tests proposed in the literature on directional statistics, see [25] (Section 12.3), our (extremely simple) quantile-based goodness-of-fit tests ϕ_{F_0} constitute an interesting alternative as they are neither tailored only for a specific null distribution nor only powerful against a specific alternative. A detailed power investigation of these tests is beyond the scope of the present paper and left for future work.

Trimming. It is obvious that our angular Mahalanobis depth lends itselve pretty well for trimming purposes. Indeed, as it provides a center-outward ordering from the median direction, trimming by cutting off the points below the τ-depth contour $C_{c_\tau \boldsymbol{\theta}_m}$ for $\tau \in (0, 1/2)$ will allow to deal only with the $1-\tau$ deepest points. Such a trimming is certainly more adapted than trimming via the angular Tukey depth which is constant on the hemisphere opposite to the deepest point. Note that this simple trimming procedure can as well be useful for constructing a bootstrap confidence region via the percentile method of [6].

5. Empirical illustration: Cosmic rays data

In this section, we make use of our new notions in order to analyze a data set which consists in 148 measurements of arrival directions of cosmic rays. [32] used these observations to study primary cosmic rays in certain energy regions. When starting the analysis of such data, a natural question arises: which model or which distribution should one use to fit the data? The first reaction of a scientist or a practitioner is to make some visual inspection of the data. In Figure 3, we constructed the median upper and lower caps and the third quartile upper and lower caps. Inspection of Figure 3 reveals a relatively low concentration of the data as the median quantile caps almost have the same volume.

Next, we performed goodness-of-fit tests based on the joint asymptotic normality of the projection quartiles $(\hat{c}_{.25}, \hat{c}_{.5}, \hat{c}_{.75})'$ for various FvML and Purkayastha spherical distributions. Of course, our previous visual inspection has led us to consider only small values for the FvML concentration parameter. Table 2 provides the asymptotic p-values associated with the tests, which reveal that, among the distributions considered, the FvML distribution with concentration .7 provides the best fit to the data. Note that we also performed the goodness-of-fit test for various linear distributions; it turns out that the null hypothesis is always rejected at the nominal asymptotic level $\alpha = 5\%$. The results of Table 2 are further corroborated by the QQ-plots in Figure 4 (we omit showing the DD-plots as they convey the same message).

6. Monte Carlo simulation studies

In the present section, our main objective is to confirm the theoretical results obtained in Section 3 and to study the moderate-to-small sample behavior of our quantiles and depth. First, we generated $N = 1,500$ independent replications

TABLE 2
p-values, for the cosmic rays data, of the goodness-of-fit tests based on the quartiles $(\hat{c}_{.25}, \hat{c}_{.5}, \hat{c}_{.75})'$

Density	p-value	Density	p-value
$f_{\mathrm{FvML}(.3)}$	.00135	$f_{\mathrm{Pur}(.3)}$	.00261
$f_{\mathrm{FvML}(.4)}$	.00839	$f_{\mathrm{Pur}(.4)}$	.01316
$f_{\mathrm{FvML}(.5)}$	.02780	$f_{\mathrm{Pur}(.5)}$	.02897
$f_{\mathrm{FvML}(.6)}$	.05274	$f_{\mathrm{Pur}(.6)}$	.02905
$f_{\mathrm{FvML}(.7)}$	.06001	$f_{\mathrm{Pur}(.7)}$	.01295
$f_{\mathrm{FvML}(.8)}$	.04163	$f_{\mathrm{Pur}(.8)}$	.00236
$f_{\mathrm{FvML}(.9)}$	.01737	$f_{\mathrm{Pur}(.9)}$	.00015
$f_{\mathrm{FvML}(1)}$	.00422	$f_{\mathrm{Pur}(1)}$	.00001

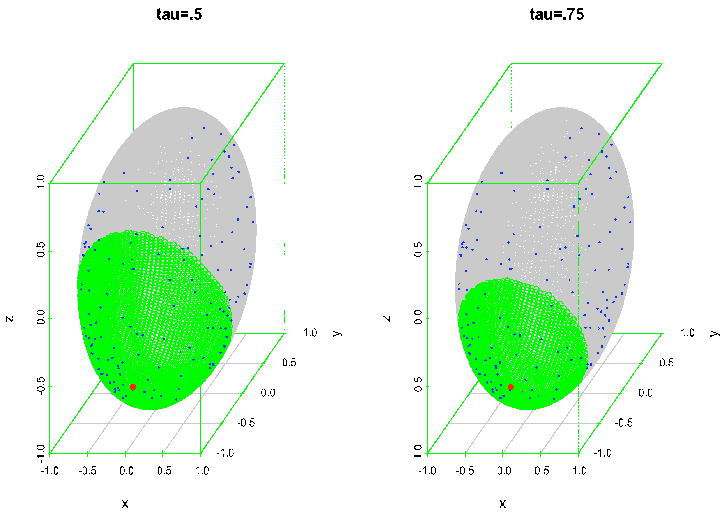

FIG 3. *The zones in grey are the* τ *(equal to* .5 *or* .75*) empirical lower quantile caps while the zones in green are the* τ *(still equal to* .5 *or* .75*) empirical upper quantile caps. The red point is the Fisher (1985) empirical median.*

of four independent samples (with sample size $n = 200$) of $(k =)$3-dimensional (spherical) random vectors

$$\mathbf{X}_{\ell;i}, \quad \ell = 1, 2, 3, 4, \quad i = 1, \ldots, n,$$

with location $\boldsymbol{\theta} = (1, 0, 0)'$ and FvML densities with concentration 1 ($\mathbf{X}_{1;i}$), 2 ($\mathbf{X}_{2;i}$), 5 ($\mathbf{X}_{3;i}$) and 10 ($\mathbf{X}_{4;i}$), respectively. Letting $\hat{c}_{.5}^{(j\ell)}$ stand for the projection median obtained from the jth replication in the design ℓ, we constructed (for all the designs $\ell = 1, 2, 3, 4$) the series

$$\delta_{j\ell} := n^{1/2}(\hat{c}_{.5}^{(j\ell)} - c_{.5}^{(\ell)}), \quad j = 1, \ldots, N, \ell = 1, \ldots, 4,$$

where $c_{.5}^{(\ell)}$ stands for the true underlying median under the ℓ^{th} design. Histograms (with the corresponding theoretical asymptotic distribution) of the se-

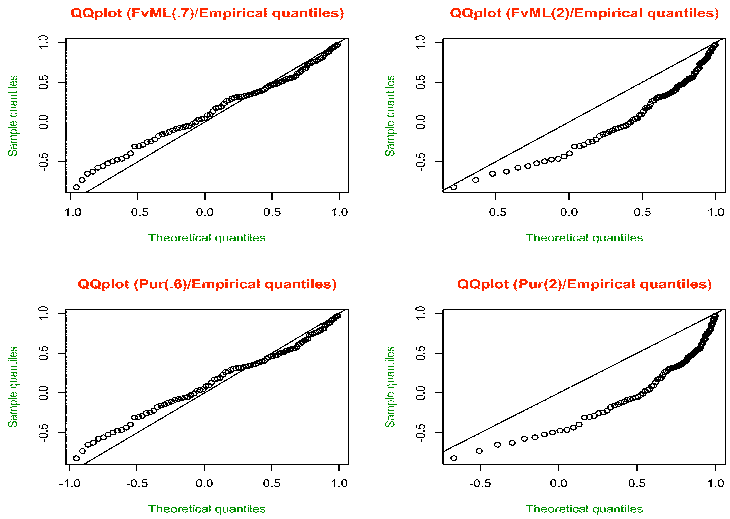

FIG 4. *QQ-plots of the quantiles of various spherical distributions (FvML(.7), FvML(2), Pur(.6) and Pur(2)) with location parameter* $\hat{\boldsymbol{\theta}}_m = (-.4009289, -.0865076, -.9120156)$ *versus the sample quantiles of the cosmic rays data set.*

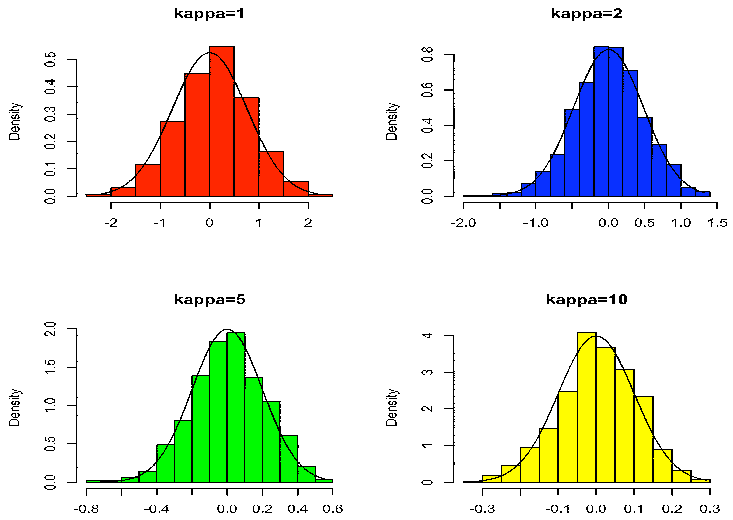

FIG 5. *Histograms of the series* $\delta_{j\ell}$ *under FvML densities with concentrations 1, 2, 5 and 10. The solid line is the (theoretical) asymptotic distribution obtained in Proposition 3.2.*

ries $\delta_{j\ell}$ are provided in Figure 5 and clearly underline the correctness of our theoretical results.

In a second simulation study, we considered circular von Mises, wrapped Cauchy and wrapped normal probability laws. More precisely, we generated

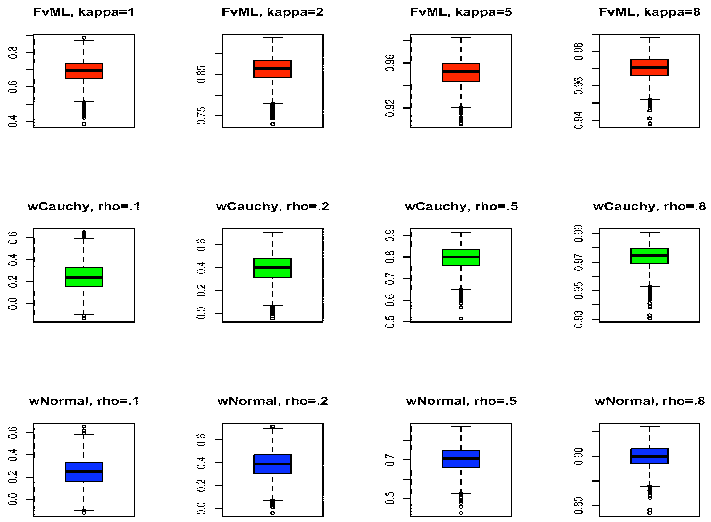

FIG 6. *Boxplots of the* $\hat{c}_{.5}^{(j\ell)}$*'s.*

$N = 2,500$ independent replications of twelve independent samples (with sample size $n = 100$) of circular ($k = 2$) random vectors

$$\mathbf{Y}_{\ell;i}, \quad \ell = 1, \ldots, 12, \quad i = 1, \ldots, n,$$

with location $\boldsymbol{\theta} = (1, 0)'$ and von Mises densities with concentration 1 ($\mathbf{Y}_{1;i}$), 2 ($\mathbf{Y}_{2;i}$), 5 ($\mathbf{Y}_{3;i}$) and 8 ($\mathbf{Y}_{4;i}$), wrapped Cauchy densities with concentration (ρ parameter) .1 ($\mathbf{Y}_{5;i}$), .2 ($\mathbf{Y}_{6;i}$), .5 ($\mathbf{Y}_{7;i}$) and .8 ($\mathbf{Y}_{8;i}$) and finally wrapped normal densities with concentration (ρ parameter) .1 ($\mathbf{Y}_{9;i}$), .2 ($\mathbf{Y}_{10;i}$), .5 ($\mathbf{Y}_{11;i}$) and .8 ($\mathbf{Y}_{12;i}$). As in the previous simulation, writing $\hat{c}_{.5}^{(j\ell)}$ the median obtained from the jth replication in the design ℓ, Figure 6 provides boxplots for the $\hat{c}_{.5}^{(j\ell)}$'s. Inspection of the boxplots reveals that they clearly look like boxplots of Gaussian distributions. As expected, $\hat{c}_{.5}$ increases with the concentration.

7. Final comments

In this paper, we have introduced a new concept of quantiles and depth for directional (circular and spherical) data. In view of the similarities with the classical Mahalanobis depth, we have called our depth function angular Mahalanobis depth. The intimate link between our quantiles and depth entails that our concept enjoys the advantages of both worlds, namely the typical depth-geometric properties (convexity, nestedness) as well as the typical quantile-asymptotics (Bahadur-type representation, asymptotic normality). Our angular Mahalanobis depth constitutes an interesting alternative to existing spherical depth functions since it is easy to compute, we have been able to establish its asymptotic properties thanks to the link with quantiles, and it is not constant on an entire hemisphere (contrarily to the angular Tukey depth). We have also discussed diverse

applications of our quantiles and depth in descriptive statistics, exploratory data analysis and statistical inference. We believe that the concepts defined in this paper shall be useful both for theoreticians and practitioners dealing with directional data, and we are planning to make our procedures available via an `R` code.

Acknowledgments

Christophe Ley thanks the Fonds National de la Recherche Scientifique, Communauté française de Belgique, for support via a Mandat de Chargé de Recherche. All three authors would like to thank the Editor, Associate Editor and an anonymous referee for constructive and helpful comments and suggestions that led to a clear improvement of the present paper.

Appendix: Proofs

Proof of Proposition 3.1. In this proof, all $o_{\mathrm{P}}(\cdot)$ quantities are taken under the joint distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_n$. We shall use the notations $q_\tau(\boldsymbol{\theta}; r) := \mathrm{E}[\rho_\tau(\mathbf{X}_i'\boldsymbol{\theta} - r)]$ and $\psi_\tau(u) := \tau - \mathbb{I}[u \leq 0]$, and consider perturbations of the spherical median of the form $\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}$ for any bounded sequence $\mathbf{t}^{(n)} \in \mathbb{R}^k$. Note that we do not assume that $\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}$ remains on $\mathcal{S}^{k-1}$ so that we can use any directional derivative evaluated at $\boldsymbol{\theta}_m$ in the sequel. In particular, this eases two-times-differentiability (in the sense of distributions) of $q_\tau(\boldsymbol{\theta}; r)$ w.r.t. $\boldsymbol{\theta}$ (two-times-differentiability w.r.t. r following by definition of $q_\tau(\boldsymbol{\theta}; r)$) at $\boldsymbol{\theta} = \boldsymbol{\theta}_m$.

It is easy to check that

$$\begin{aligned}
&\mathcal{L}^{(n)}(\mathbf{t}^{(n)}, s^{(n)}) \qquad\qquad (A.6)\\
&\quad := \sum_{i=1}^n \Big(\rho_\tau(\mathbf{X}_i'(\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}) - (c_\tau + n^{-1/2}s^{(n)})) - \rho_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau)\Big)\\
&\quad = n^{-1/2}\sum_{i=1}^n \begin{pmatrix} \psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau)\mathbf{X}_i \\ -\psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau) \end{pmatrix}' \begin{pmatrix} \mathbf{t}^{(n)} \\ s^{(n)} \end{pmatrix}\\
&\qquad + \sum_{i=1}^n R(\mathbf{X}_i, \boldsymbol{\theta}_m, c_\tau, n^{-1/2}\mathbf{t}^{(n)}, n^{-1/2}s^{(n)}),
\end{aligned}$$

where

$$\begin{aligned}
R(\mathbf{X}_i, \boldsymbol{\theta}, c, \mathbf{t}, s) \quad := \quad & \rho_\tau(\mathbf{X}_i'(\boldsymbol{\theta} + \mathbf{t}) - (c + s)) - \rho_\tau(\mathbf{X}_i'\boldsymbol{\theta} - c)\\
& - \psi_\tau(\mathbf{X}_i'\boldsymbol{\theta} - c)\mathbf{X}_i'\mathbf{t} + \psi_\tau(\mathbf{X}_i'\boldsymbol{\theta} - c)s
\end{aligned}$$

is such that

$$\begin{aligned}
&\mathrm{E}[R(\mathbf{X}_i, \boldsymbol{\theta}_m, c_\tau, n^{-1/2}\mathbf{t}^{(n)}, n^{-1/2}s^{(n)})]\\
&\quad = q_\tau(\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}; c_\tau + n^{-1/2}s^{(n)}) - q_\tau(\boldsymbol{\theta}_m; c_\tau)\\
&\qquad - \mathrm{E}[\psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau)\mathbf{X}_i' n^{-1/2}\mathbf{t}^{(n)} - \psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau)n^{-1/2}s^{(n)}].
\end{aligned}$$

Now, using the fact that $q_\tau(\boldsymbol{\theta}; r)$ is twice differentiable with respect to both $\boldsymbol{\theta}$ (at $\boldsymbol{\theta}_m$) and r, it is easy to verify that, as $n \to \infty$,

$$\sum_{i=1}^{n} \mathrm{E}[R(\mathbf{X}_i, \boldsymbol{\theta}_m, c_\tau, n^{-1/2}\mathbf{t}^{(n)}, n^{-1/2}s^{(n)})] = \frac{1}{2}\begin{pmatrix} \mathbf{t}^{(n)} \\ s^{(n)} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m} & \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau} \\ \boldsymbol{\Gamma}'_{\boldsymbol{\theta}_m, c_\tau} & \Gamma_{c_\tau} \end{pmatrix} \begin{pmatrix} \mathbf{t}^{(n)} \\ s^{(n)} \end{pmatrix} + o(1),$$

where $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_m} = \mathrm{grad}_{\boldsymbol{\theta}} \mathrm{E}[\psi_\tau(\mathbf{X}_i'\boldsymbol{\theta} - c_\tau)\mathbf{X}_i']_{|\boldsymbol{\theta}=\boldsymbol{\theta}_m}$, $\boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau} = \frac{d}{dc}\mathrm{E}[\psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c)\mathbf{X}_i]_{|c=c_\tau}$ and $\Gamma_{c_\tau} = f_{\mathrm{proj}}(c_\tau)$. Now, following equation (4.3) of [10], we also have that

$$\sum_{i=1}^{n} R(\mathbf{X}_i, \boldsymbol{\theta}_m, c_\tau, n^{-1/2}\mathbf{t}^{(n)}, n^{-1/2}s^{(n)}) = \sum_{i=1}^{n} \int_0^{\frac{1}{\sqrt{n}}(-\mathbf{X}_i'\mathbf{t}^{(n)} + s^{(n)})} (\mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \le u] - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \le 0])\, du.$$

Note that, letting $\mathbf{D}^{(n)} := \mathrm{Var}[\sum_{i=1}^{n} R(\mathbf{X}_i, \boldsymbol{\theta}_m, c_\tau, n^{-1/2}\mathbf{t}^{(n)}, n^{-1/2}s^{(n)})]$, we readily obtain that

$$\mathbf{D}^{(n)} \le n\,\mathrm{E}\left[\left(\int_0^{\frac{1}{\sqrt{n}}(-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)})} \mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m \le u + c_\tau] - \mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m \le c_\tau]\, du\right)^2\right].$$

Now a simple change of variables, the Cauchy-Schwarz inequality and the fact that $-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)} \le \|\mathbf{t}^{(n)}\| + s^{(n)}$ (and therefore $|-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)}| \le C^{(n)}$ for some sequence of bounded constants $C^{(n)}$) yield ($f_{\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau}$ stands for the density of $\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau$)

$$\begin{aligned}
\mathbf{D}^{(n)} &\le n\mathrm{E}\left[\left(\int_0^{\frac{1}{\sqrt{n}}(-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)})} \left(\mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m \le u + c_\tau] - \mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m \le c_\tau]\right) du\right)^2\right] \\
&= \mathrm{E}\left[\left(\int_0^{-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)}} \left(\mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau \le u/\sqrt{n}] - \mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau \le 0]\right) du\right)^2\right] \\
&\le \mathrm{E}\left[C^{(n)}\left|\int_0^{-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)}} \left(\mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m \le u/\sqrt{n} + c_\tau] - \mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m \le c_\tau]\right)^2 du\right|\right] \\
&\le \mathrm{E}\left[C^{(n)}\left|\int_0^{-\mathbf{X}_1'\mathbf{t}^{(n)} + s^{(n)}} \left(\mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau \le |u/\sqrt{n}|]\right)^2 du\right|\right] \\
&\le \mathrm{E}\left[C^{(n)}\int_{-C^{(n)}}^{C^{(n)}} \mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau \le |u/\sqrt{n}|] du\right] \\
&= C^{(n)}\int_{-C^{(n)}}^{C^{(n)}} \mathrm{E}\left[\mathbb{I}[\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau \le |u/\sqrt{n}|]\right] du \\
&= C^{(n)}\int_{-C^{(n)}}^{C^{(n)}} \int_{-1-c_\tau}^{1-c_\tau} \mathbb{I}[v < |u|/\sqrt{n}]\, f_{\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau}(v) dv du
\end{aligned}$$

$$\begin{aligned} &\leq C^{(n)}\|f_{\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau}\|_\infty \int_{-C^{(n)}}^{C^{(n)}} \int_{-\frac{|u|}{\sqrt{n}}}^{\frac{|u|}{\sqrt{n}}} dvdu \\ &= \frac{2}{\sqrt{n}} C^{(n)}\|f_{\mathbf{X}_1'\boldsymbol{\theta}_m - c_\tau}\|_\infty \int_{-C^{(n)}}^{C^{(n)}} |u| du \end{aligned}$$

which is clearly $o(1)$ as $n \to \infty$. Wrapping up, we obtain that

$$\begin{aligned} &\mathcal{L}^{(n)}(\mathbf{t}^{(n)}, s^{(n)}) \\ &= \sum_{i=1}^{n} \rho_\tau(\mathbf{X}_i'(\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}) - (c_\tau + n^{-1/2}s^{(n)})) - \rho_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau) \\ &= n^{-1/2}\sum_{i=1}^{n} \begin{pmatrix} \psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau)\mathbf{X}_i \\ -\psi_\tau(\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau) \end{pmatrix}' \begin{pmatrix} \mathbf{t}^{(n)} \\ s^{(n)} \end{pmatrix} \\ &\qquad + \frac{1}{2}\begin{pmatrix} \mathbf{t}^{(n)} \\ s^{(n)} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m} & \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau} \\ \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau}' & \Gamma_{c_\tau} \end{pmatrix} \begin{pmatrix} \mathbf{t}^{(n)} \\ s^{(n)} \end{pmatrix} + o_{\mathrm{P}}(1) \end{aligned}$$

as $n \to \infty$. Now, denote by $\hat{c}_\tau(\mathbf{t}^{(n)})$ the random sequence such that

$$\begin{aligned} n^{1/2}(\hat{c}_\tau(\mathbf{t}^{(n)}) - c_\tau) &= \arg\min_{s^{(n)}} \sum_{i=1}^{n} \rho_\tau[\mathbf{X}_i'(\boldsymbol{\theta}_m + n^{-1/2}\mathbf{t}^{(n)}) - (c_\tau + n^{-1/2}s^{(n)})] \\ &= \arg\min_{s^{(n)}} \mathcal{L}^{(n)}(\mathbf{t}^{(n)}, s^{(n)}). \end{aligned}$$

Evidently

$$\begin{aligned} n^{1/2}(\hat{c}_\tau - c_\tau) &= n^{1/2}(\hat{c}_\tau(n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m)) - c_\tau) \\ &= \arg\min_{s^{(n)}} \mathcal{L}^{(n)}(n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m), s^{(n)}), \end{aligned} \tag{A.7}$$

linking the desired result with our developments in this proof. Now, for any bounded $(\mathbf{t}^{(n)}, s^{(n)})$, we have that the above asymptotic expansion for $\mathcal{L}^{(n)}(\mathbf{t}^{(n)}, s^{(n)})$ holds. Thus, in particular, for $\mathbf{t}^{(n)} = n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m)$ (with $\hat{\boldsymbol{\theta}}_m$ a discretized estimator as discussed at the beginning of Section 3, see also [14], we have that

$$\begin{aligned} &\mathcal{L}^{(n)}(n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m), s^{(n)}) = \\ &\quad n^{-1/2}\sum_{i=1}^{n} \begin{pmatrix} (\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \leq 0])\mathbf{X}_i \\ -(\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \leq 0]) \end{pmatrix}' \begin{pmatrix} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}) \\ s^{(n)} \end{pmatrix} \\ &+ \frac{1}{2}\begin{pmatrix} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) \\ s^{(n)} \end{pmatrix}' \begin{pmatrix} \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m} & \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau} \\ \boldsymbol{\Gamma}_{\boldsymbol{\theta}_m, c_\tau}' & \Gamma_{c_\tau} \end{pmatrix} \begin{pmatrix} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) \\ s^{(n)} \end{pmatrix} + o_{\mathrm{P}}(1) \end{aligned}$$

as $n \to \infty$. Grouping all terms in $s^{(n)}$ together, we can rewrite the latter as

$$\mathcal{L}^{(n)}(n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m), s^{(n)}) = \mathbf{Z}_{\boldsymbol{\theta}_m}^{(n)} + -n^{-1/2}s^{(n)}\sum_{i=1}^{n}(\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \leq 0])$$

$$+ s^{(n)}\boldsymbol{\Gamma}'_{\boldsymbol{\theta}_m,c_\tau} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) + \frac{1}{2}(s^{(n)})^2\Gamma_{c_\tau} + o_{\mathrm{P}}(1),$$

where $\mathbf{Z}^{(n)}_{\boldsymbol{\theta}_m}$ is an $O_{\mathrm{P}}(1)$ sequence which does not depend on $s^{(n)}$. Applying the basic corollary of [9] to the sequence

$$\mathbf{A}_n(s^{(n)}) := s^{(n)}\left[\frac{-1}{n^{1/2}}\sum_{i=1}^{n}(\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \le 0]) + \boldsymbol{\Gamma}'_{\boldsymbol{\theta}_m,c_\tau} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m)\right] + \frac{1}{2}(s^{(n)})^2\Gamma_{c_\tau} + o_{\mathrm{P}}(1),$$

we obtain by (A.7) that

$$n^{1/2}(\hat{c}_\tau - c_\tau) = \frac{n^{-1/2}}{\Gamma_{c_\tau}}\sum_{i=1}^{n}(\tau - \mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c_\tau \le 0]) - \frac{\boldsymbol{\Gamma}'_{\boldsymbol{\theta}_m,c_\tau}}{\Gamma_{c_\tau}} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) + o_{\mathrm{P}}(1),$$

which is the desired result. □

Proof of Proposition 3.2. It directly follows from Proposition 3.1 that it suffices to show that, for $F \in \mathcal{R}$, $\boldsymbol{\Gamma}'_{\boldsymbol{\theta}_m,c_\tau} n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m)$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$. Therefore, let F belong to the class of rotationally symmetric distributions with location parameter $\boldsymbol{\theta} = \boldsymbol{\theta}_m$. For all $i = 1, \ldots, n$, letting $\mathbf{S}_{\boldsymbol{\theta}_m}(\mathbf{X}_i) := (\mathbf{X}_i - (\mathbf{X}_i'\boldsymbol{\theta}_m)\boldsymbol{\theta}_m)/\|\mathbf{X}_i - (\mathbf{X}_i'\boldsymbol{\theta}_m)\boldsymbol{\theta}_m\|$ yields

$$\mathbf{X}_i = (\mathbf{X}_i'\boldsymbol{\theta}_m)\boldsymbol{\theta}_m + \sqrt{1 - (\mathbf{X}_i'\boldsymbol{\theta}_m)^2}\,\mathbf{S}_{\boldsymbol{\theta}_m}(\mathbf{X}_i),$$

which is the so-called *tangent-normal decomposition* of $\mathbf{X}_i$. Since F belongs to $\mathcal{R}$, it follows from [34] that (i) $\mathbf{X}_i'\boldsymbol{\theta}_m$ and $\mathbf{S}_{\boldsymbol{\theta}_m}(\mathbf{X}_i)$ are stochastically independent and (ii) $\mathbf{S}_{\boldsymbol{\theta}_m}(\mathbf{X}_i)$ is uniformly distributed on $\mathcal{S}^{k-2}_{\boldsymbol{\theta}_m^\perp} := \{\mathbf{v} \in \mathcal{S}^{k-1}, \mathbf{v}'\boldsymbol{\theta}_m = 0\}$ (in particular, $\mathrm{E}[\mathbf{S}_{\boldsymbol{\theta}_m}(\mathbf{X}_i)] = \mathbf{0}$). It directly follows from (i) and (ii) that

$$\begin{aligned}\boldsymbol{\Gamma}_{\boldsymbol{\theta}_m,c_\tau} &= -\frac{d}{dc}\mathrm{E}[\mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c \le 0]\mathbf{X}_i]_{|c=c_\tau} \\ &= -\frac{d}{dc}\mathrm{E}[\mathbb{I}[\mathbf{X}_i'\boldsymbol{\theta}_m - c \le 0](\mathbf{X}_i'\boldsymbol{\theta}_m)]_{|_{c=c_\tau}}\boldsymbol{\theta}_m. \qquad \text{(A.8)}\end{aligned}$$

Now, the delta method applied to the mapping $\mathbf{x} \mapsto \mathbf{x}/\|\mathbf{x}\|$ combined with the fact that $\|\hat{\boldsymbol{\theta}}_m\| = \|\boldsymbol{\theta}_m\| = 1$ entails that

$$\begin{aligned}n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) &= n^{1/2}\left(\frac{\hat{\boldsymbol{\theta}}_m}{\|\hat{\boldsymbol{\theta}}_m\|} - \frac{\boldsymbol{\theta}_m}{\|\boldsymbol{\theta}_m\|}\right) \\ &= (\mathbf{I}_k - \boldsymbol{\theta}_m\boldsymbol{\theta}_m')n^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m) + o_{\mathrm{P}}(1) \qquad \text{(A.9)}\end{aligned}$$

as $n \to \infty$ under the joint distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_n$. The result follows by combining (A.8) and (A.9). □

References

[1] Agostinelli, C. and Romanazzi, M. (2013). Nonparametric analysis of directional data based on data depth. *Environ. Ecol. Stat.*, **20**, 253–270. MR3068658

[2] Bahadur, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.*, **37**, 577–580. MR0189095

[3] Boomsma, W., Kent, J. T., Mardia, K. V., Taylor, C. C. and Hamelryck, T. (2006). Graphical models and directional statistics capture protein structure. In S. Barber, P. D. Baxter, K. V. Mardia & R. E. Walls (Eds.), *LASR 2006–Interdisciplinary Statistics and Bioinformatics*, Leeds University Press, UK, 91–94.

[4] Bowley, A. L. (1902). *Elements of Statistics*, 2nd edition. P. S. King, London.

[5] Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.*, **91**, 862–872. MR1395753

[6] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, **7**, 1–26. MR0515681

[7] Fisher, N. I. (1985). Spherical medians. *J. Roy. Stat. Soc. B*, **47**, 342–348. MR0816099

[8] Hallin, M., Paindaveine, D. and Siman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: from L_1 optimization to halfspace depth. *Ann. Statist.*, **38**, 635–669. MR2604670

[9] Hjort, N. and Pollard, D. (1993). Asymptotics for minimisers of convex processes. Unpublished manuscript, `http://www.stat.yale.edu/~pollard/Papers/`.

[10] Koenker, R. (2005). *Quantile Regression*, 1st edition. Cambridge University Press, New York. MR2268657

[11] Koenker, R. and Bassett, G. J. (1978). Regression quantiles. *Econometrica*, **46**, 33–50. MR0474644

[12] Kong, L. and Mizera, I. (2012). Quantile tomography: Using quantiles with multivariate data. *Statist. Sinica*, **22**, 1589–1610. MR3027100

[13] Koshevoy, G. and Mosler, K. (1997). Zonoid trimming for multivariate distributions. *Ann. Statist.*, **25**, 1998–2017. MR1474078

[14] Kreiss, J. P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.*, **15**, 112–133. MR0885727

[15] LaRiccia, V. N. (1991). Smooth goodness-of-fit tests: A quantile function approach. *J. Amer. Statist. Assoc.*, **86**, 427–431. MR1137125

[16] Le Cam, L. and Yang, G. L. (2000). *Asymptotics in Statistics*, 2nd edition. Springer-Verlag, New York. MR1784901

[17] Lewis, T. and Fisher, N. I. (1982). Graphical methods for investigating the fit of a Fisher distribution to spherical data. *Geophys. J. R. astr. Soc.*, **69**, 1–13.

[18] Ley, C., Swan, Y., Thiam, B. and Verdebout, T. (2013). Optimal R-estimation of a spherical location. *Statist. Sinica*, **23**, 305–332. MR3076169

[19] Li, J., Cuesta-Albertos, J. and Liu, R. Y. (2012). Dd-classifier: Nonparametric classification procedures based on dd-plots. *J. Amer. Statist. Assoc.*, **107**, 737–753. MR2980081
[20] Liu, R. Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.*, **18**, 405–414. MR1041400
[21] Liu, R. Y. (1992). Data depth and multivariate rank tests. In Y. Dodge (Ed.), *L-1 Statistics and Related Methods*, North-Holland, Amsterdam, 279–294. MR1214839
[22] Liu, R. Y., Parelius, J. M. and Singh, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion). *Ann. Statist.*, **27**, 783–858. MR1724033
[23] Liu, R. Y., Serfling, R. J. and Souvaine, D. L. (Eds.) (2006). *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*. Amer. Math. Soc. MR2343124
[24] Liu, R. Y. and Singh, K. (1992). Ordering directional data: Concept of data depth on circles and spheres. *Ann. Statist.*, **20**, 1468–1484. MR1186260
[25] Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley, Chichester. MR1828667
[26] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *J. Roy. Stat. Soc. D*, **37**, 25–32.
[27] Mosler, K. (2013). Depth statistics. In C. Becker, R. Fried. S. Kuhnt (Eds.), *Robustness and Complex Data Structures, Festschrift in Honor of Ursula Gather*, Berlin, Springer, 17–34. MR3135871
[28] Purkayastha, S. (1991). A rotationally symmetric directional distribution: Obtained through maximum likelihood characterization. *Sankhya Ser. A*, **53**, 70–83. MR1177768
[29] Rousseeuw, P. J. and Hubert, M. (1999). Regression depth (with discussion). *J. Amer. Statist. Assoc.*, **94**, 388–433. MR1702314
[30] Serfling, R. (2002). Quantile functions for multivariate analysis: Approaches and applications. *Statist. Neerlandica*, **56**, 214–232. Special issue: Frontier research in theoretical statistics, 2000 (Eindhoven). MR1916321
[31] Small, C. G. (1990). A survey of multidimensional medians. *International Statistical Review*, **58**, 263–277.
[32] Toyoda, Y., Suga, K., Murakami, K., Hasegawa, H., Shibata, S., Domingo, V., Escobar, I., Kamata, K., Bradt, H., Clark, G. and La Pointe, M. (1965). Studies of primary cosmic rays in the energy region 10^{14} eV to 10^{17} eV (Bolivian Air Shower Joint Experiment). *Proc. Int. Conf. Cosmic Rays (London)*, **2**, 708–711.
[33] Tukey, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2, Canad. Math. Congress, Montreal, Quebec*, 523–531. MR0426989
[34] Watson, G. S. (1983). *Statistics on Spheres*. Wiley, New York. MR0709262
[35] Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function. *Ann. Statist.*, **28**, 461–482. MR1790005