Journal of Multivariate Analysis 159 (2017) 67–81

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Skew-rotationally-symmetric distributions and related efficient inferential procedures

Christophe Ley [a], Thomas Verdebout [b,*]

[a] Department of Applied Mathematics, Computer Science and Statistics, Ghent University (UGent), Belgium

[b] Département de Mathématique and ECARES, Université libre de Bruxelles (ULB), Belgium

ARTICLE INFO

Article history:
Received 22 September 2016
Available online 4 May 2017

AMS subject classifications:
62H11
62H15
62F05

Keywords:
Directional statistics
Rotationally symmetric distributions
Skew-symmetric distributions
Tests for rotational symmetry

ABSTRACT

Most commonly used distributions on the unit hypersphere $\mathcal{S}^{k-1} = \{\mathbf{v} \in \mathbb{R}^k : \mathbf{v}^\top\mathbf{v} = 1\}$, $k \geq 2$, assume that the data are rotationally symmetric about some direction $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$. However, there is empirical evidence that this assumption often fails to describe reality. We study in this paper a new class of skew-rotationally-symmetric distributions on $\mathcal{S}^{k-1}$ that enjoy numerous good properties. We discuss the Fisher information structure of the model and derive efficient inferential procedures. In particular, we obtain the first semi-parametric test for rotational symmetry about a known direction. We also propose a second test for rotational symmetry, obtained through the definition of a new measure of skewness on the hypersphere. We investigate the finite-sample behavior of the new tests through a Monte Carlo simulation study. We conclude the paper with a discussion about some intriguing open questions related to our new models.

1. Introduction

Directional data $\mathbf{X}_1, \ldots, \mathbf{X}_n$ on unit spheres, or simply *spherical data*, are observations taking values on the non-linear manifold $\mathcal{S}^{k-1} = \{\mathbf{v} \in \mathbb{R}^k : \mathbf{v}^\top\mathbf{v} = 1\}$ for some integer $k \geq 2$. Over the past decade, there has been a strong surge of interest in directional statistics, thanks in part to the publication of cornerstone reference books [8,20] and the emergence of new applications in structural bioinformatics, genetics, cosmology and machine learning. Meanwhile, the use of spherical data continues to spread in more traditional fields such as paleomagnetism, meteorology or studies of animal behavior.

In the literature on spherical data, the distribution of the $\mathbf{X}_i$'s is commonly assumed to be rotationally symmetric about some location parameter $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$. The probability density function (pdf) with respect to the usual surface area measure on spheres is then taken to be of the form

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

where the angular function $f : [-1, 1] \to \mathbb{R}^+$ is absolutely continuous and $c_{f,k}$ is a normalizing constant. The terminology "angular function" is closely related to rotational symmetry as it reflects the fact that the distribution of $\mathbf{X}_i$ only depends on the angle (colatitude angle in case $k = 3$) between $\boldsymbol{\theta}$ and $\mathbf{X}_i$ for each $i \in \{1, \ldots, n\}$. A classical example of such a distribution is the *Fisher–von Mises–Langevin (FvML)* distribution with density

$$\mathbf{x} \mapsto \left(\frac{\kappa}{2}\right)^{k/2-1} \frac{1}{2\pi^{k/2} I_{k/2-1}(\kappa)} \exp(\kappa\mathbf{x}^\top\boldsymbol{\theta}), \quad \mathbf{x} \in \mathcal{S}^{k-1}$$

where $\kappa > 0$ is a concentration parameter and $I_{k/2-1}$ the modified Bessel function of the first kind and of order $k/2 - 1$.

* Corresponding author.
E-mail address: tverdebo@ulb.ac.be (T. Verdebout).

http://dx.doi.org/10.1016/j.jmva.2017.02.010
0047-259X/

In practice, however, not all real-life phenomena can be represented by symmetric models. For instance, Leong and Carlile [13] provide evidence that rotational symmetry is a too strong assumption in neurosciences while Mardia [19] explains that in bioinformatics, especially in protein structure prediction, data can be skewed. Motivated by these examples, we study in the present paper a spherical adaptation of the celebrated skew-symmetric distributions on $\mathbb{R}^k$. The vast research stream related to these distributions was initiated in the seminal paper [3] by Azzalini, who investigated the scalar skew-normal density $2\phi(x-\mu)\Phi\{\delta(x-\mu)\}$, $x, \mu, \delta \in \mathbb{R}$, with ϕ and Φ respectively the standard Gaussian density and distribution function. Here μ is a location parameter and δ is a skewness parameter. As an upshot of several generalization efforts, [4,27] proposed the aforementioned multivariate skew-symmetric distributions with pdf

$$2f_k(\mathbf{x}-\boldsymbol{\mu})\Pi_k(\mathbf{x}-\boldsymbol{\mu}, \boldsymbol{\delta}), \quad \mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\delta} \in \mathbb{R}^k \tag{2}$$

where f_k is a centrally symmetric pdf (i.e., $f_k(-\mathbf{x}) = f_k(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^k$) and $\Pi_k : \mathbb{R}^k \times \mathbb{R}^k \to [0, 1]$ satisfies $\Pi_k(-\mathbf{x}, \boldsymbol{\delta}) + \Pi_k(\mathbf{x}, \boldsymbol{\delta}) = 1$ and $\Pi_k(\mathbf{x}, \mathbf{0}) = 1/2$ for all $\mathbf{x}, \boldsymbol{\delta} \in \mathbb{R}^k$. This multiplicative perturbation of symmetry enjoys numerous attractive features, including elegant random number generation procedures and a very simple normalizing constant. Our spherical adaptation will enjoy the same advantages. The usefulness of the perturbation approach on the sphere was already put forward by Jupp et al. [10], who recently provided a general analysis of this approach.

We also derive in this paper various results that are important when asymptotic inference within the new model is considered. More precisely, we show that (i) the Fisher information for this model is singular if and only if the kernel rotationally symmetric density is FvML and (ii) how to construct a new semi-parametric test for rotational symmetry about a fixed center $\boldsymbol{\theta}$ that will be optimal within the entire class of rotationally symmetric distributions. More precisely we show that the classical Watson score test (see [28]) for spherical location is locally and asymptotically optimal, in the Le Cam sense, for testing the null hypothesis of rotational symmetry against the proposed skew alternatives. To the best of our knowledge this result is the first to consider the problem of rotational symmetry from a semi-parametric angle. Moreover, the test we obtain is uniformly optimal against the new class of skew distributions on the sphere. We accompany this test by a natural competitor, obtained through the definition of a novel measure of skewness on the sphere. The derivation of its asymptotic distribution is of independent interest.

The paper is organized as follows. In Section 2 we formally define the skew distributions and discuss how our construction is linked to skew densities on the circle; we further establish a stochastic representation allowing us to generate data from our model. In Section 3 we provide the score functions for location and skewness and the corresponding Fisher information matrix; we also investigate the underlying Fisher singularity issue. Turning our attention towards inferential issues, we build in Section 4 the announced uniformly optimal semi-parametric test for rotational symmetry about a fixed location $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$. In Section 5 we propose a measure of skewness on the sphere and derive from this measure another test for rotational symmetry. The finite-sample behavior of the new tests is investigated through a Monte Carlo simulation study in Section 6. We discuss interesting open questions related to our skew-rotationally-symmetric distributions in Section 7. Finally, Appendix A contains a crucial theoretical development required for the tests of Section 4, and Appendix B collects the technical proofs.

2. Skew-rotationally-symmetric distributions

As mentioned in the Introduction, we adapt to the spherical setting the skew-symmetric construction from $\mathbb{R}^k$, yielding the *skew-rotationally-symmetric (SRS)* distributions. Starting with a rotationally symmetric density (kernel) $f_{\boldsymbol{\theta};k}$ with central direction $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ as defined in (1), the idea of the construction consists in nesting $f_{\boldsymbol{\theta};k}$ into a larger family of distributions whose only rotationally symmetric member is the kernel $f_{\boldsymbol{\theta};k}$. Let $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}$ stand for a $k \times (k-1)$ semi-orthogonal matrix such that

$$\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top = \mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top \quad \text{and} \quad \boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\boldsymbol{\Upsilon}_{\boldsymbol{\theta}} = \mathbf{I}_{k-1},$$

where $\mathbf{I}_\ell$ is the $\ell \times \ell$ identity matrix. Consider a skewing function $\Pi : \mathbb{R} \to [0, 1]$, i.e., a monotone increasing continuous function satisfying $\Pi(-y) + \Pi(y) = 1$ for all $y \in \mathbb{R}$. Skewness is introduced by multiplication of the rotationally symmetric kernel with such a skewing function, turning $f_{\boldsymbol{\theta};k}$ into

$$\mathbf{x} \mapsto f_{\boldsymbol{\theta},\boldsymbol{\delta};k}(\mathbf{x}) = 2c_{f,k}f(\mathbf{x}^\top\boldsymbol{\theta})\Pi(\boldsymbol{\delta}^\top\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{x}), \quad \mathbf{x} \in \mathcal{S}^{k-1} \tag{3}$$

with $\boldsymbol{\delta} \in \mathbb{R}^{k-1}$. The following lemma shows that (3) is a density on $\mathcal{S}^{k-1}$.

Lemma 1. *Let $d\sigma_{k-1}(\mathbf{x})$ stand for the usual surface measure on $\mathcal{S}^{k-1}$. Then for all $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ and all $\boldsymbol{\delta} \in \mathbb{R}^{k-1}$,*

$$\int_{\mathcal{S}^{k-1}} f_{\boldsymbol{\theta},\boldsymbol{\delta};k}(\mathbf{x})\, d\sigma_{k-1}(\mathbf{x}) = 1.$$

A short proof of this result is given in Appendix B. This now allows us to formally define our family of skew distributions on $\mathcal{S}^{k-1}$.

Definition 1. Let $c_{f,k}f(\mathbf{x}^\top\boldsymbol{\theta})$, $\mathbf{x}\in\mathcal{S}^{k-1}$, be a rotationally symmetric density with location $\boldsymbol{\theta}\in\mathcal{S}^{k-1}$, absolutely continuous angular function $f:[-1,1]\to\mathbb{R}^+$ and normalizing constant $c_{f,k}$. Let $\Pi:\mathbb{R}\to[0,1]$ be a skewing function, i.e., a monotone increasing continuous function satisfying $\Pi(-y)+\Pi(y)=1$ for all $y\in\mathbb{R}$. Then the spherical distribution with pdf $f_{\boldsymbol{\theta},\boldsymbol{\delta};k}(\mathbf{x})$ is skew-rotationally-symmetric (SRS) with skewness parameter $\boldsymbol{\delta}\in\mathbb{R}^{k-1}$.

Typical examples of skewing functions Π are cumulative distribution functions (cdfs) G of univariate symmetric (about 0) random variables, the most classical being $G=\Phi$. See Section 6.2 of [7] for further examples. Obviously, at $\boldsymbol{\delta}=\mathbf{0}$, we retrieve $c_{f,k}f(\mathbf{x}^\top\boldsymbol{\theta})$ since $\Pi(0)=1/2$.

Let us now comment on the construction underpinning our pdf (3). In view of (2) it is tempting to define SRS densities as $2c_{f,k}f(\mathbf{x}^\top\boldsymbol{\theta})\Pi(\mathbf{x}^\top\boldsymbol{\delta})$. While this construction is appropriate for breaking elliptical symmetry or central symmetry about $\boldsymbol{\mu}$ in $\mathbb{R}^k$, it makes no sense here because $\mathbf{x}^\top\boldsymbol{\delta}$ does not take into account rotational symmetry around $\boldsymbol{\theta}$, entailing that this function does not even integrate to 1 over $\mathcal{S}^{k-1}$. This is why we rather use $\Pi(\boldsymbol{\delta}^\top\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{x})$, as $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top=\mathbf{I}_k-\boldsymbol{\theta}\boldsymbol{\theta}^\top$ is a projection matrix onto the tangent space to $\mathcal{S}^{k-1}$ at $\boldsymbol{\theta}$. Now, every rotationally symmetric random vector $\mathbf{X}$ admits the tangent–normal decomposition

$$\begin{aligned}\mathbf{X}&=(\mathbf{X}^\top\boldsymbol{\theta})\boldsymbol{\theta}+(1-(\mathbf{X}^\top\boldsymbol{\theta})^2)^{1/2}\mathbf{S}_{\boldsymbol{\theta}}(\mathbf{X})\\&=(\mathbf{X}^\top\boldsymbol{\theta})\boldsymbol{\theta}+(1-(\mathbf{X}^\top\boldsymbol{\theta})^2)^{1/2}\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{X}),\end{aligned}$$

where the sign vector $\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{X})$ is uniformly distributed on $\mathcal{S}^{k-2}$. Thus our construction breaks rotational symmetry at the level of $\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{X})$, as it should. Therefore $\boldsymbol{\delta}$ is a $(k-1)$-dimensional vector. Moreover, note that $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{x}=(1-(\mathbf{x}^\top\boldsymbol{\theta})^2)^{1/2}\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{x})$; thus, SRS pdfs could as well be spelled out as

$$2c_{f,k}f(\mathbf{x}^\top\boldsymbol{\theta})\Pi[\{1-(\mathbf{x}^\top\boldsymbol{\theta})^2\}^{1/2}\boldsymbol{\delta}^\top\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{x})],\tag{4}$$

which shows perhaps in a clearer way the reasoning behind the skewing process.

Besides their intuitive construction, the densities (3) enjoy further nice properties. Calculating normalizing constants is often a very delicate and tedious task on spheres; see, e.g., [11]. This problematic issue is completely avoided here. Furthermore, the following stochastic representation makes it easy to generate random vectors from (3).

Lemma 2 (*Stochastic Representation*). *Generate* $\mathbf{Y}\sim f_{\boldsymbol{\theta};k}$ *rotationally symmetric. Then the uniformly distributed sign vector* $\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{Y})$ *is transformed into*

$$\mathbf{U}_{\boldsymbol{\theta};\Pi}(\mathbf{Y})=\begin{cases}\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{Y}) & \text{if } U\le\Pi[\{1-(\mathbf{Y}^\top\boldsymbol{\theta})^2\}^{1/2}\boldsymbol{\delta}^\top\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{Y})],\\ -\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{Y}) & \text{if } U>\Pi[\{1-(\mathbf{Y}^\top\boldsymbol{\theta})^2\}^{1/2}\boldsymbol{\delta}^\top\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{Y})],\end{cases}$$

where the random variable U *is uniformly distributed on* $(0,1)$ *and independent of* $\mathbf{Y}$. *The* SRS *vector* $\mathbf{X}$ *with density* $f_{\boldsymbol{\theta},\boldsymbol{\delta};k}$ *is obtained as*

$$\mathbf{X}=(\mathbf{Y}^\top\boldsymbol{\theta})\boldsymbol{\theta}+\{1-(\mathbf{Y}^\top\boldsymbol{\theta})^2\}^{1/2}\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}\mathbf{U}_{\boldsymbol{\theta};\Pi}(\mathbf{Y}).\tag{5}$$

This lemma is proved in Appendix B. Finally, as an illustration of our skew distributions we provide in Fig. 1 contour plots of various skew-FvML densities.

We conclude this section with a comparison to the circular setting. Umbach and Jammalamadaka [25] suggest skew-circular densities of the type $2f_0(x-\theta)G\{\omega(x-\theta)\}$, $x\in[-\pi,\pi)$, with location $\theta\in[-\pi,\pi)$, where $G(x)=\int_{-\pi}^x g(y)dy$ is the cdf of some circular symmetric density g and ω is a weighting function satisfying the following three conditions for all $x\in[-\pi,\pi)$:

(i) $\omega(-x)=-\omega(x)$;
(ii) $\omega(x+2\pi p)=\omega(x)$ for all $p\in\mathbb{Z}$;
(iii) $|\omega(x)|\le\pi$.

Abe and Pewsey [1] have thoroughly studied a particular instance of this construction, namely $G(x)=(\pi+x)/(2\pi)$ and $\omega(x)=\delta\pi\sin(x)$ with skewness parameter $\delta\in(-1,1)$. The resulting densities

$$x\mapsto f_0(x-\theta)\{1+\delta\sin(x-\theta)\},\quad x\in[-\pi,\pi)$$

are said to be sine-skewed. Now, expressing sine-skewed densities in terms of our SRS density (4), we find that $\Pi(x)=(1+x)/2$ for $x\in[-1,1]$, $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{x}=\mathbf{1}(x\ge\theta)-\mathbf{1}(x\le\theta)$ (meaning here that x lies on the right, respectively left, hemisphere when θ is considered as north pole), the scalar product $\mathbf{x}^\top\boldsymbol{\theta}$ is $\cos(x-\theta)$, whence $\sqrt{1-(\mathbf{x}^\top\boldsymbol{\theta})^2}$ here equals $|\sin(x-\theta)|$. Thus, sine-skewed densities are special cases of our SRS distributions on the unit circle $\mathcal{S}^1$, and our general construction sheds new light on these skew-circular densities. Moreover, our Fisher information singularity analysis from the subsequent Section 3 is in agreement with that in [16] on sine-skewed densities.

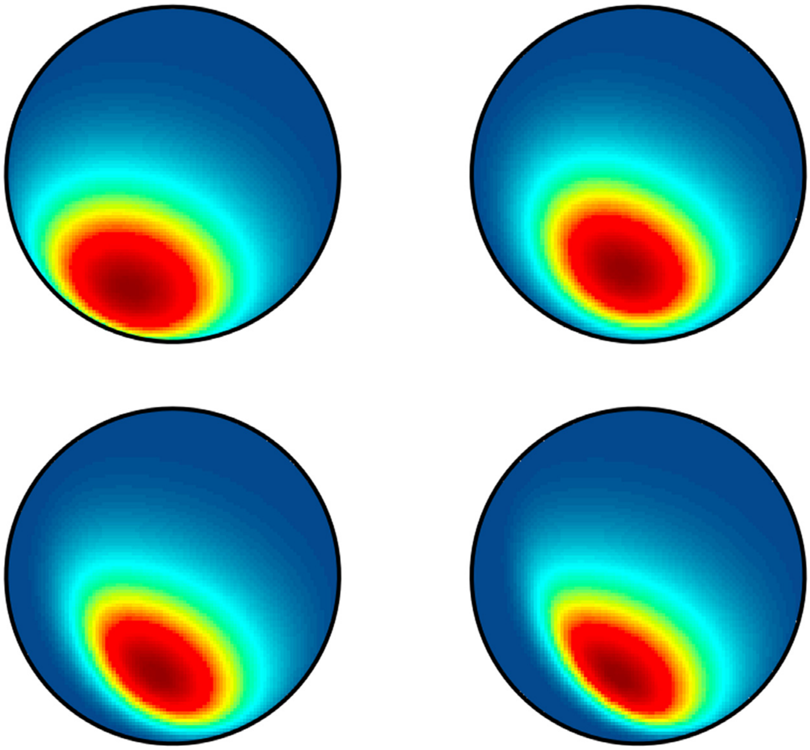

Fig. 1. Contour plots of the skew-FvML density with concentration $\kappa = 5$ and varying skewness parameter $\boldsymbol{\delta} = (2(d-1), 2(d-1))^\top$ for $d \in \{1, 2, 3, 4\}$; top left: $d = 1$, top right: $d = 2$, bottom left: $d = 3$, bottom right: $d = 4$. The skewing function Π is the standard Gaussian cdf.

3. Score functions and Fisher information matrix in the vicinity of symmetry

In this section we provide the expression for the location and skewness scores and the associated Fisher information matrix when $\boldsymbol{\delta} = \mathbf{0}$. By absolute continuity (see Definition 1), we know that the pdf f is almost everywhere (a.e.) continuously differentiable. In order to proceed, we need the following essentially technical assumption.

Assumption A. The skewing function Π is a.e. continuously differentiable. Moreover, if $\dot{f}$ denotes the a.e.-derivative of f and $\varphi_f = \dot{f}/f$, the quantities $\mathcal{J}_k(f) = \int_{-1}^{1} \varphi_f^2(t)(1-t^2)\tilde{f}(t)dt$ and $\mathcal{A}_k(f) = \int_{-1}^{1}(1-t^2)\tilde{f}(t)dt$ are finite, with

$$t \mapsto \tilde{f}(t) = \frac{\omega_k\, c_{f,k}}{\mathcal{B}\left[\frac{1}{2}, \frac{1}{2}(k-1)\right]} f(t)(1-t^2)^{(k-3)/2}, \quad -1 \le t \le 1 \tag{6}$$

where $\omega_k = 2\pi^{k/2}/\Gamma(k/2)$ is the surface area measure of $\mathcal{S}^{k-1}$ and $\mathcal{B}(\cdot,\cdot)$ is the beta function.

Assumption A implies that the score vector exists and that the Fisher information matrix is finite. The function $\tilde{f}$ corresponds to the density of $\mathbf{X}^\top\boldsymbol{\theta}$ if $\mathbf{X} \sim c_{f,k} f(\mathbf{x}^\top\boldsymbol{\theta})$. For the sake of convenience, we write $\boldsymbol{\vartheta} = (\boldsymbol{\theta}^\top, \boldsymbol{\delta}^\top)^\top \in \mathcal{S}^{k-1} \times \mathbb{R}^{k-1}$ and $f_{\boldsymbol{\vartheta};k}$ for $f_{\boldsymbol{\theta},\boldsymbol{\delta};k}$. The score vector $\boldsymbol{\ell}_{f,\Pi;\boldsymbol{\vartheta}}(\mathbf{x})$ at $(\boldsymbol{\theta}^\top, \mathbf{0}^\top)^\top = \boldsymbol{\vartheta}_0$ takes the form

$$\begin{aligned}\boldsymbol{\ell}_{f,\Pi;\boldsymbol{\vartheta}_0}(\mathbf{x}) &= \mathrm{grad}_{\boldsymbol{\vartheta}} \log f_{\boldsymbol{\vartheta};k}(\mathbf{x})|_{\boldsymbol{\vartheta}_0} = \left((\boldsymbol{\ell}^1_{f;\boldsymbol{\vartheta}_0}(\mathbf{x}))^\top,\ (\boldsymbol{\ell}^2_{\Pi;\boldsymbol{\vartheta}_0}(\mathbf{x}))^\top\right)^\top \\ &= \begin{pmatrix} \varphi_f(\mathbf{x}^\top\boldsymbol{\theta})\{1-(\mathbf{x}^\top\boldsymbol{\theta})^2\}^{1/2}\mathbf{S}_{\boldsymbol{\theta}}(\mathbf{x}) \\ 2\Pi'(0)\{1-(\mathbf{x}^\top\boldsymbol{\theta})^2\}^{1/2}\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{x}) \end{pmatrix},\end{aligned}$$

where the factor 2 in $\boldsymbol{\ell}^2_{\Pi;\boldsymbol{\vartheta}_0}(\mathbf{x})$ follows from the fact that $\Pi(0) = 1/2$. The corresponding $(2k-1) \times (2k-1)$ Fisher information matrix $\boldsymbol{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0}$ is then given by

$$\begin{aligned}\boldsymbol{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0} &= \begin{pmatrix} \boldsymbol{\gamma}^{11}_{f;\boldsymbol{\vartheta}_0} & \boldsymbol{\gamma}^{12}_{f,\Pi;\boldsymbol{\vartheta}_0} \\ (\boldsymbol{\gamma}^{12}_{f,\Pi;\boldsymbol{\vartheta}_0})^\top & \boldsymbol{\gamma}^{22}_{f,\Pi;\boldsymbol{\vartheta}_0} \end{pmatrix} \\ &= (k-1)^{-1} \begin{pmatrix} \mathcal{J}_k(f)(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top) & 2\Pi'(0)\mathcal{I}_k(f)\boldsymbol{\Upsilon}_{\boldsymbol{\theta}} \\ 2\Pi'(0)\mathcal{I}_k(f)\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top & 4(\Pi'(0))^2\mathcal{A}_k(f)\mathbf{I}_{k-1} \end{pmatrix}\end{aligned}$$

where $\mathcal{I}_k(f) = \int_{-1}^{1} \varphi_f(t)(1-t^2)\tilde{f}(t)dt$. The expression of $\mathbf{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0}$ follows from the independence between $\mathbf{X}^\top\boldsymbol{\theta}$ and $\mathbf{S}_{\boldsymbol{\theta}}(\mathbf{X})$ and the fact that the density of $\mathbf{X}^\top\boldsymbol{\theta}$ over $[-1, 1]$ is given by $\tilde{f}$ in (6). We have left out f and Π as indices whenever they do not appear in the related expressions.

A large amount of literature on skew-symmetric models on $\mathbb{R}^k$ has been dedicated to Fisher information singularity issues, in the vicinity of symmetry, which arise from an unfortunate combination of symmetric kernel f and skewing function $\Pi(\cdot,\cdot)$ in densities (2); see, e.g., [6]. We expose ourselves to the same risk here with the Fisher information for location and skewness. Therefore, an investigation of its singularity is necessary in order to avoid using mismatches between the angular function f on the one hand and Π on the other hand.

The Fisher information matrix $\mathbf{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0}$ is never full rank due to the curved nature of the parameter space ($\boldsymbol{\theta} \in \mathcal{S}^{k-1}$). Its rank can at most be $2k-2$; any lower rank implies a non-trivial singularity. Let $\ker(\mathbf{A})$ denote the kernel of some matrix $\mathbf{A}$. Considering

$$\mathbf{\Gamma}^{1.2}_{f,\Pi;\boldsymbol{\vartheta}_0} = \boldsymbol{\gamma}^{11}_{f;\boldsymbol{\vartheta}_0} - \boldsymbol{\gamma}^{12}_{f,\Pi;\boldsymbol{\vartheta}_0}\left(\boldsymbol{\gamma}^{22}_{f,\Pi;\boldsymbol{\vartheta}_0}\right)^{-1}\left(\boldsymbol{\gamma}^{12}_{f,\Pi;\boldsymbol{\vartheta}_0}\right)^\top,$$

we note that $\ker(\mathbf{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0})$ and $\ker(\mathbf{\Gamma}^{1.2}_{f,\Pi;\boldsymbol{\vartheta}_0})$ have the same dimension. This follows from similar arguments as in Lemma 2.1 of [14]. Easy computations yield that the rank of $\mathbf{\Gamma}^{1.2}_{f,\Pi;\boldsymbol{\vartheta}_0}$ is the highest possible (i.e., $k-1$) if and only if $\mathcal{J}_k(f)\mathcal{A}_k(f) - (\mathcal{I}_k(f))^2 \neq 0$. By a simple Cauchy–Schwarz argument, it can be seen that

$$\left\{\int_{-1}^{1} \varphi_f(t)(1-t^2)\tilde{f}(t)dt\right\}^2 \le \left\{\int_{-1}^{1} \varphi_f^2(t)(1-t^2)\tilde{f}(t)dt\right\}\left\{\int_{-1}^{1}(1-t^2)\tilde{f}(t)dt\right\}$$

with equality if and only if $\varphi_f(t) = c$ for some constant $c \in \mathbb{R}$. In other words, Fisher information singularity occurs if and only if $f(t) \propto \exp(ct)$, an FvML angular function with concentration parameter $c \in \mathbb{R}$ (negative values of c simply mean that data aggregate around $-\boldsymbol{\theta}$). Thus the famous FvML distribution is the only rotationally symmetric distribution to give rise to Fisher information singularity in our SRS models, mimicking the multivariate normal that also plays a central role in terms of singularity in certain skew-symmetric models; see [14,6].

4. Efficient tests for rotational symmetry about a fixed location

In this section we focus on the problem of testing rotational symmetry around a fixed center $\boldsymbol{\theta}$ within the class of SRS distributions defined in Section 2. More precisely, we build optimal testing procedures for the null hypothesis of rotational symmetry that can be written as $\mathcal{H}_0 : \boldsymbol{\delta} = \mathbf{0}$ within the class of distributions of Section 2 assuming that the location parameter $\boldsymbol{\theta}$ is known. This problem has been considered in the circular ($k = 2$) case in [16,23]. The construction of locally and asymptotically optimal testing procedures relies on the Uniform Local Asymptotic Normality (ULAN) property of a sequence of SRS models in the vicinity of symmetry. We establish this property in detail in Appendix A. For the sake of presentation, we restrict ourselves here to the description of inferential procedures that we can derive from ULAN.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ form a random sample on $\mathcal{S}^{k-1}$ with common density (3), and define

$$\mathcal{F} = \{f : [-1, 1] \to \mathbb{R}^+, f \text{ absolutely continuous}\}$$

the class of angular functions defining rotationally symmetric densities $f_{\boldsymbol{\theta};k}$. For any kernel $f_{\boldsymbol{\theta};k}$ with angular function f and any skewing function Π, denote by $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f,\Pi}$, with $\boldsymbol{\vartheta}$ still representing $(\boldsymbol{\theta}^\top, \boldsymbol{\delta}^\top)^\top \in \mathcal{S}^{k-1} \times \mathbb{R}^{k-1}$, the joint distribution of the n-tuple $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$. Since the density $f_{\boldsymbol{\theta},\boldsymbol{\delta};k}$ reduces to $c_{f,k}f(\mathbf{x}^\top\boldsymbol{\theta})$ when $\boldsymbol{\delta} = \mathbf{0}$ and hence does not depend on Π in this case, we drop the index Π and simply write $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f}$ at $\boldsymbol{\vartheta}_0 = (\boldsymbol{\theta}^\top, \mathbf{0}^\top)^\top$. Any couple (f, Π) then induces the parametric location-skewness model

$$\mathcal{P}^{(n)}_{f,\Pi} = \left\{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta};f,\Pi} : \boldsymbol{\vartheta} \in \mathcal{S}^{k-1} \times \mathbb{R}^{k-1}\right\}.$$

In Appendix A, we prove that this model is ULAN in the vicinity of symmetry (i.e., at $\boldsymbol{\delta} = \mathbf{0}$) for the score function and Fisher information matrix calculated in Section 3. The locally and asymptotically optimal testing procedures we shall provide are therefore based on the score-function type quantities $\mathbf{\Delta}^{(n)}_{f,\Pi}(\boldsymbol{\theta}) = n^{-1/2}\sum_{i=1}^{n} \boldsymbol{\ell}_{f,\Pi;\boldsymbol{\vartheta}_0}(\mathbf{X}_i)$, termed central sequences.

The construction is based on the $\boldsymbol{\delta}$-part of the central sequence $\mathbf{\Delta}^{(n)}_{f,\Pi}(\boldsymbol{\theta})$, namely

$$\begin{aligned}\mathbf{\Delta}^{(n)}_{\Pi;2}(\boldsymbol{\theta}) &= n^{-1/2}\sum_{i=1}^{n} \boldsymbol{\ell}^2_{\Pi;\boldsymbol{\vartheta}_0}(\mathbf{X}_i) = 2\Pi'(0)n^{-1/2}\sum_{i=1}^{n}\{1-(\mathbf{X}_i^\top\boldsymbol{\theta})^2\}^{1/2}\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{X}_i)\\ &= 2\Pi'(0)n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{X}_i = 2\Pi'(0)n^{1/2}\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\bar{\mathbf{X}},\end{aligned}$$

where $\bar{\mathbf{X}} = \sum_{i=1}^n \mathbf{X}_i/n$. Note that $\mathbf{\Delta}^{(n)}_{\Pi;2}(\boldsymbol{\theta})$ does not depend on the original angular function f. This is a crucial property, as we shall see in a few lines. Following Section 11.9 of [12], and in view of the fact that $\mathbf{\Delta}^{(n)}_{\Pi;2}(\boldsymbol{\theta})$ is asymptotically normal (see Theorem 1) with mean zero and covariance matrix

$$\boldsymbol{\gamma}^{22}_{f,\Pi;\boldsymbol{\vartheta}_0} = \frac{4(\Pi'(0))^2\mathcal{A}_k(f)}{k-1}\,\mathbf{I}_{k-1},$$

the locally and asymptotically maximin f-parametric procedure $\phi^{(n)}_f$ for testing $\mathcal{H}^f_0 : \boldsymbol{\delta} = \mathbf{0}$ against $\mathcal{H}^f_1 : \boldsymbol{\delta} \neq \mathbf{0}$ consists in rejecting $\mathcal{H}^f_0$ in favor of $\mathcal{H}^f_1$ at asymptotic level α whenever

$$\begin{aligned} T^{(n)}_f(\boldsymbol{\theta}) &= \{\mathbf{\Delta}^{(n)}_{\Pi;2}(\boldsymbol{\theta})\}^\top (\boldsymbol{\gamma}^{22}_{f,\Pi;\boldsymbol{\vartheta}_0})^{-1}\mathbf{\Delta}^{(n)}_{\Pi;2}(\boldsymbol{\theta}) \\ &= \frac{n(k-1)}{\mathcal{A}_k(f)}\,\bar{\mathbf{X}}^\top(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top)\bar{\mathbf{X}}, \end{aligned}$$

exceeds $\chi^2_{k-1;1-\alpha}$, the α-upper quantile of the chi-square distribution with $k-1$ degrees of freedom.

Now, in order to tackle the semi-parametric problem of interest $\mathcal{H}_0 = \cup_{f\in\mathcal{F}}\mathcal{H}^f_0$, we need to estimate the quantity $\mathcal{A}_k(f)$, in other words, to make the statistic $T^{(n)}_f(\boldsymbol{\theta})$ semi-parametric. This can be achieved via the consistent (under any f) estimator $\hat{\mathcal{A}}_k = 1 - \sum_{i=1}^n(\mathbf{X}_i^\top\boldsymbol{\theta})^2/n$, so that the resulting semi-parametric testing procedure $\phi^{(n)}_{\mathrm{Wat}}$ rejects $\mathcal{H}_0 : \boldsymbol{\delta} = \mathbf{0}$ at asymptotic level α whenever

$$T^{(n)} = \frac{n(k-1)}{\hat{\mathcal{A}}_k}\bar{\mathbf{X}}^\top(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top)\bar{\mathbf{X}} \tag{7}$$

exceeds $\chi^2_{k-1;1-\alpha}$. Quite interestingly, the test statistic (7) is exactly the statistic used in [28] to address the spherical location problem under rotational symmetry, hence our notation $\phi^{(n)}_{\mathrm{Wat}}$. Paindaveine and Verdebout [21] have established that the Watson test is the locally and asymptotically optimal pseudo-FvML test for the spherical location problem. Hence the Watson test, well known to be efficient for the spherical location problem in the FvML case, also happens to be an efficient test for rotational symmetry against SRS alternatives under specified $\boldsymbol{\theta}$. This result is further complemented by the fact that $T^{(n)}$ is the most efficient procedure under any angular function f; hence it is the uniformly optimal semi-parametric test for rotational symmetry against SRS alternatives under fixed $\boldsymbol{\theta}$. All this is summarized in the following proposition, where we also provide the asymptotic properties of $T^{(n)}$ under the null and under local alternatives.

Proposition 1. *Let $f \in \mathcal{F}$ and $\boldsymbol{\vartheta}_0 = (\boldsymbol{\theta}^\top, \mathbf{0}^\top)^\top$ with $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ fixed, and assume that Assumption A is valid. Then the following statements hold true.*

(i) *Under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f}$, $T^{(n)}$ is asymptotically chi-square with $k-1$ degrees of freedom.*

(ii) *Under $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0+(\mathbf{0},n^{-1/2}\mathbf{t}_2^{(n)});f,\Pi}$, with bounded sequence $\mathbf{t}_2^{(n)} \in \mathbb{R}^{k-1}$, $T^{(n)}$ is asymptotically non-central chi-square, still with $k-1$ degrees of freedom, and non-centrality parameter*

$$\frac{4(\Pi'(0))^2\mathcal{A}_k(f)}{k-1}\,\|\mathbf{t}_2\|^2,$$

with $\mathbf{t}_2 = \lim_{n\to\infty}\mathbf{t}_2^{(n)}$.

(iii) *The sequence of tests $\phi^{(n)}_{\mathrm{Wat}}$ has asymptotic size α under the entire null hypothesis $\mathcal{H}_0$.*

(iv) *$\phi^{(n)}_{\mathrm{Wat}}$ is locally asymptotically maximin, at asymptotic level α, when testing $\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_0;f}$ against alternatives of the form $\bigcup_{\boldsymbol{\delta}\neq\mathbf{0}}\mathrm{P}^{(n)}_{(\boldsymbol{\theta},\boldsymbol{\delta});f,\Pi}$.*

The proof is provided in Appendix B. The uniformly optimal test $\phi^{(n)}_{\mathrm{Wat}}$ extends the optimal test developed by Ley and Verdebout in [16] to any dimension and to the best of our knowledge, it is the first semi-parametric test for rotational symmetry.

5. Measure of skewness and related test for rotational symmetry about a fixed location

A natural way to measure rotational symmetry/asymmetry around $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ is to use a traditional multivariate measure of spherical symmetry/asymmetry computed on $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{X}$. As a consequence, there exist many distinct ways to measure skewness on the sphere; see [24] for examples. The following proposal is based on the Mardia [18] measure of skewness that is related to third-order moments. It leads to the measure

$$\rho_{\boldsymbol{\theta}} = \mathrm{E}[\{\mathbf{X}^\top(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top)\mathbf{Y}\}^3],$$

where $\mathbf{X}$ and $\mathbf{Y}$ are i.i.d. random vectors taking values on $\mathcal{S}^{k-1}$. Clearly if $\mathbf{X}$ and $\mathbf{Y}$ are rotationally symmetric around $\boldsymbol{\theta}$, then $\rho_{\boldsymbol{\theta}} = 0$. An empirical version of $\rho_{\boldsymbol{\theta}}$ is given by the U-statistic

$$\hat{\rho}_{\boldsymbol{\theta}} = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \{\mathbf{X}_i^\top (\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top)\mathbf{X}_j\}^3.$$

The following result provides the asymptotic behavior of $\hat{\rho}_{\boldsymbol{\theta}}$ under rotational symmetry.

Proposition 2. *Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ form a random sample on $\mathcal{S}^{k-1}$ with a common rotationally symmetric distribution with location $\boldsymbol{\theta}$ and angular function $f \in \mathcal{F}$. Letting $\mathrm{P}^{(n)}_{\boldsymbol{\theta};f}$ stand for the joint distribution of the $\mathbf{X}_i$'s and $e_\ell = \mathrm{E}_{\mathrm{P}^{(n)}_{\boldsymbol{\theta};f}}[\{1 - (\mathbf{X}_1^\top \boldsymbol{\theta})^2\}^{\ell/2}]$, we have that*

$$S_n = \frac{n\hat{\rho}_{\boldsymbol{\theta}}}{\sqrt{\frac{30 e_6^2}{(k-1)(k+1)(k+3)}}}$$

is asymptotically standard normal as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};f}$.

The proof is provided in Appendix B. Proposition 2 naturally yields a test of rotational symmetry around a fixed center $\boldsymbol{\theta}$. Since

$$\hat{e}_6 = \frac{1}{n}\sum_{i=1}^{n} \{1 - (\mathbf{X}_i^\top \boldsymbol{\theta})^2\}^3$$

is a consistent estimator of e_6 under the null hypothesis of rotational symmetry around $\boldsymbol{\theta}$, Slutsky's Lemma directly entails that

$$\hat{S}_n = \frac{n\hat{\rho}_{\boldsymbol{\theta}}}{\sqrt{\frac{30 \hat{e}_6^2}{(k-1)(k+1)(k+3)}}}$$

is still asymptotically standard normal under the null hypothesis of rotational symmetry. As a consequence, the test $\phi^{(n)}_{\mathrm{Skew}}$ that rejects (at the asymptotic nominal level α) the null hypothesis when

$$|\hat{S}_n| \ge z_{1-\alpha/2},$$

where z_ν stand for the quantile of order ν of a standard Gaussian random variable, is a valid test for rotational symmetry about $\boldsymbol{\theta}$. It thus represents an interesting alternative to the uniformly optimal test constructed in the previous section, and we shall compare their behaviors in the next section.

6. Simulation study

In this section, the objectives of our Monte Carlo simulation study are twofold: (i) show that both tests $\phi^{(n)}_{\mathrm{Wat}}$ and $\phi^{(n)}_{\mathrm{Skew}}$ hold their nominal level at finite sample sizes, and (ii) exhibit their empirical powers under various alternatives to rotational symmetry.

To do so, we generated, for $n = 50$ and $n = 100$, $M = 2\,500$ independent random samples

$$\mathbf{X}^{(j)}_{i,\ell}, \quad i \in \{1, \ldots, n\},\ j \in \{1, 2\},\ \ell \in \{0, \ldots, 5\},$$

from the following spherical (3-dimensional) distributions:

(i) the $\mathbf{X}^{(1)}_{i,\ell}$'s have a common skew-FvML distribution with location $\boldsymbol{\theta} = (1, 0, 0)^\top$, concentration $\kappa = 5$ and skewness parameter $\boldsymbol{\delta} = (\ell/5, \ell/5)^\top$ (the skewing function is the standard Gaussian cdf);

(ii) the $\mathbf{X}^{(2)}_{i,\ell}$'s have a common mixture of FvML distributions of the form

$$\left(1 - \frac{\ell}{10}\right)\mathbf{Y} + \frac{\ell}{10}\mathbf{Z},$$

where $\mathbf{Y}$ is a FvML random vector with location $\boldsymbol{\theta} = (1, 0, 0)^\top$ and concentration $\kappa = 5$ and $\mathbf{Z}$ is a FvML random vector with location $(\sqrt{2}/2, \sqrt{2}/2, 0)^\top$ and concentration $\kappa = 5$.

In both cases (i) and (ii), the value $\ell = 0$ corresponds to the null hypothesis of rotational symmetry around $\boldsymbol{\theta} = (1, 0, 0)^\top$ while the values $\ell = 1, \ldots, 5$ represent distributions that are increasingly skewed.

In Figs. 2–5, we provide the empirical rejection frequencies of the tests $\phi^{(n)}_{\mathrm{Wat}}$ and $\phi^{(n)}_{\mathrm{Skew}}$ performed at the asymptotic nominal level $\alpha = .05$ for sample sizes $n = 50$ and $n = 100$. Inspection of those Figures reveals that both $\phi^{(n)}_{\mathrm{Wat}}$ and $\phi^{(n)}_{\mathrm{Skew}}$ reach the nominal level constraint. While as expected $\phi^{(n)}_{\mathrm{Wat}}$ dominates $\phi^{(n)}_{\mathrm{Skew}}$ under the SRS alternatives of Section 2, $\phi^{(n)}_{\mathrm{Skew}}$ slightly dominates $\phi^{(n)}_{\mathrm{Wat}}$ under mixtures of FvML distributions.

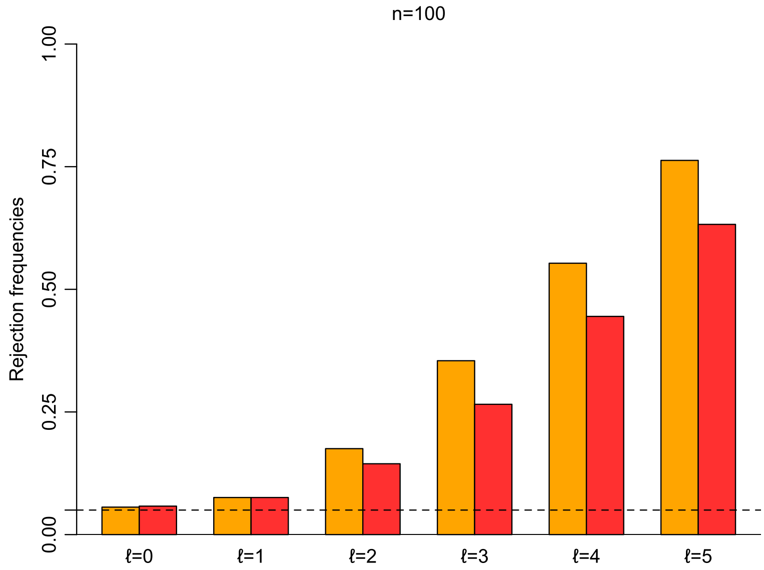

Fig. 2. Empirical rejection frequencies of the uniformly optimal test $\phi^{(n)}_{\mathrm{Wat}}$ (in orange) and of the skewness measure based test $\phi^{(n)}_{\mathrm{Skew}}$ (in red) under 3-dimensional skew-FvML distributions with location $\boldsymbol{\theta} = (1, 0, 0)^\top$, concentration $\kappa = 5$ and skewness parameter $\boldsymbol{\delta} = (\ell/5, \ell/5)^\top$, $\ell = 0, \ldots, 5$. The sample size is $n = 50$ and the nominal level is $\alpha = .05$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

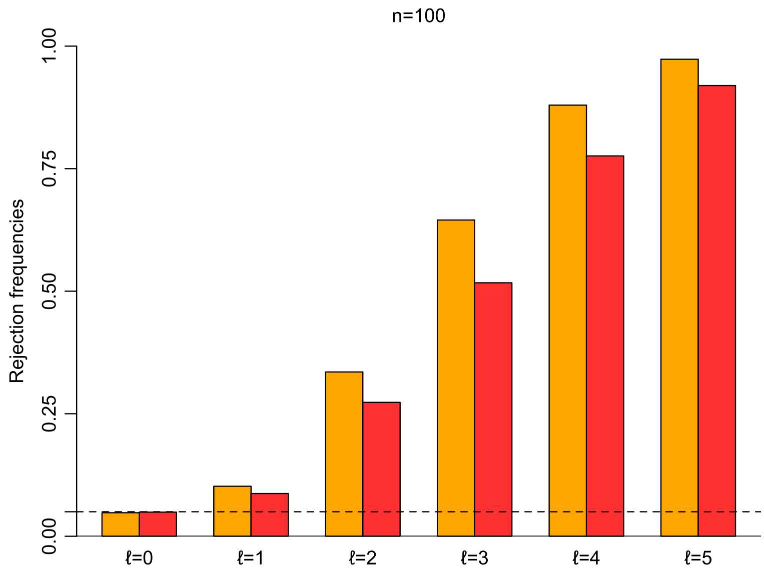

Fig. 3. Empirical rejection frequencies of the uniformly optimal test $\phi^{(n)}_{\mathrm{Wat}}$ (in orange) and of the skewness measure based test $\phi^{(n)}_{\mathrm{Skew}}$ (in red) under 3-dimensional skew-FvML distributions with location $\boldsymbol{\theta} = (1, 0, 0)^\top$, concentration $\kappa = 5$ and skewness parameter $\boldsymbol{\delta} = (\ell/5, \ell/5)^\top$, $\ell = 0, \ldots, 5$. The sample size is $n = 100$ and the nominal level is $\alpha = .05$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

7. Discussion

We conclude the paper by discussing open questions and future research issues related to skew-rotationally-symmetric models.

We have not investigated the number of modes of our skew models. Starting from a unimodal rotationally symmetric density with angular function f, will the resulting skew-f density be unimodal for all values of the skewness parameter $\boldsymbol{\delta}$? Or will multi-modality occur above a certain threshold value? In the circular setting, Abe and Pewsey [1] have shown that the sine-skewed von Mises density is either unimodal or bimodal, depending on the amount of skewness, while the sine-skewed wrapped Cauchy density is always unimodal. Obtaining the number of modes in the circular case is not an easy task

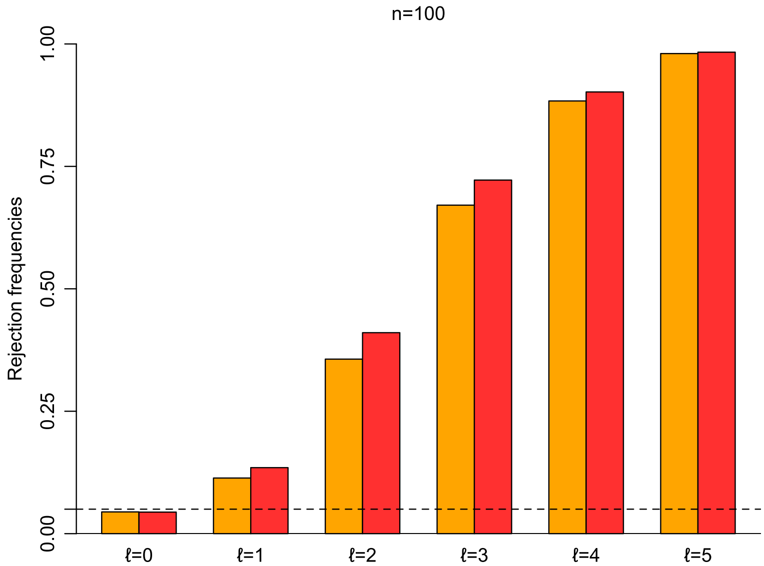

Fig. 4. Empirical rejection frequencies of the uniformly optimal test $\phi^{(n)}_{\mathrm{Wat}}$ (in orange) and of the skewness measure based test $\phi^{(n)}_{\mathrm{Skew}}$ (in red) under the 3-dimensional FvML mixture $(1-\ell/10)\,\mathbf{Y}+\ell\mathbf{Z}/10$, where $\mathbf{Y}$ has location $\boldsymbol{\theta}=(1,0,0)^\top$ and concentration $\kappa=5$ and $\mathbf{Z}$ has location $(\sqrt{2}/2,\sqrt{2}/2,0)^\top$ and concentration $\kappa=5$, for $\ell=0,\ldots,5$. The sample size is $n=50$ and the nominal level is $\alpha=.05$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

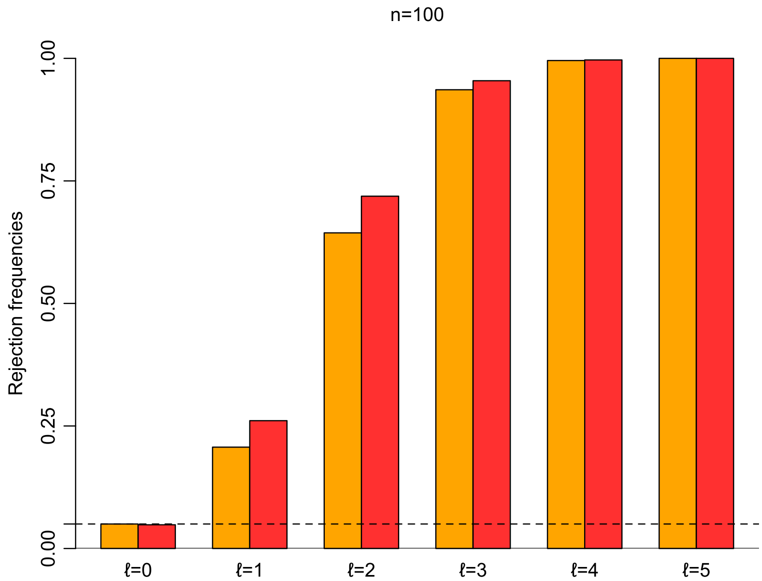

Fig. 5. Empirical rejection frequencies of the uniformly optimal test $\phi^{(n)}_{\mathrm{Wat}}$ (in orange) and of the skewness measure based test $\phi^{(n)}_{\mathrm{Skew}}$ (in red) under the 3-dimensional FvML mixture $(1-\ell/10)\,\mathbf{Y}+\ell\mathbf{Z}/10$, where $\mathbf{Y}$ has location $\boldsymbol{\theta}=(1,0,0)^\top$ and concentration $\kappa=5$ and $\mathbf{Z}$ has location $(\sqrt{2}/2,\sqrt{2}/2,0)^\top$ and concentration $\kappa=5$, for $\ell=0,\ldots,5$. The sample size is $n=100$ and the nominal level is $\alpha=.05$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

since it is determined by studying the discriminant of a quartic equation. The problem is even more complicated to solve in the $k\geq 3$ case and is left for future research.

A further issue of interest is the existence of mechanistic derivations of the SRS distributions. Skew-symmetric distributions on $\mathbb{R}^k$ enjoy such properties. For example, the multivariate skew-normal can be built as follows. Consider a scalar random variable Z_0 and a random k-vector $\mathbf{Z}$ such that $(Z_0,\mathbf{Z})^\top$ follows a $(k+1)$-dimensional normal distribution. Then $\mathbf{Z}\mid Z_0>0$ is k-dimensional skew-normal. Deriving a similar generating mechanism for SRS densities is a challenging problem. The additional difficulty here lies in the fact that simple conditioning leads to a vector that does not necessarily lie on the unit sphere. A potential solution might be to consider nested spheres as defined in [9] in the context of principal nested spheres.

Finally, we return to the Fisher information singularity studied in Section 3. As for the normal distribution on $\mathbb{R}^k$, the FvML is the most used distribution on the sphere. It is therefore natural to search for ways to overcome this singularity, as was done for the multivariate skew-normal; see, e.g., Arellano-Valle and Azzalini [2]. A reparameterization à la Hallin and Ley [7] could solve the problem, but does not seem straightforward. The natural curved parameter space ($\boldsymbol{\theta} \in \mathcal{S}^{k-1}$) used here makes the problem more challenging and is therefore left as an open question. An alternative approach would consist in using a distinct skewing function for FvML densities. Recall that the argument of the skewing function Π in (4) is given by $\{1-(\mathbf{X}^\top\boldsymbol{\theta})^2\}^{1/2}\boldsymbol{\delta}^\top\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{X}) = \boldsymbol{\delta}^\top\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{X}$. Since, under rotational symmetry, $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{X}$ is spherically symmetric on the unit ball of $\mathbb{R}^{k-1}$, the mechanism we define in (4) actually breaks this spherical symmetry. Now, the way we skew $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{X}$ does not depend on the value of the projection $\mathbf{X}^\top\boldsymbol{\theta}$ of $\mathbf{X}$ along $\boldsymbol{\theta}$. Of course one could consider skewing mechanisms that depend on $\mathbf{X}^\top\boldsymbol{\theta}$; skewing the distribution more in the vicinity of $\boldsymbol{\theta}$ than around the opposite pole $-\boldsymbol{\theta}$ may be more appropriate in certain situations. A possible way to achieve this is to consider a skewing function of the form (4) with

$$(\mathbf{X}^\top\boldsymbol{\theta})\boldsymbol{\delta}^\top\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top\mathbf{X}$$

for the argument of the function Π (of course any increasing function of $(\mathbf{X}^\top\boldsymbol{\theta})$ can be considered). For such a mechanism, the score function in the vicinity of $\boldsymbol{\delta} = \mathbf{0}$ takes the form

$$\begin{aligned}\boldsymbol{\ell}_{f,\Pi;\boldsymbol{\vartheta}_0}(\mathbf{x}) &= \mathrm{grad}_{\boldsymbol{\vartheta}}\log f_{\boldsymbol{\vartheta};k}(\mathbf{x})|_{\boldsymbol{\vartheta}_0} = \left(\{\boldsymbol{\ell}^1_{f;\boldsymbol{\vartheta}_0}(\mathbf{x})\}^\top,\ \{\boldsymbol{\ell}^2_{\Pi;\boldsymbol{\vartheta}_0}(\mathbf{x})\}^\top\right)^\top \\ &= \begin{pmatrix}\varphi_f(\mathbf{x}^\top\boldsymbol{\theta})\{1-(\mathbf{x}^\top\boldsymbol{\theta})^2\}^{1/2}\mathbf{S}_{\boldsymbol{\theta}}(\mathbf{x}) \\ 2\Pi'(\mathbf{x}^\top\boldsymbol{\theta})\{1-(\mathbf{x}^\top\boldsymbol{\theta})^2\}^{1/2}\mathbf{U}_{\boldsymbol{\theta}}(\mathbf{x})\end{pmatrix}\end{aligned}$$

and the corresponding Fisher information matrix is given by

$$(k-1)^{-1}\begin{pmatrix}\mathcal{J}_k(f)(\mathbf{I}_k-\boldsymbol{\theta}\boldsymbol{\theta}^\top) & 2\Pi'(0)\mathcal{I}_{k;mod}(f)\boldsymbol{\Upsilon}_{\boldsymbol{\theta}} \\ 2\Pi'(0)\mathcal{I}_{k;mod}(f)\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top & 4(\Pi'(0))^2\mathcal{A}_{k;mod}(f)\mathbf{I}_{k-1}\end{pmatrix},$$

where the modified quantities $\mathcal{I}_{k;mod}(f)$ and $\mathcal{A}_{k;mod}(f)$ are now respectively equal to

$$\int_{-1}^1 \varphi_f(t)t(1-t^2)\tilde{f}(t)dt \quad \text{and} \quad \int_{-1}^1 t^2(1-t^2)\tilde{f}(t)dt.$$

This Fisher information matrix is no longer singular at the FvML distribution so that such a choice of mechanism cancels the singularity issue uncovered in the previous section for the FvML. Given the popularity of the FvML distribution among theoreticians and practitioners, it seems worthwhile to elaborate a meaningful skew-FvML distribution.

Acknowledgments

Thomas Verdebout thanks the Banque Nationale de Belgique for support via a research grant. Both authors thank the Editor-in-Chief, an Associate Editor and two anonymous referees for helpful comments that led to a clear improvement of the paper.

Appendix A. The ULAN property of SRS distributions

The main technical tool in our construction of tests for rotational symmetry consists in establishing the uniform local asymptotic normality (ULAN) property, in the vicinity of symmetry (i.e., at $\boldsymbol{\delta} = \mathbf{0}$), of the parametric model $\mathcal{P}^{(n)}_{f,\Pi}$. This property will allow us to build optimal parametric testing procedures in line with the Le Cam methodology; see [12]. This is not an easy task due to the "curved" nature of the parameter space for $\boldsymbol{\theta}$: the unit hypersphere $\mathcal{S}^{k-1}$ 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. An extension of Le Cam's theory to this spherical setting has been given in [15], allowing us to establish the following ULAN property for SRS distributions.

Theorem 1. *Let $f \in \mathcal{F}$ and assume that Assumption A holds. Then, for any $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$, the parametric family of densities $\mathcal{P}^{(n)}_{f,\Pi}$ is ULAN at $\boldsymbol{\vartheta}_0 = (\boldsymbol{\theta}^\top, \mathbf{0}^\top)^\top$ with central sequence $\boldsymbol{\Delta}^{(n)}_{f,\Pi}(\boldsymbol{\theta}) = \sum_{i=1}^n \boldsymbol{\ell}_{f,\Pi;\boldsymbol{\vartheta}_0}(\mathbf{X}_i)/\sqrt{n}$ and corresponding Fisher information matrix $\boldsymbol{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0}$. More precisely, for any $\boldsymbol{\theta}^{(n)} \in \mathcal{S}^{k-1}$ such that $\boldsymbol{\theta}^{(n)} - \boldsymbol{\theta} = O(n^{-1/2})$, any bounded sequence $\mathbf{t}_1^{(n)} \in \mathbb{R}^k$ such that $\boldsymbol{\theta}^{(n)} + \mathbf{t}_1^{(n)}/\sqrt{n}$ remains in $\mathcal{S}^{k-1}$ and any bounded sequence $\mathbf{t}_2^{(n)} \in \mathbb{R}^{k-1}$, we have, letting*

$$\lambda^{(n)} = \ln\left(d\mathrm{P}^{(n)}_{(\boldsymbol{\theta}^{(n)}+n^{-1/2}\mathbf{t}_1^{(n)},n^{-1/2}\mathbf{t}_2^{(n)});f,\Pi}/d\mathrm{P}^{(n)}_{(\boldsymbol{\theta}^{(n)},\mathbf{0});f}\right),$$

$$\lambda^{(n)} = ((\mathbf{t}_1^{(n)})^\top, (\mathbf{t}_2^{(n)})^\top)\boldsymbol{\Delta}^{(n)}_{f,\Pi}(\boldsymbol{\theta}^{(n)}) - \frac{1}{2}((\mathbf{t}_1^{(n)})^\top, (\mathbf{t}_2^{(n)})^\top)\boldsymbol{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0}((\mathbf{t}_1^{(n)})^\top, (\mathbf{t}_2^{(n)})^\top)^\top + o_{\mathrm{P}}(1)$$

and $\boldsymbol{\Delta}^{(n)}_{f,\Pi}(\boldsymbol{\theta}^{(n)}) \rightsquigarrow \mathcal{N}_{2k-1}(\mathbf{0}, \boldsymbol{\Gamma}_{f,\Pi;\boldsymbol{\vartheta}_0})$, both under $\mathrm{P}^{(n)}_{(\boldsymbol{\theta}^{(n)},\mathbf{0});f}$ as $n \to \infty$.

Proof of Theorem 1. 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

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

with associated Jacobian matrix $D\hbar(\boldsymbol{\eta})$; see [15] for its expression. In a first step, we need to prove ULAN for the $(\boldsymbol{\eta}, \boldsymbol{\delta})$-parameterization. By Theorem 7.2 of [26], it suffices to show that

$$(\boldsymbol{\eta}, \boldsymbol{\delta}) \mapsto f^{1/2}_{\boldsymbol{\eta},\boldsymbol{\delta};k}(\mathbf{x}) = (2c_{f,k})^{1/2} f^{1/2}\{\mathbf{x}^\top h(\boldsymbol{\eta})\} \Pi^{1/2}(\boldsymbol{\delta}^\top \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta})}\mathbf{x})$$

is quadratic mean differentiable at $(\boldsymbol{\eta}^\top, \mathbf{0}^\top)^\top$ for any $\boldsymbol{\eta} \in \mathbb{R}^{k-1}$. This is achieved in the subsequent lemma.

Lemma 3. *Let $f \in \mathcal{F}$ and assume that Assumption A holds. Define*

$$\mathrm{grad}_{\boldsymbol{\eta}} f^{1/2}_{\boldsymbol{\eta},\mathbf{0};k}(\mathbf{x}) = \frac{1}{2} f^{1/2}_{\boldsymbol{\eta};k}(\mathbf{x}) \varphi_f(\mathbf{x}^\top \hbar(\boldsymbol{\eta})) D\hbar(\boldsymbol{\eta})^\top \mathbf{x},$$

and

$$\mathrm{grad}_{\boldsymbol{\delta}} f^{1/2}_{\boldsymbol{\eta},\boldsymbol{\delta};k}(\mathbf{x})|_{\boldsymbol{\delta}=\mathbf{0}} = f^{1/2}_{\boldsymbol{\eta};k}(\mathbf{x}) \Pi'(0) \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta})}\mathbf{x}.$$

Then, for any $\boldsymbol{\eta} \in \mathbb{R}^{k-1}$, we have that, as $(\mathbf{e}, \mathbf{d}) \to (\mathbf{0}, \mathbf{0})$,

(i) $\int_{\mathcal{S}^{k-1}} \left\{ f^{1/2}_{\boldsymbol{\eta}+\mathbf{e},\mathbf{0};k}(\mathbf{x}) - f^{1/2}_{\boldsymbol{\eta},\mathbf{0};k}(\mathbf{x}) - \mathbf{e}^\top \mathrm{grad}_{\boldsymbol{\eta}} f^{1/2}_{\boldsymbol{\eta},\mathbf{0};k}(\mathbf{x}) \right\}^2 d\sigma_{k-1}(\mathbf{x}) = o(\|\mathbf{e}\|^2)$;

(ii) $\int_{\mathcal{S}^{k-1}} \left\{ f^{1/2}_{\boldsymbol{\eta}+\mathbf{e},\mathbf{d};k}(\mathbf{x}) - f^{1/2}_{\boldsymbol{\eta}+\mathbf{e},\mathbf{0};k}(\mathbf{x}) - \mathbf{d}^\top \mathrm{grad}_{\boldsymbol{\delta}} f^{1/2}_{\boldsymbol{\eta}+\mathbf{e},\boldsymbol{\delta};k}(\mathbf{x})|_{\boldsymbol{\delta}=\mathbf{0}} \right\}^2 d\sigma_{k-1}(\mathbf{x}) = o(\|\mathbf{d}\|^2)$;

(iii) $\int_{\mathcal{S}^{k-1}} \left[\mathbf{d}^\top \left\{ \mathrm{grad}_{\boldsymbol{\delta}} f^{1/2}_{\boldsymbol{\eta}+\mathbf{e},\boldsymbol{\delta};k}(\mathbf{x})|_{\boldsymbol{\delta}=\mathbf{0}} - \mathrm{grad}_{\boldsymbol{\delta}} f^{1/2}_{\boldsymbol{\eta},\boldsymbol{\delta};k}(\mathbf{x})|_{\boldsymbol{\delta}=\mathbf{0}} \right\} \right]^2 d\sigma_{k-1}(\mathbf{x}) = o(\|\mathbf{d}\|^2)$;

(iv) $\int_{\mathcal{S}^{k-1}} \left\{ f^{1/2}_{\boldsymbol{\eta}+\mathbf{e},\mathbf{d};k}(\mathbf{x}) - f^{1/2}_{\boldsymbol{\eta},\mathbf{0};k}(\mathbf{x}) - \begin{pmatrix} \mathbf{e} \\ \mathbf{d} \end{pmatrix}^\top \begin{pmatrix} \mathrm{grad}_{\boldsymbol{\eta}} f^{1/2}_{\boldsymbol{\eta},\mathbf{0};k}(\mathbf{x}) \\ \mathrm{grad}_{\boldsymbol{\delta}} f^{1/2}_{\boldsymbol{\eta},\boldsymbol{\delta};k}(\mathbf{x})|_{\boldsymbol{\delta}=\mathbf{0}} \end{pmatrix} \right\}^2 d\sigma_{k-1}(\mathbf{x}) = o(\|(\mathbf{e}^\top, \mathbf{d}^\top)^\top\|^2)$.

Proof of Lemma 3. (i) This result has been established in [15].

(ii) By definition of $f_{\boldsymbol{\eta}+\mathbf{e},\mathbf{d};k}$, we can rewrite the left-hand side integral in (ii) as

$$c_{f,k} \int_{\mathcal{S}^{k-1}} f\{\mathbf{x}^\top h(\boldsymbol{\eta}+\mathbf{e})\} \left\{ 2^{1/2} \Pi^{1/2}(\mathbf{d}^\top \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta}+\mathbf{e})}\mathbf{x}) - 1 - \mathbf{d}^\top \Pi'(0) \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta}+\mathbf{e})}\mathbf{x} \right\}^2 d\sigma_{k-1}(\mathbf{x}).$$

The absolute continuity (and hence a.e.-differentiability) of the skewing function Π yields that

$$\sup_{\mathbf{x}\in\mathcal{S}^{k-1}} \left| 2^{1/2} \Pi^{1/2}(\mathbf{d}^\top \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta}+\mathbf{e})}\mathbf{x}) - 1 - \mathbf{d}^\top \Pi'(0) \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta}+\mathbf{e})}\mathbf{x} \right|$$

is $o(\|\mathbf{d}\|)$ uniformly in $\mathbf{x}$. Consequently, the second factor of the integrand is $o(\|\mathbf{d}\|^2)$ uniformly in $\mathbf{x}$. The result then follows since $\int_{\mathcal{S}^{k-1}} f\{\mathbf{x}^\top h(\boldsymbol{\eta}+\mathbf{e})\} d\sigma_{k-1}(\mathbf{x})$ is finite.

(iii) The left-hand side in (iii) equals

$$(\Pi'(0))^2 \int_{\mathcal{S}^{k-1}} \left[\mathbf{d}^\top \left\{ f^{1/2}_{\boldsymbol{\eta}+\mathbf{e};k}(\mathbf{x}) \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta}+\mathbf{e})}\mathbf{x} - f^{1/2}_{\boldsymbol{\eta};k}(\mathbf{x}) \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta})}\mathbf{x} \right\} \right]^2 d\sigma_{k-1}(\mathbf{x}).$$

Since $f^{1/2}_{\boldsymbol{\eta};k}(\mathbf{x}) \mathbf{d}^\top \boldsymbol{\Upsilon}^\top_{h(\boldsymbol{\eta})}\mathbf{x}$ is square-integrable by Assumption A, the quadratic mean continuity entails that the above expression is an $o(\|\mathbf{d}\|^2)$ quantity.

(iv) The left-hand side in (iv) is bounded by a constant times each of the three integrals treated before. The result then follows from (i), (ii) and (iii). □

Proof of Theorem 1 (*Continued*)**.** The quadratic mean differentiability for the $(\boldsymbol{\eta}, \boldsymbol{\delta})$-parameterization yields the ULAN result in that parameterization. The transposition from here to ULAN of the targeted $(\boldsymbol{\theta}, \boldsymbol{\delta})$-parameterization is obtained exactly along the same lines as in [15], which concludes the proof of the ULAN property. □

Appendix B. Proofs

Proof of Lemma 1. Writing, for the sake of simplicity, $g_{\boldsymbol{\theta},\boldsymbol{\delta}}(\mathbf{x})$ for $\boldsymbol{\delta}^\top \boldsymbol{\Upsilon}^\top_{\boldsymbol{\theta}}\mathbf{x}$, straightforward calculations give

$$\int_{\mathcal{S}^{k-1}} f_{\boldsymbol{\theta},\boldsymbol{\delta};k}(\mathbf{x}) \, d\sigma_{k-1}(\mathbf{x}) = \int_{g_{\boldsymbol{\theta},\boldsymbol{\delta}}(\mathbf{x})\geq 0} f_{\boldsymbol{\theta},\boldsymbol{\delta};k}(\mathbf{x}) \, d\sigma_{k-1}(\mathbf{x}) + \int_{g_{\boldsymbol{\theta},\boldsymbol{\delta}}(\mathbf{x})<0} f_{\boldsymbol{\theta},\boldsymbol{\delta};k}(\mathbf{x}) \, d\sigma_{k-1}(\mathbf{x})$$

$$
\begin{aligned}
&= 2\int_{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})\geq 0} c_{f,k} f(\mathbf{x}^\top\boldsymbol\theta)\Pi\{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})\}\, d\sigma_{k-1}(\mathbf{x})\\
&\quad + 2\int_{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})> 0} c_{f,k} f(\mathbf{x}^\top\boldsymbol\theta)[1-\Pi\{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})\}]\, d\sigma_{k-1}(\mathbf{x})\\
&= 2\int_{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})= 0} \frac{1}{2} c_{f,k} f(\mathbf{x}^\top\boldsymbol\theta)\, d\sigma_{k-1}(\mathbf{x}) + 2\int_{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})> 0} c_{f,k} f(\mathbf{x}^\top\boldsymbol\theta)\, d\sigma_{k-1}(\mathbf{x})\\
&= \int_{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})= 0} c_{f,k} f(\mathbf{x}^\top\boldsymbol\theta)\, d\sigma_{k-1}(\mathbf{x}) + \int_{g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})\neq 0} c_{f,k} f(\mathbf{x}^\top\boldsymbol\theta)\, d\sigma_{k-1}(\mathbf{x})\\
&= 1,
\end{aligned}
$$

where we have strongly used the fact that $f(\mathbf{x}^\top\boldsymbol\theta)$ takes the same values on both hemispheres $\{\mathbf{x}\in\mathcal{S}^{k-1} : g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})>0\}$ and $\{\mathbf{x}\in\mathcal{S}^{k-1} : g_{\boldsymbol\theta,\boldsymbol\delta}(\mathbf{x})<0\}$. □

Proof of Lemma 2. We prove that the pdf of $\mathbf{X}$ defined via (5) is $f_{\boldsymbol\theta,\boldsymbol\delta;k}$. First, conditionally on $\mathbf{Y}^\top\boldsymbol\theta = \mathbf{y}^\top\boldsymbol\theta = t\in[-1,1]$, the density of the sign vector $\mathbf{U}_{\boldsymbol\theta;\Pi}(\mathbf{Y})$ on $\mathcal{S}^{k-2}$ corresponds to $2(\omega_{k-1})^{-1}\Pi\{(1-t^2)^{1/2}\boldsymbol\delta^\top\mathbf{s}\}$ with $\mathbf{s}=\mathbf{U}_{\boldsymbol\theta}(\mathbf{y})$; this follows from skew-symmetric theory on $\mathbb{R}^k$; see, e.g., [27]. In what follows we will write out the probability that $\mathbf{X}$ belongs to a certain set $\mathcal{B}\subseteq\mathcal{S}^{k-1}$. To this end, we write $\ell_{\mathcal{B}}$ as the subset of $[-1,1]$ that the projection $\mathbf{X}^\top\boldsymbol\theta \overset{d}{=} \mathbf{Y}^\top\boldsymbol\theta$ has to cover, and we denote by $\mathcal{B}_t^\perp$ the subset of $\mathcal{S}^{k-2}$ that the sign vector $\mathbf{U}_{\boldsymbol\theta;\Pi}(\mathbf{Y})$ has to cover when $\mathbf{X}^\top\boldsymbol\theta = t\in\ell_{\mathcal{B}}$. With these notations and with the definition of $\tilde f$ from (6), we have

$$
\begin{aligned}
\Pr[\mathbf{X}\in\mathcal{B}] &= \Pr[(\mathbf{Y}^\top\boldsymbol\theta)\boldsymbol\theta + \{1-(\mathbf{Y}^\top\boldsymbol\theta)^2\}^{1/2}\boldsymbol\Upsilon_{\boldsymbol\theta}\mathbf{U}_{\boldsymbol\theta;\Pi}(\mathbf{Y})\in\mathcal{B}]\\
&= \int_{\ell_{\mathcal{B}}} \Pr\{t\boldsymbol\theta + (1-t^2)^{1/2}\boldsymbol\Upsilon_{\boldsymbol\theta}\mathbf{U}_{\boldsymbol\theta;\Pi}(\mathbf{Y})\in\mathcal{B}\mid \mathbf{Y}^\top\boldsymbol\theta = t\}\tilde f(t)dt\\
&= \int_{\ell_{\mathcal{B}}} \Pr\{\mathbf{U}_{\boldsymbol\theta;\Pi}(\mathbf{Y})\in\mathcal{B}_t^\perp\mid \mathbf{Y}^\top\boldsymbol\theta = t\}\tilde f(t)dt\\
&= \int_{\ell_{\mathcal{B}}}\int_{\mathcal{B}_t^\perp} 2(\omega_{k-1})^{-1}\Pi\{(1-t^2)^{1/2}\boldsymbol\delta^\top\mathbf{s}\}\tilde f(t)d\sigma_{k-2}(\mathbf{s})dt\\
&= \int_{\ell_{\mathcal{B}}}\int_{\mathcal{B}_t^\perp} \frac{1}{\omega_{k-1}}2c_{k,f}f(t)\Pi\{(1-t^2)^{1/2}\boldsymbol\delta^\top\mathbf{s}\}\frac{\omega_k(1-t^2)^{(k-3)/2}}{\mathcal{B}(\frac12,\frac{k-1}{2})}d\sigma_{k-2}(\mathbf{s})dt\\
&= \int_{\ell_{\mathcal{B}}}\int_{\mathcal{B}_t^\perp} 2c_{k,f}f(t)\Pi\{(1-t^2)^{1/2}\boldsymbol\delta^\top\mathbf{s}\}(1-t^2)^{(k-3)/2}d\sigma_{k-2}(\mathbf{s})dt,
\end{aligned}
$$

where we have used the fact that $\omega_k/\omega_{k-1} = \mathcal{B}(\frac12,\frac{k-1}{2})$ in the last equality. The final integral expression corresponds exactly to what one would obtain by starting from the pdf $f_{\boldsymbol\theta,\boldsymbol\delta;k}$ and performing the change of variables $d\sigma_{k-1}(\mathbf{y}) = (1-t^2)^{(k-3)/2}dtd\sigma_{k-2}(\mathbf{s})$ (see [28, p. 44, Eq. (2.2.2)]), hence proving the stochastic representation to be correct. □

Proof of Proposition 1. Point (i) follows directly from Slutsky's Lemma combined with

$$
\hat{\mathcal{A}}_k - \mathcal{A}_k(f) = o_{\mathrm{P}}(1) \tag{B.1}
$$

as $n\to\infty$ and the fact that $\boldsymbol\Delta^{(n)}_{\Pi;2}(\boldsymbol\theta)$ is asymptotically normal with mean zero and covariance matrix $4\{\Pi'(0)\}^2\mathcal{A}_k(f)\,\mathbf{I}_{k-1}/(k-1)$, both under $\mathrm{P}^{(n)}_{(\boldsymbol\theta,\mathbf{0});f}$. Now the multivariate central limit theorem entails that

$$
\left((\boldsymbol\Delta^{(n)}_{\Pi;2}(\boldsymbol\theta))^\top, (\mathbf{t}_2^{(n)})^\top\boldsymbol\Delta^{(n)}_{\Pi;2}(\boldsymbol\theta)\right)^\top
$$

is asymptotically normal with mean zero and covariance matrix

$$
\begin{pmatrix}
\boldsymbol\gamma^{22}_{f,\Pi;\boldsymbol\vartheta_0} & \boldsymbol\gamma^{22}_{f,\Pi;\boldsymbol\vartheta_0}\mathbf{t}_2\\
\mathbf{t}_2^\top\boldsymbol\gamma^{22}_{f,\Pi;\boldsymbol\vartheta_0} & \mathbf{t}_2^\top\boldsymbol\gamma^{22}_{f,\Pi;\boldsymbol\vartheta_0}\mathbf{t}_2
\end{pmatrix}
$$

under $\mathrm{P}^{(n)}_{(\boldsymbol\theta,\mathbf{0});f}$. Le Cam's Third Lemma then implies that the central sequence $\boldsymbol\Delta^{(n)}_{\Pi;2}(\boldsymbol\theta)$ is asymptotically $\mathcal{N}_{k-1}\big(\boldsymbol\gamma^{22}_{f,\Pi;\boldsymbol\vartheta_0}\mathbf{t}_2, \boldsymbol\gamma^{22}_{f,\Pi;\boldsymbol\vartheta_0}\big)$ under $\mathrm{P}^{(n)}_{(\boldsymbol\theta,n^{-1/2}\mathbf{t}_2^{(n)});f}$.

Since (B.1) holds as well asymptotically under $\mathrm{P}^{(n)}_{(\boldsymbol\theta,n^{-1/2}\mathbf{t}_2^{(n)});f}$ by contiguity, Point (ii) follows. Finally, the fact that $\phi^{(n)}_{\mathrm{Wat}}$ has asymptotic level α is a direct consequence of the asymptotic null distribution given in Part (i), while local asymptotic maximinity is a consequence of the weak convergence of the local experiments to the Gaussian shift experiment, see the explanations in [17] for more details. □

Proof of Proposition 2. In this proof, we put $u_i = \sqrt{1-(\mathbf{X}_i^\top \boldsymbol{\theta})^2}$. First note that we have

$$\hat{\rho}_{\boldsymbol{\theta}} = \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \{\mathbf{X}_i^\top(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}^\top)\mathbf{X}_j\}^3$$
$$= \frac{2}{n(n-1)} \sum_{1\le i<j\le n} u_i^3 u_j^3 \{(\mathbf{U}_i(\boldsymbol{\theta}))^\top \mathbf{U}_j(\boldsymbol{\theta})\}^3.$$

In the rest of the proof we put $\rho_{ij} = \{\mathbf{U}_i(\boldsymbol{\theta})\}^\top \mathbf{U}_j(\boldsymbol{\theta})$. The independence between the u_i's and the $\mathbf{U}_i(\boldsymbol{\theta})$'s together with the fact that $\mathrm{E}(\rho_{ij}^3) = 0$ (Lemma A.1(ii) in [22]) entail that $\mathrm{E}(u_i^3 u_j^3 \rho_{ij}^3) = 0$.

Now, from Lemma A.1(iv) of [22] the ρ_{ij}'s are pairwise independent and from Lemma A.1(iii) of [22] we know that

$$\mathrm{E}(\rho_{ij}^6) = \frac{15}{(k-1)(k+1)(k+3)}.$$

These results combined yield

$$\mathrm{E}(\hat{\rho}_{\boldsymbol{\theta}}^2) = \frac{4}{n^2(n-1)^2} \sum_{1\le i<j<n} \sum_{1\le i'<j'<n} \mathrm{E}(u_i^3 u_j^3 u_{i'}^3 u_{j'}^3 \rho_{ij}^3 \rho_{i'j'}^3)$$
$$= \frac{2}{n(n-1)} \mathrm{E}^2(u_1^6)\mathrm{E}(\rho_{12}^6)$$
$$= \frac{2}{n(n-1)} \left\{ \frac{15\mathrm{E}^2(u_1^6)}{(k-1)(k+1)(k+3)} \right\}$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};f}$. In view of what precedes, our goal is to show that $\hat{\rho}_{\boldsymbol{\theta}}/\sigma_n$ is asymptotically standard normal where

$$\sigma_n = \sqrt{\frac{30\mathrm{E}^2(u_1^6)}{n(n-1)(k-1)(k+1)(k+3)}}.$$

First note that

$$\frac{\hat{\rho}_{\boldsymbol{\theta}}}{\sigma_n} = \frac{2}{n(n-1)\sigma_n} \sum_{\ell=2}^{n} \sum_{i=1}^{\ell-1} (\mathbf{Y}_i^\top \mathbf{Y}_\ell)^3,$$

where the $\mathbf{Y}_1 = \boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top \mathbf{X}_1, \ldots, \mathbf{Y}_n = \boldsymbol{\Upsilon}_{\boldsymbol{\theta}}^\top \mathbf{X}_n$ are i.i.d. spherically symmetric bounded random vectors under $\mathrm{P}^{(n)}_{\boldsymbol{\theta};f}$. Write

$$\frac{\hat{\rho}_{\boldsymbol{\theta}}}{\sigma_n} = \sum_{\ell=2}^{n} D_\ell,$$

where

$$D_\ell = \frac{2}{n(n-1)\sigma_n} \sum_{i=1}^{\ell-1} (\mathbf{Y}_i^\top \mathbf{Y}_\ell)^3$$

is a martingale difference with respect to the natural filtration defined by the sequence of σ-algebras $\mathcal{F}_\ell = \sigma(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell)$. Our proof then relies on the central limit theorem for martingale differences; see, e.g., [5]. More precisely, the asymptotic normality of $\hat{\rho}_{\boldsymbol{\theta}}/\sigma_n$ follows if

$$\sum_{\ell=1}^{n} D_\ell^2 \to 1 \tag{B.2}$$

in probability as $n \to \infty$ and, for all $\epsilon > 0$,

$$\sum_{\ell=1}^{n} \mathrm{E}\{D_\ell^2 \mathbf{1}(|D_\ell| > \epsilon)\} \to 0. \tag{B.3}$$

For (B.2) first note that, from the pairwise independence, $\mathrm{E}(\rho_{i\ell}^3 \rho_{j\ell}^3) = 0$ if $i \neq j$. From this we can deduce that

$$\mathrm{E}(D_\ell^2) = \frac{4}{n^2(n-1)^2\sigma_n^2} \mathrm{E}\left(\sum_{i,j=1}^{\ell-1} u_\ell^6 u_i^3 u_j^3 \rho_{i\ell}^3 \rho_{j\ell}^3 \right)$$

$$
\begin{aligned}
&= \frac{4}{n^2(n-1)^2\sigma_n^2}(\ell-1)\mathrm{E}^2(u_\ell^6)\mathrm{E}(\rho_{1\ell}^6)\\
&= \frac{4}{n^2(n-1)^2\sigma_n^2}(\ell-1)\frac{15\mathrm{E}^2(u_1^6)}{(k-1)(k+1)(k+3)}\\
&= \frac{2}{n(n-1)}(\ell-1) \qquad \text{(B.4)}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{E}(D_\ell^4) &= \frac{16}{n^4(n-1)^4\sigma_n^4}\mathrm{E}\left(\sum_{i,j,r,s=1}^{\ell-1} u_\ell^{12}u_i^3u_j^3u_r^3u_s^3\rho_{i\ell}^3\rho_{j\ell}^3\rho_{r\ell}^3\rho_{s\ell}^3\right)\\
&= \frac{16}{n^4(n-1)^4\sigma_n^4}\sum_{i,j,r,s=1}^{\ell-1}\mathrm{E}(u_\ell^{12}u_i^3u_j^3u_r^3u_s^3)\mathrm{E}(\rho_{i\ell}^3\rho_{j\ell}^3\rho_{r\ell}^3\rho_{s\ell}^3)\\
&= \frac{16}{n^4(n-1)^4\sigma_n^4}\left\{(\ell-1)\mathrm{E}^2(u_\ell^{12})\mathrm{E}(\rho_{1\ell}^{12})+(\ell-1)(\ell-2)\mathrm{E}(u_\ell^{12})\mathrm{E}^2(u_\ell^6)\mathrm{E}^2(\rho_{1\ell}^6)\right\} \qquad \text{(B.5)}
\end{aligned}
$$

since all expectations are bounded. Note that

$$
\sum_{\ell=1}^n \mathrm{E}(D_\ell^4) = o(1) \qquad \text{(B.6)}
$$

as $n\to\infty$ and from (B.4) that

$$
\sum_{\ell=1}^n \mathrm{E}(D_\ell^2) = \frac{2}{n(n-1)}\sum_{\ell=1}^n(\ell-1) = 1. \qquad \text{(B.7)}
$$

As a consequence it remains to show that $\mathrm{var}(\sum_{\ell=1}^n D_\ell^2)$ is $o(1)$ as $n\to\infty$ to obtain (B.2). First note that, for $\ell<\ell'$ and using the same arguments as above, we get

$$
\begin{aligned}
\mathrm{E}(D_\ell^2D_{\ell'}^2) &= \frac{16}{n^4(n-1)^4\sigma_n^4}\mathrm{E}\left(\sum_{i,j=1}^{\ell-1}\sum_{r,s=1}^{\ell'-1} u_\ell^6u_i^3u_j^3\rho_{i\ell}^3\rho_{j\ell}^3u_{\ell'}^6u_r^3u_s^3\rho_{r\ell'}^3\rho_{s\ell'}^3\right)\\
&= \frac{16}{n^4(n-1)^4\sigma_n^4}\mathrm{E}\left(\sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell'-1} u_\ell^6u_{\ell'}^6u_i^6u_j^6\rho_{i\ell}^6\rho_{j\ell'}^6\right)\\
&= \frac{16}{n^4(n-1)^4\sigma_n^4}\left\{(\ell-1)\mathrm{E}^2(u_1^6)\mathrm{E}(u_1^{12})\mathrm{E}^2(\rho_{12}^6)+(\ell'-2)(\ell-1)\mathrm{E}^4(u_1^6)\mathrm{E}^2(\rho_{12}^6)\right\}\\
&= \frac{16}{n^4(n-1)^4\sigma_n^4}(\ell-1)\left\{\frac{15}{(k-1)(k+1)(k+3)}\right\}^2\left\{\mathrm{E}^2(u_1^6)\mathrm{E}(u_1^{12})+(\ell'-2)\mathrm{E}^4(u_1^6)\right\}\\
&= \frac{4}{n^2(n-1)^2}(\ell-1)\left\{\frac{\mathrm{E}(u_1^{12})}{\mathrm{E}^2(u_1^6)}+(\ell'-2)\right\} \qquad \text{(B.8)}
\end{aligned}
$$

and therefore from (B.6) and (B.8) that

$$
\begin{aligned}
\mathrm{E}\left(\sum_{\ell,\ell'=1}^n D_\ell^2D_{\ell'}^2\right) &= \sum_{\ell=1}^n\mathrm{E}(D_\ell^4)+2\sum_{1\le\ell<\ell'\le n}\mathrm{E}(D_\ell^2D_{\ell'}^2)\\
&= 2\sum_{\ell'=2}^n\sum_{\ell=1}^{\ell'-1}\mathrm{E}(D_\ell^2D_{\ell'}^2)+o(1)\\
&= \frac{8}{n^2(n-1)^2}\sum_{\ell'=2}^n\sum_{\ell=1}^{\ell'-1}(\ell-1)\left\{\frac{\mathrm{E}(u_1^{12})}{\mathrm{E}^2(u_1^6)}+(\ell'-2)\right\}+o(1)\\
&= \frac{4}{n^2(n-1)^2}\left\{\sum_{\ell'=2}^n(\ell'-2)(\ell'-1)\frac{\mathrm{E}(u_1^{12})}{\mathrm{E}^2(u_1^6)}+\sum_{\ell'=2}^n(\ell'-2)^2(\ell'-1)\right\}+o(1)
\end{aligned}
$$

$$= \frac{4}{n^2(n-1)^2}\sum_{\ell'=2}^{n}(\ell'-2)^2(\ell'-1)+o(1)$$
$$= 1+o(1) \tag{B.9}$$

as $n \to \infty$. Therefore it follows from (B.7) and (B.9) that $\mathrm{var}[\sum_{\ell=1}^{n} D_\ell^2] = o(1)$ as $n \to \infty$ so that (B.2) holds. For (B.3) the Cauchy–Schwarz and the Markov inequalities together with (B.4) and (B.5) imply that, for some constant C,

$$\sum_{\ell=1}^{n} \mathrm{E}\{D_\ell^2 \mathbf{1}(|D_\ell| > \epsilon)\} \le \sum_{\ell=1}^{n}\sqrt{\mathrm{E}(D_\ell^4)\mathrm{E}(D_\ell^2)}$$
$$\le \frac{C}{n^{5/2}(n-1)^{5/2}\sigma_n^2}\sum_{\ell=1}^{n}(\ell-1)^{3/2},$$

which is $o(1)$ as $n \to \infty$; (B.3) follows. □

References

[1] T. Abe, A. Pewsey, Sine-skewed circular distributions, Statist. Papers 52 (2011) 683–707.
[2] R.B. Arellano-Valle, A. Azzalini, The centred parametrization for the multivariate skew-normal distribution, J. Multivariate Anal. 99 (2008) 1362–1382.
[3] A. Azzalini, A class of distributions which includes the normal ones, Scand. J. Stat. 12 (1985) 171–178.
[4] A. Azzalini, A. Capitanio, Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution, J. R. Stat. Soc. Ser. B 65 (2003) 367–389.
[5] P. Hall, C.C. Heyde, Martingale Limit Theory and its Application, Academic Press, New York, 2014.
[6] M. Hallin, C. Ley, Skew-symmetric distributions and Fisher information—a tale of two densities, Bernoulli 18 (2012) 747–763.
[7] M. Hallin, C. Ley, Skew-symmetric distributions and Fisher information: the double sin of the skew-normal, Bernoulli 20 (2014) 1432–1453.
[8] S.R. Jammalamadaka, A. SenGupta, Topics in Circular Statistics, World Scientific, 2001.
[9] S. Jung, I.L. Dryden, J.S. Marron, Analysis of principal nested spheres, Biometrika 99 (2012) 551–568.
[10] P.E. Jupp, G. Regoli, A. Azzalini, A general setting for symmetric distributions and their relationship to general distributions, J. Multivariate Anal. 148 (2016) 107–119.
[11] A. Kume, A.T.A. Wood, Saddlepoint approximations for the Bingham and Fisher–Bingham normalising constants, Biometrika 92 (2005) 465–476.
[12] L. Le Cam, Asymptotic Methods in Statistical Decision Theory, Springer, New York, 1986.
[13] P. Leong, S. Carlile, Methods for spherical data analysis and visualization, J. Neurosci. Meth. 80 (1998) 191–200.
[14] C. Ley, D. Paindaveine, On the singularity of multivariate skew-symmetric models, J. Multivariate Anal. 101 (2010) 1434–1444.
[15] C. Ley, Y. Swan, B. Thiam, T. Verdebout, Optimal R-estimation of a spherical location, Statist. Sinica 23 (2013) 305–332.
[16] C. Ley, T. Verdebout, Simple optimal tests for circular reflective symmetry about a specified median direction, Statist. Sinica 14 (2014) 1319–1340.
[17] C. Ley, T. Verdebout, Local powers of optimal one- and multi-sample tests for the concentration of Fisher–von Mises–Langevin distributions, Int. Stat. Rev. 82 (2014) 440–456.
[18] K.V. Mardia, Measures of multivariate skewness and kurtosis with applications, Biometrika 57 (1970) 519–530.
[19] K.V. Mardia, Some aspects of geometry driven statistical models, in: Annual LASR 2013 Proceedings, Leeds University Press, UK, 2013, pp. 7–15.
[20] K.V. Mardia, P.E. Jupp, Directional Statistics, Wiley, New York, 2000.
[21] D. Paindaveine, T. Verdebout, Optimal rank-based tests for the location parameter of a rotationally symmetric distribution on the hypersphere, in: M. Hallin, D. Mason, D. Pfeifer, J. Steinebach (Eds.), Mathematical Statistics and Limit Theorems: Festschrift in Honor of Paul Deheuvels, Springer, New York, 2014, pp. 249–270.
[22] D. Paindaveine, T. Verdebout, On high-dimensional sign tests, Bernoulli 22 (2016) 1745–1769.
[23] A. Pewsey, Testing for circular reflective symmetry about a known median axis, J. Appl. Stat. 31 (2004) 575–585.
[24] R.J. Serfling, Multivariate symmetry and asymmetry, in: S. Kotz, N. Balakrishnan, C.B. Read, B. Vidakovic (Eds.), Encyclopedia of Statistical Sciences, Vol. 8, second ed., Wiley, 2006, pp. 5338–5345.
[25] D. Umbach, S.R. Jammalamadaka, Building asymmetry into circular distributions, Statist. Probab. Lett. 79 (2009) 659–663.
[26] A. van der Vaart, Asymptotic Statistics, Cambridge University Press, Cambridge, 1998.
[27] J. Wang, J. Boyer, M.G. Genton, A skew-symmetric representation of multivariate distributions, Statist. Sinica 14 (2004) 1259–1270.
[28] G.S. Watson, Statistics on Spheres, Wiley, New York, 1983.