Optimal Rank-Based Tests for the Location Parameter of a Rotationally Symmetric Distribution on the Hypersphere

Davy Paindaveine and Thomas Verdebout

Université libre de Bruxelles, Brussels, Belgium

Laboratoire EQUIPPE, Université Lille Nord de France, France

Abstract

Rotationally symmetric distributions on the unit hyperpshere are among the most commonly met in directional statistics. These distributions involve a finite-dimensional parameter $\boldsymbol{\theta}$ and an infinite-dimensional parameter g, that play the role of "location" and "angular density" parameters, respectively. In this paper, we focus on hypothesis testing on $\boldsymbol{\theta}$, under unspecified g. We consider (i) the problem of testing that $\boldsymbol{\theta}$ is equal to some given $\boldsymbol{\theta}_0$, and (ii) the problem of testing that $\boldsymbol{\theta}$ belongs to some given great "circle". Using the uniform local and asymptotic normality result from Ley et al. (2013), we define parametric tests that achieve Le Cam optimality at a target angular density f. To improve on the poor robustness of these parametric procedures, we then introduce a class of rank tests for these problems. Parallel to parametric tests, the proposed rank tests achieve Le Cam optimality under correctly specified angular densities. We derive the asymptotic properties of the various tests and investigate their finite-sample behaviour in a Monte Carlo study.

Keywords and phrases: Group invariance, rank-based tests, rotationally symmetric distributions, spherical statistics, uniform local and asymptotic normality

1 Introduction

Spherical or directional data naturally arise in a plethora of earth sciences such as geology (see, e.g., Fisher (1989)), seismology (Storetvedt and Scheidegger (1992)), astrophysics (Briggs (1993)), oceanography (Bowers and Mould (2000)) or meteorology (Fisher (1987)), as well as in studies of animal behavior (Fisher and Embleton (1987)) or even in neuroscience (Leong and Carlile (1998)). For decades, spherical data were explored through linear approximations trying to circumvent the "curved" nature of the data. Then the seminal paper Fisher (1953) showed that linearization hampers a correct study of several phenomena (such as, e.g., the remanent magnetism found in igneous or sedimentary rocks) and that it was therefore crucial to take into account the non-linear, spherical, nature of the data. Since then, a huge literature has been dedicated to a more appropriate study of spherical data; we refer to Mardia (1975), Jupp and Mardia (1989), Mardia and Jupp (2000) or to the second chapter of Merrifield (2006) for a detailed overview.

Spherical data are commonly viewed as realizations of a random vector $\mathbf{X}$ taking values in the unit hypersphere $\mathcal{S}^{k-1} := \left\{\mathbf{x} \in \mathbb{R}^k \,|\, \|\mathbf{x}\| := \sqrt{\mathbf{x}'\mathbf{x}} = 1\right\}$ ($k \geq 2$). In the last decades, numerous (classes of) distributions on $\mathcal{S}^{k-1}$ have been proposed and investigated. In this paper, we focus on the class of *rotationally symmetric distributions* on $\mathcal{S}^{k-1}$, that were introduced in Saw (1978). This class is characterized by the fact that the probability mass at $\mathbf{x}$ is a monotone nondecreasing function of the "spherical distance" $\mathbf{x}'\boldsymbol{\theta}$ between $\mathbf{x}$ and a given $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$. This implies that the resulting equiprobability contours are the $(k-2)$-hyperspheres $\mathbf{x}'\boldsymbol{\theta} = c$ ($c \in [-1, 1]$), and that this "north pole" $\boldsymbol{\theta}$ may be considered as a "spherical mode", hence may be interpreted as a *location* parameter.

Of course, this assumption of rotational symmetry may seem very restrictive. Yet, the latter is often used to model real phenomena. Indeed, according to Jupp and

Mardia (1989), rotationally symmetric spherical data appear *inter alia* in situations where the observation process imposes such symmetrization (e.g., the rotation of the earth; see Mardia and Edwards (1982)). Another instance where rotational symmetry is appropriate is obtained when the observation scheme does not allow to make a distinction between the measurements $\mathbf{x}$ and $\mathbf{O}_{\boldsymbol{\theta}}\mathbf{x}$ for any rotation matrix $\mathbf{O}_{\boldsymbol{\theta}}$ such that $\mathbf{O}_{\boldsymbol{\theta}}\boldsymbol{\theta} = \boldsymbol{\theta}$. In such a case, indeed, only the projection of $\mathbf{x}$ onto the modal axis $\boldsymbol{\theta}$ can be observed; see Clark (1983).

In the absolutely continuous case (with the dominating measure being the uniform distribution on $\mathcal{S}^{k-1}$), rotationally symmetric distributions have a probability density function (pdf) of the form $\mathbf{x} \mapsto cg(\mathbf{x}'\boldsymbol{\theta})$, for some nondecreasing function $g : [-1, 1] \to \mathbb{R}^+$. Hence, this model is intrinsically of a semiparametric nature. While inference about $\boldsymbol{\theta}$ has been considered in many papers (see, among others, Chang (2004) and Tsai and Sen (2007)), semiparametrically efficient inference procedures in the rotationally symmetric case have not been developed in the literature. The only exception is the very recent contribution by Ley et al. (2013), where rank-based estimators of $\boldsymbol{\theta}$ that achieve semiparametric efficiency at a target angular density are defined. Their methodology, that builds on Hallin and Werker (2003), relies on invariance arguments and on the uniform local and asymptotic normality—with respect to $\boldsymbol{\theta}$, at a fixed g—of the model considered.

Ley et al. (2013), however, considers point estimation only, hence does not address situations where one would like to test the null hypothesis that the location parameter $\boldsymbol{\theta}$ is equal to a given $\boldsymbol{\theta}_0$. In this paper, we therefore extend the results from Ley et al. (2013) to hypothesis testing. This leads to a class of rank tests for the aforementioned testing problem, that, when based on correctly specified scores, are semiparametrically optimal. The proposed tests are invariant both with respect to the group of continuous monotone increasing transformations (of spherical distances) and with respect to the group of orthogonal transformations fixing the null value $\boldsymbol{\theta}_0$. Their main advantage over "studentized" parametric tests is that they are not only validity-robust but are also efficiency-robust. We also treat a more involved testing problem, in which one needs to test the hypothesis that $\boldsymbol{\theta}$ belongs to some given great "circle"—more precisely, to the intersection of $\mathcal{S}^{k-1}$ with a given vectorial subspace of $\mathbb{R}^k$.

The outline of the paper is as follows. In Section 2, we carefully define the class of rotationally symmetric distributions considered, introduce the main assumptions needed, and state the uniform local and asymptotic normality result that will be the main technical tool for this work. In Section 3, we focus on the problem of testing that $\boldsymbol{\theta}$ is equal to some given $\boldsymbol{\theta}_0$, derive optimal parametric tests and study their asymptotic behaviour. In Section 4, we discuss the group invariance structure of this testing problem, propose a class of (invariant) rank tests, and study their asymptotic properties. In Section 5, we treat the problem of testing that $\boldsymbol{\theta}$ belongs to a given great circle. We conduct in Section 6 a Monte Carlo study to investigate the finite-sample behaviour of the proposed tests. Finally, an Appendix collects technical proofs.

2 Rotationally symmetric distributions and ULAN

The random vector $\mathbf{X}$, with values in the unit sphere $\mathcal{S}^{k-1}$ of $\mathbb{R}^k$, is said to be *rotationally symmetric* about $\boldsymbol{\theta}(\in \mathcal{S}^{k-1})$ if and only if, for all orthogonal $k \times k$ matrices $\mathbf{O}$ satisfying $\mathbf{O}\boldsymbol{\theta} = \boldsymbol{\theta}$, the random vectors $\mathbf{OX}$ and $\mathbf{X}$ are equal in distribution. If $\mathbf{X}$ is further absolutely continuous (with respect to the usual surface area measure on $\mathcal{S}^{k-1}$), then the corresponding density is of the form

$$\begin{aligned} f_{\boldsymbol{\theta},g} : \mathcal{S}^{k-1} &\to \mathbb{R}^k \\ \mathbf{x} &\mapsto c_{k,g}\, g(\mathbf{x}'\boldsymbol{\theta}), \end{aligned} \tag{2.1}$$

where $c_{k,g}(> 0)$ is a normalization constant and $g : [-1, 1] \to \mathbb{R}$ is some nonnegative function—called an *angular function* in the sequel. Throughout the paper, we then (tacitly) adopt the following assumption on the data generating process.

Assumption (A). The observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are mutually independent and admit a common density of the form (2.1), for some $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ and some angular function g in the collection $\mathcal{F}$ of functions from $[-1, 1]$ to $\mathbb{R}^+$ that are positive and monotone nondecreasing.

The notation f (instead of g) will be used when considering a fixed angular density. An angular function that plays a fundamental role in directional statistics is then

$$t \mapsto f_{\exp,\kappa}(t) = \exp(\kappa t), \tag{2.2}$$

for some "concentration" parameter $\kappa(>0)$. Clearly, $f_{\exp,\kappa}$ satisfies the conditions in Assumption (A). The resulting rotationally symmetric distribution was introduced in Fisher (1953) and is known as the Fisher-von Mises-Langevin (FvML(κ)) distribution. Other examples are the so-called "linear" rotationally symmetric distributions (LIN(a)), that are obtained for angular densities defined by $f(t) = t + a$, with $a > 1$.

In the sequel, the joint distribution of $\mathbf{X}_1, \ldots, \mathbf{X}_n$ under Assumption (A) will be denoted as $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$. Note that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, the random variables $\mathbf{X}_1'\boldsymbol{\theta}, \ldots, \mathbf{X}_n'\boldsymbol{\theta}$ are mutually independent and admit the common density (with respect to the Lebesgue measure over the real line)

$$t \mapsto \tilde{g}(t) := \frac{\omega_k \, c_{k,g}}{B(\frac{1}{2}, \frac{1}{2}(k-1))} \, g(t)(1-t^2)^{(k-3)/2} \, \mathbb{I}_{[-1,1]}(t), \tag{2.3}$$

where $B(\cdot,\cdot)$ is the beta function, $\omega_k = 2\pi^{k/2}/\Gamma(k/2)$ is the surface area of $\mathcal{S}^{k-1}$, and $\mathbb{I}_A(\cdot)$ stands for the indicator function of the set A. The corresponding cdf will be denoted by $t \mapsto \tilde{G}(t) = \int_{-1}^t \tilde{g}(t)dt$. Still under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, the random vectors $\mathbf{S}_1(\boldsymbol{\theta}), \ldots, \mathbf{S}_n(\boldsymbol{\theta})$, where we let

$$\mathbf{S}_i(\boldsymbol{\theta}) := \frac{\mathbf{X}_i - (\mathbf{X}_i'\boldsymbol{\theta})\boldsymbol{\theta}}{\|\mathbf{X}_i - (\mathbf{X}_i'\boldsymbol{\theta})\boldsymbol{\theta}\|}, \qquad i = 1, \ldots, n, \tag{2.4}$$

are independent of the $\mathbf{X}_i'\boldsymbol{\theta}$'s, and are i.i.d., with a common distribution that is uniform over the unit $(k-2)$-sphere $\mathcal{S}^{k-1}(\boldsymbol{\theta}^\perp) := \{\mathbf{x} \in \mathcal{S}^{k-1} : \mathbf{x}'\boldsymbol{\theta} = 0\}$. It is easy to check that the common mean vector and covariance matrix of the $\mathbf{S}_i(\boldsymbol{\theta})$'s are given by $\mathbf{0}$ and $(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')/(k-1)$, respectively.

Fix now an angular density f and consider the parametric family of probability measures $\mathcal{P}^{(n)}_f := \{\mathrm{P}^{(n)}_{\boldsymbol{\theta},f} \,|\, \boldsymbol{\theta} \in \mathcal{S}^{k-1}\}$. For $\mathcal{P}^{(n)}_f$ to be *uniformly locally and asymptotically normal (ULAN)*, the angular density f needs to satisfy some mild regularity conditions; more precisely, as we will state in Proposition 2.1 below, ULAN holds if f belongs to the collection $\mathcal{F}_{\rm ULAN}$ of angular densities in $\mathcal{F}$ (see Assumption (A) above) that (i) are absolutely continuous (with a.e. derivative f', say) and such that (ii), letting $\varphi_f := f'/f$, the quantity

$$\mathcal{J}_k(f) := \int_{-1}^{1} \varphi_f^2(t)(1-t^2)\tilde{f}(t)\, dt = \int_0^1 \varphi_f^2(\tilde{F}^{-1}(u))(1-(\tilde{F}^{-1}(u))^2)\, du$$

is finite.

As usual, uniform local and asymptotic normality describes the asymptotic behaviour of local likelihood ratios of the form

$$\frac{{\rm P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},f}}{{\rm P}^{(n)}_{\boldsymbol{\theta},f}},$$

where the sequence $(\boldsymbol{\tau}^{(n)})$ is bounded. In the present *curved* setup, $(\boldsymbol{\tau}^{(n)})$ should be such that $\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)} \in \mathcal{S}^{k-1}$ for all n, which imposes that $\boldsymbol{\theta}'\boldsymbol{\tau}^{(n)} = O(n^{-1/2})$. For the sake of simplicity, we will assume throughout that $\boldsymbol{\tau}^{(n)} = \boldsymbol{\tau} + O(n^{-1/2})$, with $\boldsymbol{\theta}'\boldsymbol{\tau} = 0$.

We then have the following result (see Ley et al. (2013) for the proof).

Proposition 2.1 *Fix* $f \in \mathcal{F}_{\rm ULAN}$. *Then the family* $\mathcal{P}_f^{(n)} = \left\{{\rm P}^{(n)}_{\boldsymbol{\theta},f} \mid \boldsymbol{\theta} \in \mathcal{S}^{k-1}\right\}$ *is ULAN, with* central sequence

$$\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},f} := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_f(\mathbf{X}_i'\boldsymbol{\theta})\sqrt{1-(\mathbf{X}_i'\boldsymbol{\theta})^2}\ \mathbf{S}_i(\boldsymbol{\theta}) \tag{2.5}$$

and Fisher information matrix

$$\boldsymbol{\Gamma}_{\boldsymbol{\theta},f} := \frac{\mathcal{J}_k(f)}{k-1}\,(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}'). \tag{2.6}$$

More precisely, (i) for any sequence $(\boldsymbol{\tau}^{(n)})$ *as above,*

$$\log\left(\frac{{\rm P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},f}}{{\rm P}^{(n)}_{\boldsymbol{\theta},f}}\right) = (\boldsymbol{\tau}^{(n)})'\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},f} - \frac{1}{2}(\boldsymbol{\tau}^{(n)})'\boldsymbol{\Gamma}_{\boldsymbol{\theta},f}(\boldsymbol{\tau}^{(n)}) + o_{\rm P}(1)$$

as $n \to \infty$ *under* ${\rm P}^{(n)}_{\boldsymbol{\theta},f}$, *and (ii)* $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},f}$, *still under* ${\rm P}^{(n)}_{\boldsymbol{\theta},f}$, *is asymptotically normal with mean zero and covariance matrix* $\boldsymbol{\Gamma}_{\boldsymbol{\theta},f}$.

As we show in the next section, this ULAN result allows to define Le Cam optimal tests for $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ *under specified angular density* f.

3 Optimal parametric tests for $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$

For some fixed $\boldsymbol{\theta}_0 \in \mathcal{S}^{k-1}$ and $f \in \mathcal{F}_{\rm ULAN}$, consider the problem of testing $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ in $\mathcal{P}_f^{(n)}$, that is, consider the testing problem

$$\begin{cases} \mathcal{H}_0 : \{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,f}\} \\ \mathcal{H}_1 : \bigcup_{\boldsymbol{\theta}\neq\boldsymbol{\theta}_0} \{\mathrm{P}^{(n)}_{\boldsymbol{\theta},f}\}. \end{cases} \tag{3.7}$$

For this problem, we define the test $\phi_f^{(n)}$ that, at asymptotic level α, rejects the null of (3.7) whenever

$$Q_f^{(n)} := (\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,f})' \boldsymbol{\Gamma}^-_{\boldsymbol{\theta}_0,f} \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,f} \tag{3.8}$$

$$= \frac{k-1}{n\mathcal{J}_k(f)} \sum_{i,j=1}^n \varphi_f(\mathbf{X}_i'\boldsymbol{\theta}_0)\varphi_f(\mathbf{X}_j'\boldsymbol{\theta}_0)\sqrt{1-(\mathbf{X}_i'\boldsymbol{\theta}_0)^2} \tag{3.9}$$

$$\times \sqrt{1-(\mathbf{X}_j'\boldsymbol{\theta}_0)^2}\,(\mathbf{S}_i(\boldsymbol{\theta}_0))'\mathbf{S}_j(\boldsymbol{\theta}_0) > \chi^2_{k-1,1-\alpha},$$

where $\mathbf{A}^-$ denotes the Moore-Penrose inverse of $\mathbf{A}$ and $\chi^2_{k-1,1-\alpha}$ stands for the α-upper quantile of a chi-square distribution with $k-1$ degrees of freedom. Applying in the present context the general results in Hallin et al. (2010) about hypothesis testing in curved ULAN families, yields that $\phi_f^{(n)}$ is Le Cam optimal—more precisely, locally and asymptotically maximin—at asymptotic level α for the problem (3.7). The asymptotic properties of this test are stated in the following result.

Theorem 3.1 *Fix $\boldsymbol{\theta}_0 \in \mathcal{S}^{k-1}$ and $f \in \mathcal{F}_{\rm ULAN}$. Then, (i) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,f}$, $Q_f^{(n)}$ is asymptotically chi-square with $k-1$ degrees of freedom; (ii) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)},f}$, where the sequence $(\boldsymbol{\tau}^{(n)})$ in $\mathbb{R}^k$ satisfies $\boldsymbol{\tau}^{(n)} = \boldsymbol{\tau} + O(n^{-1/2})$, with $\boldsymbol{\theta}_0'\boldsymbol{\tau} = 0$, $Q_f^{(n)}$ is asymptotically non-central chi-square, still with $k-1$ degrees of freedom, and non-centrality parameter*

$$\boldsymbol{\tau}'\,\boldsymbol{\Gamma}_{\boldsymbol{\theta}_0,f}\boldsymbol{\tau} = \frac{\mathcal{J}_k(f)}{k-1}\,\|\boldsymbol{\tau}\|^2; \tag{3.10}$$

(iii) the sequence of tests $\phi_f^{(n)}$ has asymptotic size α under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,f}$; (iv) $\phi_f^{(n)}$ is locally asymptotically maximin, at asymptotic level α, when testing $\{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,f}\}$ against alternatives of the form $\bigcup_{\boldsymbol{\theta}\neq\boldsymbol{\theta}_0}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},f}\}$.

For the particular case of the fixed-κ Fisher-von Mises-Langevin (FvML(κ)) distribution (obtained for $f_{\exp,\kappa}$; see (2.2)), we obtain

$$\begin{aligned} Q^{(n)}_{f_{\exp,\kappa}} &= \frac{\kappa^2(k-1)}{n\mathcal{J}_k(f_{\exp,\kappa})} \sum_{i,j=1}^{n} (\mathbf{X}_i - (\mathbf{X}_i'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0)'(\mathbf{X}_j - (\mathbf{X}_j'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0) \\ &= \frac{\kappa^2(k-1)}{n\mathcal{J}_k(f_{\exp,\kappa})} \sum_{i,j=1}^{n} \mathbf{X}_i'(\mathbf{I}_k - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\mathbf{X}_j \\ &=: \frac{\kappa^2(k-1)n}{\mathcal{J}_k(f_{\exp,\kappa})} \bar{\mathbf{X}}'(\mathbf{I}_k - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\bar{\mathbf{X}}. \end{aligned} \tag{3.11}$$

The main drawback of the parametric tests $\phi^{(n)}_f$ is their lack of (validity-)robustness : under angular density $g \neq f$, there is no guarantee that $\phi^{(n)}_f$ asymptotically meets the nominal level constraint. Indeed $Q^{(n)}_f$ is, in general, not asymptotically χ^2_{k-1} under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,g}$. In practice, however, the underlying angular density may hardly be assumed to be known, and it is therefore needed to define robustified versions of $\phi^{(n)}_f$ that will combine (a) Le Cam optimality at f (Theorem 3.1(iv)) and (b) validity under a broad collection of angular densities $\{g\}$.

A first way to perform such a robustification is to rely on "studentization". This simply consists in considering test statistics of the form

$$Q^{(n)}_{f;\mathrm{Stud}} := (\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,f})' \, (\hat{\boldsymbol{\Gamma}}^{g}_{\boldsymbol{\theta}_0,f})^{-} \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,f},$$

where $\hat{\boldsymbol{\Gamma}}^{g}_{\boldsymbol{\theta}_0,f}$ is an arbitrary consistent estimator of the covariance matrix in the asymptotic multinormal distribution of $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}_0,f}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0,g}$. The resulting tests, that reject the null $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ (with unspecified angular density) whenever $Q^{(n)}_{f;\mathrm{Stud}} > \chi^2_{k-1,1-\alpha}$, are *validity-robust* — that is, they asymptotically meet the level constraint under a broad range of angular densities — and remain Le Cam optimal at f.

Of special interest is the FvML studentized test—$\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$, say—that rejects the null hypothesis whenever

$$Q^{(n)}_{f_{\exp};\mathrm{Stud}} = \frac{k-1}{n\hat{\mathcal{L}}_k} \sum_{i,j=1}^{n} \mathbf{X}_i'(\mathbf{I}_k - \boldsymbol{\theta}_0\boldsymbol{\theta}_0')\mathbf{X}_j, \tag{3.12}$$

where $\hat{\mathcal{L}}_k := 1 - \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i'\boldsymbol{\theta}_0)^2$ is a consistent estimator of $\mathcal{L}_k(g) := 1 - \mathrm{E}^{(n)}_{\boldsymbol{\theta}_0,g}[(\mathbf{X}_i'\boldsymbol{\theta}_0)^2]$ (this quantity does not depend on $\boldsymbol{\theta}_0$, which justifies the notation); this test was studied in Watson (1983). From studentization, this test is valid under any rotationally symmetric distribution; moreover, since $Q^{(n)}_{f_{\exp};\mathrm{Stud}} = Q^{(n)}_{f_{\exp,\kappa}} + o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0, f_{\exp,\kappa}}$ for any κ, this test is also optimal in the Le Cam sense under any FvML distribution.

Studentization, however, typically leads to tests that fail to be *efficiency-robust*, in the sense that the resulting type 2 risk may dramatically increase when the underlying angular density g much deviates from the target density—or target densities, in the case of the FvML studentized test $\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$—at which they are optimal. That is why studentization will not be considered in this paper. Instead, we will take advantage of the group invariance structure of the testing problem considered, in order to introduce invariant tests that are both validity- and efficiency-robust. As we will see, invariant tests in the present context are *rank* tests.

4 Optimal rank tests for $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$

We start by describing the group invariance structure of the testing problem considered above (Section 4.1). Then we introduce (and study the properties of) rank-based versions of the central sequences from Proposition 2.1 (Section 4.2). This will allow us to develop the resulting (optimal) rank tests and to derive their asymptotic properties (Section 4.3).

4.1 Group invariance structure

Still for some given $\boldsymbol{\theta}_0 \in \mathcal{S}^{k-1}$, consider the problem of testing $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ *under unspecified angular density* g, that is, consider the testing problem

$$\begin{cases} \mathcal{H}_0 : \bigcup_{g \in \mathcal{F}} \{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0, g}\} \\ \mathcal{H}_1 : \bigcup_{\boldsymbol{\theta} \neq \boldsymbol{\theta}_0} \bigcup_{g \in \mathcal{F}} \{\mathrm{P}^{(n)}_{\boldsymbol{\theta}, g}\}. \end{cases} \tag{4.13}$$

This testing problem is invariant under two groups of transformations, which we now quickly describe.

(i) To define the first group, we introduce the *tangent-normal decomposition*

$$\mathbf{X}_i = (\mathbf{X}_i'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0 + \|\mathbf{X}_i - (\mathbf{X}_i'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0\|\, \mathbf{S}_i(\boldsymbol{\theta}_0), \qquad i = 1, \ldots, n$$

of the observations $\mathbf{X}_i$, $i = 1, \ldots, n$. The first group of transformations we consider is then $\mathcal{G} = \{g_h : h \in \mathcal{H}\}, \circ$, with

$$\begin{array}{rcl} g_h : \quad (\mathcal{S}^{k-1})^n & \to & (\mathcal{S}^{k-1})^n \\ (\mathbf{X}_1, \ldots, \mathbf{X}_n) & \mapsto & \big(h(\mathbf{X}_1'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0 + \|\mathbf{X}_1 - h(\mathbf{X}_1'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0\|\, \mathbf{S}_1(\boldsymbol{\theta}_0), \ldots, \\ & & h(\mathbf{X}_n'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0 + \|\mathbf{X}_n - h(\mathbf{X}_n'\boldsymbol{\theta}_0)\boldsymbol{\theta}_0\|\, \mathbf{S}_n(\boldsymbol{\theta}_0)\big), \end{array}$$

where $\mathcal{H}$ is the collection of mappings $h : [-1, 1] \mapsto [-1, 1]$ that are continuous, monotone increasing, and satisfy $h(\pm 1) = \pm 1$.

The null hypothesis of (4.13) is clearly invariant under the group $\mathcal{G}, \circ$. The invariance principle therefore suggests restricting to tests that are invariant with respect to this group. As it was shown in Ley et al. (2013), the maximal invariant $\mathbf{I}^{(n)}(\boldsymbol{\theta}_0)$ associated with $\mathcal{G}, \circ$ is the sign-and-rank statistic $(\mathbf{S}_1(\boldsymbol{\theta}_0), \ldots, \mathbf{S}_n(\boldsymbol{\theta}_0), R_1(\boldsymbol{\theta}_0), \ldots, R_n(\boldsymbol{\theta}_0))$, where $R_i(\boldsymbol{\theta}_0)$ denotes the rank of $\mathbf{X}_i'\boldsymbol{\theta}_0$ among $\mathbf{X}_1'\boldsymbol{\theta}_0, \ldots, \mathbf{X}_n'\boldsymbol{\theta}_0$. Consequently, the class of invariant tests coincides with the collection of tests that are measurable with respect to $\mathbf{I}^{(n)}(\boldsymbol{\theta}_0)$, in short, with the class of *(sign-and-)rank*—or, simply, *rank*—tests.

It is easy to check that $\mathcal{G}, \circ$ is actually a generating group for the null hypothesis $\bigcup_{g \in \mathcal{F}} \{\mathrm{P}^{(n)}_{\boldsymbol{\theta}_0, g}\}$ in (4.13). As a direct corollary, rank tests are distribution-free under the whole null hypothesis. This explains why rank tests will be validity-robust.

(ii) Of course, the null hypothesis in (4.13) is also invariant under orthogonal transformations fixing the null location value $\boldsymbol{\theta}_0$. More precisely, it is invariant under the group $\mathcal{G}_{\mathrm{rot}} = \{g_{\mathbf{O}} : \mathbf{O} \in \mathcal{O}_{\boldsymbol{\theta}_0}\}, \circ$, with

$$\begin{array}{rcl} g_{\mathbf{O}} : \quad (\mathcal{S}^{k-1})^n & \to & (\mathcal{S}^{k-1})^n \\ (\mathbf{X}_1, \ldots, \mathbf{X}_n) & \mapsto & (\mathbf{O}\mathbf{X}_1, \ldots, \mathbf{O}\mathbf{X}_n), \end{array}$$

where $\mathcal{O}_{\boldsymbol{\theta}_0}$ is the collection of all $k \times k$ matrices $\mathbf{O}$ satisfying $\mathbf{O}\boldsymbol{\theta}_0 = \boldsymbol{\theta}_0$. Clearly, the vectors of signs and ranks above is not invariant under $\mathcal{G}_{\mathrm{rot}}, \circ$, but the statistic

$$\Big(\big(\mathbf{S}_1(\boldsymbol{\theta}_0)\big)'\mathbf{S}_2(\boldsymbol{\theta}_0), \big(\mathbf{S}_1(\boldsymbol{\theta}_0)\big)'\mathbf{S}_3(\boldsymbol{\theta}_0), \ldots, \big(\mathbf{S}_{n-1}(\boldsymbol{\theta}_0)\big)'\mathbf{S}_n(\boldsymbol{\theta}_0), R_1(\boldsymbol{\theta}_0), \ldots, R_n(\boldsymbol{\theta}_0) \Big) \tag{4.14}$$

is. Tests that are measurable with respect to the statistic in (4.14) will therefore be invariant with respect to both groups considered above.

4.2 Rank-based central sequences

To combine validity-robustness/invariance with Le Cam optimality at a target angular density f, we introduce rank-based versions of the central sequences that appear in the ULAN property above (Proposition 2.1). More precisely, we consider rank statistics of the form

$$\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta},K} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} K\left(\frac{R_i(\boldsymbol{\theta})}{n+1}\right)\mathbf{S}_i(\boldsymbol{\theta}),$$

where the *score function* $K : [0,1] \to \mathbb{R}$ is throughout assumed to be continuous (which implies that it is bounded and square-integrable over $[0,1]$).

In order to state the asymptotic properties of the rank-based random vector $\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta},K}$, we introduce the following notation. For any $g \in \mathcal{F}$, write

$$\mathbf{\Delta}^{(n)}_{\boldsymbol{\theta},K,g} := \frac{1}{\sqrt{n}}\sum_{i=1}^{n} K\big(\tilde{G}(\mathbf{X}_i'\boldsymbol{\theta})\big)\mathbf{S}_i(\boldsymbol{\theta}),$$

where $\tilde{G}$ denotes the cdf of $\mathbf{X}_i'\boldsymbol{\theta}$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$. For any $g \in \mathcal{F}_{\mathrm{ULAN}}$, define further

$$\mathbf{\Gamma}_{\boldsymbol{\theta},K} := \frac{\mathcal{J}_k(K)}{k-1}\left(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}'\right) \quad \text{and} \quad \mathbf{\Gamma}_{\boldsymbol{\theta},K,g} := \frac{\mathcal{J}_k(K,g)}{k-1}\left(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}'\right),$$

with $\mathcal{J}_k(K) := \int_0^1 K^2(u)du$ and $\mathcal{J}_k(K,g) := \int_0^1 K(u)K_g(u)du$, where we wrote $K_g(u) := \varphi_g(\tilde{G}^{-1}(u))\sqrt{1-(\tilde{G}^{-1}(u))^2}$ for any $u \in [0,1]$. We then have the following result (see Ley et al. (2013)).

Proposition 4.1 *Fix $\boldsymbol{\theta} \in \mathcal{S}^{k-1}$ and let $(\boldsymbol{\tau}^{(n)})$ be a sequence in $\mathbb{R}^k$ that satisfies $\boldsymbol{\tau}^{(n)} = \boldsymbol{\tau} + O(n^{-1/2})$, with $\boldsymbol{\theta}'\boldsymbol{\tau} = 0$. Then (i) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, with $g \in \mathcal{F}$, $\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta},K} = \mathbf{\Delta}^{(n)}_{\boldsymbol{\theta},K,g} + o_{L^2}(1)$ as $n \to \infty$; (ii) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, with $g \in \mathcal{F}$, $\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta},K}$ is asymptotically multinormal with mean zero and covariance matrix $\mathbf{\Gamma}_{\boldsymbol{\theta},K}$; (iii) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$, with $g \in \mathcal{F}_{\mathrm{ULAN}}$, $\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta},K}$ is asymptotically multinormal with mean $\mathbf{\Gamma}_{\boldsymbol{\theta},K,g}\boldsymbol{\tau}$ and covariance matrix $\mathbf{\Gamma}_{\boldsymbol{\theta},K}$; (iv) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, with $g \in \mathcal{F}_{\mathrm{ULAN}}$, $\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},K} = \underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\theta},K} - \mathbf{\Gamma}_{\boldsymbol{\theta},K,g}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1)$ as $n \to \infty$.*

For any $f \in \mathcal{F}^{C^1}_{\mathrm{ULAN}}$, where $\mathcal{F}^{C^1}_{\mathrm{ULAN}}$ denotes the collection of angular densities in $\mathcal{F}_{\mathrm{ULAN}}$ that are continuously differentiable over $[-1,1]$, the function $u \mapsto K_f(u) := \varphi_f(\tilde{F}^{-1}(u))\sqrt{1-(\tilde{F}^{-1}(u))^2}$ is a valid score function to be used in rank-based central sequences. Proposition 4.1(i) then entails that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},f}$, $\underset{\sim}{\mathbf{\Delta}}^{(n)}_{\boldsymbol{\vartheta},K_f}$ is asymptotically equivalent—in L^2, hence also in probability—to the parametric f-central sequence $\mathbf{\Delta}^{(n)}_{\boldsymbol{\theta},f} = \mathbf{\Delta}^{(n)}_{\boldsymbol{\theta},K_f,f}$. This provides the key to develop rank tests that are Le Cam optimal at any given $f \in \mathcal{F}^{C^1}_{\rm ULAN}$. As for Proposition 4.1(ii)-(iii), they allow to derive the asymptotic properties of the resulting optimal rank tests.

4.3 Rank tests for $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$

Fix a score function K as above. The previous sections then make it natural to consider the rank test—$\underset{\sim}{\phi}{}^{(n)}_K$, say—that, at asymptotic level α, rejects the null hypothesis $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ (with unspecified angular density g) whenever

$$\begin{aligned} \underset{\sim}{Q}{}^{(n)}_K &= (\underset{\sim}{\mathbf{\Delta}}{}^{(n)}_{\boldsymbol{\theta}_0,K})' \mathbf{\Gamma}^-_{\boldsymbol{\theta}_0,K} \underset{\sim}{\mathbf{\Delta}}{}^{(n)}_{\boldsymbol{\theta}_0,K} \\ &= \frac{k-1}{n\mathcal{J}_k(K)} \sum_{i,j=1}^n K\left(\frac{R_i(\boldsymbol{\theta}_0)}{n+1}\right) K\left(\frac{R_j(\boldsymbol{\theta}_0)}{n+1}\right) (\mathbf{S}_i(\boldsymbol{\theta}_0))'\mathbf{S}_j(\boldsymbol{\theta}_0) \\ &> \chi^2_{k-1,1-\alpha}. \end{aligned}$$

Clearly, this test is invariant with respect to both groups introduced in Section 4.1, since it is measurable with respect to the statistic in (4.14). The following result, that summarizes the asymptotic properties of $\underset{\sim}{\phi}{}^{(n)}_K$, easily follows from Proposition 4.1.

Theorem 1 *Let $(\boldsymbol{\tau}^{(n)})$ be a sequence in $\mathbb{R}^k$ that satisfies $\boldsymbol{\tau}^{(n)} = \boldsymbol{\tau} + O(n^{-1/2})$, with $\boldsymbol{\theta}'_0\boldsymbol{\tau} = 0$. Then, (i) under $\bigcup_{g\in\mathcal{F}}\{{\rm P}^{(n)}_{\boldsymbol{\theta}_0,g}\}$, $\underset{\sim}{Q}{}^{(n)}_K$ is asymptotically chi-square with $k-1$ degrees of freedom; (ii) under ${\rm P}^{(n)}_{\boldsymbol{\theta}_0+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$, with $g \in \mathcal{F}_{\rm ULAN}$, is asymptotically non-central chi-square, still with $k-1$ degrees of freedom, and non-centrality parameter*

$$\boldsymbol{\tau}'\mathbf{\Gamma}_{\boldsymbol{\theta},K,g}\mathbf{\Gamma}^-_{\boldsymbol{\theta}_0,K}\mathbf{\Gamma}_{\boldsymbol{\theta},K,g}\boldsymbol{\tau} = \frac{\mathcal{J}^2_k(K,g)}{(k-1)\mathcal{J}_k(K)}\,\|\boldsymbol{\tau}\|^2; \tag{4.15}$$

(iii) the sequence of tests $\underset{\sim}{\phi}{}^{(n)}_K$ has asymptotic size α under $\bigcup_{g\in\mathcal{F}}\{{\rm P}^{(n)}_{\boldsymbol{\theta}_0,g}\}$; (iv) $\underset{\sim}{\phi}{}^{(n)}_{K_f}$, with $f \in \mathcal{F}^{C^1}_{\rm ULAN}$, is locally and asymptotically maximin, at asymptotic level α, when testing $\bigcup_{g\in\mathcal{F}}\{{\rm P}^{(n)}_{\boldsymbol{\theta}_0,g}\}$ against alternatives of the form $\bigcup_{\boldsymbol{\theta}\neq\boldsymbol{\theta}_0}\{{\rm P}^{(n)}_{\boldsymbol{\theta},f}\}$.

Note that, for $K = K_f$ and $g = f$ (with $f \in \mathcal{F}^{C^1}_{\rm ULAN}$), the non-centrality parameter in (4.15) above rewrites

$$\frac{\mathcal{J}^2_k(K_f,f)}{(k-1)\mathcal{J}_k(K_f)}\,\|\boldsymbol{\tau}\|^2 = \frac{\mathcal{J}_k(f)}{k-1}\,\|\boldsymbol{\tau}\|^2,$$

hence coincides with the non-centrality parameter in (3.10). This establishes the optimality statement in Theorem 1(iv).

5 Testing great circle hypotheses

In this section, we turn to another classical testing problem involving rotationally symmetric distributions, namely to the problem of testing that $\boldsymbol{\theta}$ belongs to some given "great circle" of $\mathcal{S}^{k-1}$, where the term *great circle* here refers to the intersection of $\mathcal{S}^{k-1}$ with a vectorial subspace of $\mathbb{R}^k$. In other words, letting $\boldsymbol{\Upsilon}$ be some given full-rank $k \times s$ ($s < k$) matrix, we consider the problem of testing $\mathcal{H}_0^{\boldsymbol{\Upsilon}} : \boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$ against $\mathcal{H}_1^{\boldsymbol{\Upsilon}} : \boldsymbol{\theta} \notin \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$, where $\mathcal{M}(\boldsymbol{\Upsilon})$ denotes the s-dimensional subspace of $\mathbb{R}^k$ that is spanned by the columns of $\boldsymbol{\Upsilon}$. More precisely, the testing problem is

$$\begin{cases} \mathcal{H}_0^{\boldsymbol{\Upsilon}} : \bigcup_{\boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})} \bigcup_g \left\{ \mathrm{P}_{\boldsymbol{\theta},g}^{(n)} \right\} \\ \mathcal{H}_1^{\boldsymbol{\Upsilon}} : \bigcup_{\boldsymbol{\theta} \notin \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})} \bigcup_g \left\{ \mathrm{P}_{\boldsymbol{\theta},g}^{(n)} \right\}. \end{cases} \tag{5.16}$$

This problem has been studied by Watson (1983), that provided a FvML score test, and in Fujikoshi and Watamori (1992) and Watamori (1992), that investigated the properties of the FvML likelihood ratio test.

For any $f \in \mathcal{F}_{\mathrm{ULAN}}$, the construction of f-optimal parametric tests for this problem proceeds as follows. Fix $\boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$ (the unspecification of $\boldsymbol{\theta}$ under the null will be taken care of later on) and consider a local perturbation of the form $\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}^{(n)}$ where $\boldsymbol{\tau}^{(n)}$ is such that $\boldsymbol{\tau}^{(n)} = \boldsymbol{\tau} + O(n^{-1/2})$, with $\boldsymbol{\theta}'\boldsymbol{\tau} = 0$ (see the comment just before Proposition 2.1). It directly follows from Hallin et al. (2010) that a locally (in the vicinity of $\boldsymbol{\theta}$) and asymptotically most stringent test can be obtained by considering the most stringent test for the linear constraint $\boldsymbol{\tau}^{(n)} \in \mathcal{M}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}') \cap \mathcal{M}(\boldsymbol{\Upsilon})$. Letting $\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}$ be such that $\mathcal{M}(\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}) = \mathcal{M}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}') \cap \mathcal{M}(\boldsymbol{\Upsilon})$, the resulting f-optimal test therefore rejects the null hypothesis $\mathcal{H}_{0,f}^{\boldsymbol{\Upsilon}} : \bigcup_{\boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})} \left\{ \mathrm{P}_{\boldsymbol{\theta},f}^{(n)} \right\}$ for large values of

$$Q_{\boldsymbol{\theta},f}^{(n)} = (\boldsymbol{\Delta}_{\boldsymbol{\theta},f}^{(n)})' \left(\boldsymbol{\Gamma}_{\boldsymbol{\theta},f}^{-} - \boldsymbol{\Upsilon}_{\boldsymbol{\theta}} (\boldsymbol{\Upsilon}_{\boldsymbol{\theta}}' \boldsymbol{\Gamma}_{\boldsymbol{\theta},f} \boldsymbol{\Upsilon}_{\boldsymbol{\theta}})^{-} \boldsymbol{\Upsilon}_{\boldsymbol{\theta}}' \right) \boldsymbol{\Delta}_{\boldsymbol{\theta},f}^{(n)}.$$

Using the identity $(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\boldsymbol{\Upsilon}_{\boldsymbol{\theta}} = \boldsymbol{\Upsilon}_{\boldsymbol{\theta}}$ (which follows from the fact that $\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}'$ is the projection matrix onto $\mathcal{M}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')$, that, by definition, contains every column vector of $\mathbf{\Upsilon}_{\boldsymbol{\theta}}$), then the fact that $\mathbf{\Upsilon}_{\boldsymbol{\theta}}(\mathbf{\Upsilon}_{\boldsymbol{\theta}}'\mathbf{\Upsilon}_{\boldsymbol{\theta}})^{-}\mathbf{\Upsilon}_{\boldsymbol{\theta}}'(\mathbf{I}_k-\boldsymbol{\theta}\boldsymbol{\theta}') = \mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}'(\mathbf{I}_k-\boldsymbol{\theta}\boldsymbol{\theta}')$, we obtain

$$\begin{aligned} Q_{\boldsymbol{\theta},f}^{(n)} &= \frac{k-1}{\mathcal{J}_k(f)}\,(\mathbf{\Delta}_{\boldsymbol{\theta},f}^{(n)})'\left(\mathbf{I}_k-\mathbf{\Upsilon}_{\boldsymbol{\theta}}(\mathbf{\Upsilon}_{\boldsymbol{\theta}}'\mathbf{\Upsilon}_{\boldsymbol{\theta}})^{-}\mathbf{\Upsilon}_{\boldsymbol{\theta}}'\right)\mathbf{\Delta}_{\boldsymbol{\theta},f}^{(n)} \\ &= \frac{k-1}{\mathcal{J}_k(f)}\,(\mathbf{\Delta}_{\boldsymbol{\theta},f}^{(n)})'(\mathbf{I}_k-\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-}\mathbf{\Upsilon}')\mathbf{\Delta}_{\boldsymbol{\theta},f}^{(n)}. \end{aligned} \tag{5.17}$$

We will show below that, under $\mathrm{P}_{\boldsymbol{\theta},f}^{(n)}$, $Q_{\boldsymbol{\theta},f}^{(n)}$ is asymptotically chi-square with $k-s$ degrees of freedom, so that the resulting test rejects the null, at asymptotic level α, whenever $Q_{\boldsymbol{\theta},f}^{(n)}$ exceeds the corresponding upper α-quantile $\chi^2_{k-s,1-\alpha}$.

Since $\boldsymbol{\theta}\in\mathcal{M}(\mathbf{\Upsilon})$ implies that $\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}'\boldsymbol{\theta}=\boldsymbol{\theta}$ (or equivalently, that $(\mathbf{I}_k-\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}')\boldsymbol{\theta}=\mathbf{0}$), the FvML($\kappa$) version of $Q_{\boldsymbol{\theta},f}^{(n)}$ is given by

$$Q_{\boldsymbol{\theta},f_{\exp,\kappa}}^{(n)} = \frac{n\kappa^2(k-1)}{\mathcal{J}_k(f_{\exp,\kappa})}\,\bar{\mathbf{X}}'(\mathbf{I}_k-\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}')\bar{\mathbf{X}}. \tag{5.18}$$

As in Section 3, this leads to defining the FvML studentized test — $\phi_{f_{\exp};\mathrm{Stud}}^{(n)}$, say — that rejects $\mathcal{H}_0^{\mathbf{\Upsilon}}$ whenever

$$Q_{f_{\exp};\mathrm{Stud}}^{(n)} = \frac{n(k-1)}{\hat{\mathcal{L}}_k(\hat{\boldsymbol{\theta}})}\,\bar{\mathbf{X}}'(\mathbf{I}_k-\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}')\bar{\mathbf{X}} > \chi^2_{k-s,1-\alpha},$$

where $\hat{\mathcal{L}}_k(\boldsymbol{\theta}) = 1-\frac{1}{n}\sum_{i=1}^n(\mathbf{X}_i'\boldsymbol{\theta})^2$ is evaluated at an arbitrary consistent estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$. When based on $\hat{\boldsymbol{\theta}}=\bar{\mathbf{X}}$, this test is actually the Watson (1983) score test. Consequently, the following result, that states the asymptotic and optimality properties of $\phi_{f_{\exp};\mathrm{Stud}}^{(n)}$, clarifies the exact optimality properties of this classical test.

Theorem 5.1 *Fix $\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\mathbf{\Upsilon})$ and let $(\boldsymbol{\tau}^{(n)})$ be a sequence in $\mathbb{R}^k$ that satisfies $\boldsymbol{\tau}^{(n)}=\boldsymbol{\tau}+O(n^{-1/2})$, with $\boldsymbol{\theta}'\boldsymbol{\tau}=0$. Then, (i) under $\mathrm{P}_{\boldsymbol{\theta},g}^{(n)}$, with $g\in\mathcal{F}$, $Q_{f_{\exp};\mathrm{Stud}}^{(n)}$ is asymptotically chi-square with $k-s$ degrees of freedom; (ii) under $\mathrm{P}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}^{(n)}$, with $g\in\mathcal{F}_{\mathrm{ULAN}}$, $Q_{f_{\exp};\mathrm{Stud}}^{(n)}$ is asymptotically non-central chi-square, still with $k-s$ degrees of freedom, and non-centrality parameter*

$$\frac{k-1}{\mathcal{L}_k(g)}\,\boldsymbol{\tau}'(\mathbf{I}_k-\mathbf{\Upsilon}(\mathbf{\Upsilon}'\mathbf{\Upsilon})^{-1}\mathbf{\Upsilon}')\boldsymbol{\tau}; \tag{5.19}$$

(iii) the sequence of tests $\phi_{f_{\exp};\mathrm{Stud}}^{(n)}=\mathbb{I}[Q_{f_{\exp};\mathrm{Stud}}^{(n)}>\chi^2_{k-s,1-\alpha}]$ has asymptotic size α under $\cup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\mathbf{\Upsilon})}\cup_{g\in\mathcal{F}}\{\mathrm{P}_{\boldsymbol{\theta},g}^{(n)}\}$; (iv) $\phi_{f_{\exp};\mathrm{Stud}}^{(n)}$ is locally asymptotically most stringent,

at asymptotic level α, when testing $\cup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\boldsymbol{\Upsilon})}\cup_{g\in\mathcal{F}}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}\}$ against alternatives of the form $\cup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\setminus\mathcal{M}(\boldsymbol{\Upsilon})}\cup_{\kappa>0}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},f_{\exp,\kappa}}\}$.

This test is therefore valid under any rotationally symmetric distribution, hence is validity-robust. It is optimal under any FvML distribution, but is not efficiency-robust outside the class of FvML distributions. As for the first testing problem we considered in the previous sections, this motivates building rank-based tests that combine validity- and efficiency-robustness.

The appropriate rank test statistics are obtained by replacing in (5.17) the parametric central sequence $\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},f}$ with its rank-based counterpart $\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta},f}=\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta},K_f}$. More generally, we will consider general-score rank statistics of the form

$$\underset{\sim}{Q}^{(n)}_{\boldsymbol{\theta},K}=\frac{k-1}{\mathcal{J}_k(K)}\left(\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta},K}\right)'\left(\mathbf{I}_k-\boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right)\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta},K}.$$

As already mentioned, $\boldsymbol{\theta}$ is not specified under the null hypothesis. We will therefore rather consider the test — $\underset{\sim}{\phi}^{(n)}_{K}$, say — that rejects the null at asymptotic level α whenever $\underset{\sim}{Q}^{(n)}_{K}:=\underset{\sim}{Q}^{(n)}_{\hat{\boldsymbol{\theta}},K}>\chi^2_{k-s,1-\alpha}$. The estimator $\hat{\boldsymbol{\theta}}$ to be used needs to be part of a sequence of estimators that satisfies the following assumption.

Assumption (B). The sequence of estimators $\hat{\boldsymbol{\theta}}=\hat{\boldsymbol{\theta}}^{(n)}$ is (i) *root-n consistent:* $\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}=O_{\mathrm{P}}(n^{-1/2})$ under $\bigcup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\boldsymbol{\Upsilon})}\bigcup_{g\in\mathcal{F}}\mathrm{P}^{(n)}_{\boldsymbol{\theta};g}$; (ii) *locally and asymptotically discrete:* for all $\boldsymbol{\theta}$ and for all $C>0$, there exists a positive integer $M=M(C)$ such that the number of possible values of $\hat{\boldsymbol{\theta}}^{(n)}$ in balls of the form $\{\boldsymbol{\theta}'\in\mathbb{R}^k:\sqrt{n}\|\boldsymbol{\theta}'-\boldsymbol{\theta}\|\leq C\}$ is bounded by M, uniformly as $n\to\infty$; (iii) *constrained:* $\hat{\boldsymbol{\theta}}$ takes its values in $\mathcal{M}(\boldsymbol{\Upsilon})(\cap\mathcal{S}^{k-1})$.

The following result then states the asymptotic properties of the rank tests $\underset{\sim}{\phi}^{(n)}_{K}$.

Theorem 5.2 *Let Assumption (B) hold, fix $g\in\mathcal{F}_{\mathrm{ULAN}}$, $\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\boldsymbol{\Upsilon})$, and let $(\boldsymbol{\tau}^{(n)})$ be a sequence in $\mathbb{R}^k$ that satisfies $\boldsymbol{\tau}^{(n)}=\boldsymbol{\tau}+O(n^{-1/2})$, with $\boldsymbol{\theta}'\boldsymbol{\tau}=0$. Then, (i) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, $\underset{\sim}{Q}^{(n)}_{K}$ is asymptotically chi-square with $k-s$ degrees of freedom; (ii) under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$, $\underset{\sim}{Q}^{(n)}_{K}$ is asymptotically non-central chi-square, still with $k-s$*

degrees of freedom, and non-centrality parameter

$$\frac{\mathcal{J}_k^2(K,g)}{(k-1)\mathcal{J}_k(K)}\boldsymbol{\tau}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\boldsymbol{\tau}; \tag{5.20}$$

(iii) the sequence of tests $\underset{\sim}{\phi}_K^{(n)} = \mathbb{I}[\underset{\sim}{Q}_K^{(n)} > \chi^2_{k-s,1-\alpha}]$ *has asymptotic size* α *under* $\cup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\boldsymbol{\Upsilon})}\cup_{g\in\mathcal{F}}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}\}$*; (iv)* $\underset{\sim}{\phi}_{K_f}^{(n)}$*, with* $f \in \mathcal{F}^{C^1}_{\mathrm{ULAN}}$*, is locally asymptotically most stringent, at asymptotic level* α*, when testing* $\cup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\cap\mathcal{M}(\boldsymbol{\Upsilon})}\cup_{g\in\mathcal{F}_{\mathrm{ULAN}}}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}\}$ *against alternatives of the form* $\cup_{\boldsymbol{\theta}\in\mathcal{S}^{k-1}\setminus\mathcal{M}(\boldsymbol{\Upsilon})}\{\mathrm{P}^{(n)}_{\boldsymbol{\theta},f}\}$.

Theorems 5.1- 5.2 allow to compute in a straightforward way (as usual, as the ratios of the non-centrality parameters in the asymptotic non-null distributions of the corresponding statistics) the asymptotic relative efficiencies (AREs) of the proposed rank tests with respect to their FvML-score competitors from Watson (1983). These AREs are given by

$$\mathrm{ARE}_g\big[\,\underset{\sim}{\phi}_K^{(n)}/\phi^{(n)}_{f_{\exp};\mathrm{Stud}}\,\big] = \frac{\mathcal{L}_k^2(g)\mathcal{J}_k^2(K,g)}{(k-1)^2\mathcal{J}_k(K)},$$

and do not depend on $\boldsymbol{\theta}$ nor on $\boldsymbol{\tau}$. It is easy to check that these AREs, that coincides with the ones obtained in Ley et al. (2013) for point estimation, also hold for the testing problem considered in the previous sections.

6 Simulations

In this final section, we conduct a Monte Carlo study to investigate the finite-sample behaviour of the rank tests proposed in Sections 4.3 and 5. Letting

$$\boldsymbol{\theta}_0 = \begin{pmatrix}1\\0\\0\end{pmatrix} \in \mathbb{R}^k = \mathbb{R}^3 \quad \text{and} \quad \boldsymbol{\Upsilon} = \begin{pmatrix}1 & 0\\0 & 0\\0 & 1\end{pmatrix},$$

we considered the problems of testing $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $\mathcal{H}_0 : \boldsymbol{\theta} \in \mathcal{S}^{k-1}\cap\mathcal{M}(\boldsymbol{\Upsilon})$, respectively. We generated $M = 10{,}000$ (mutually independent) random samples of rotationally symmetric random vectors

$$\boldsymbol{\varepsilon}_i^{(\ell)}, \quad i = 1,\ldots,n = 250, \quad \ell = 1,2,3,4,$$

with location $\boldsymbol{\theta}_0$ and with angular densities that are FvML(1), FvML(3), LIN(1.1), and LIN(2), for $\ell = 1, 2, 3$, and 4, respectively; see Section 2. Each random vector $\boldsymbol{\varepsilon}_i^{(\ell)}$ was then transformed into

$$\mathbf{X}_{i;\omega}^{(\ell)} := \mathbf{O}_\omega \boldsymbol{\varepsilon}_i^{(\ell)}, \quad i = 1, \ldots, n, \quad \ell = 1, 2, 3, 4, \quad \omega = 0, 1, 2, 3,$$

with

$$\mathbf{O}_\omega = \begin{pmatrix} \cos(\pi\omega/50) & -\sin(\pi\omega/50) & 0 \\ \sin(\pi\omega/50) & \cos(\pi\omega/50) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For both testing problems considered, the value $\omega = 0$ corresponds to the null hypothesis, whereas $\omega = 1, 2, 3$ correspond to increasingly severe alternatives.

On each replication of the samples $(\mathbf{X}_{1;\omega}^{(\ell)}, \ldots, \mathbf{X}_{n;\omega}^{(\ell)})$, $\ell = 1, 2, 3, 4$, $\omega = 0, 1, 2, 3$, we performed the following tests for $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ and for $\mathcal{H}_0 : \boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$, all at asymptotic level $\alpha = 5\%$: (i) the tests $\phi_{f_{\exp,3}}^{(n)}$, that is, the parametric tests $\phi_f^{(n)}$ using a FvML(3) angular target density, (ii) the tests $\phi_{f_{\exp};\mathrm{Stud}}^{(n)}$ that are based on the FvML studentized statistics $Q_{f_{\exp};\mathrm{Stud}}^{(n)}$, and (iii) the rank-based tests $\underset{\sim}{\phi}_{K_{\mathrm{FvML}(1)}}^{(n)}$, $\underset{\sim}{\phi}_{K_{\mathrm{FvML}(3)}}^{(n)}$, $\underset{\sim}{\phi}_{K_{\mathrm{LIN}(2)}}^{(n)}$, and $\underset{\sim}{\phi}_{K_{\mathrm{LIN}(2)}}^{(n)}$ that are Le Cam optimal at FvML(1), FvML(3), Lin(1.1), and Lin(2) distributions, respectively. The resulting rejection frequencies are provided in Tables 1 and 2, for $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ and for $\mathcal{H}_0 : \boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$, respectively.

These empirical results perfectly agree with the asymptotic theory : the parametric tests $\phi_{f_{\exp,3}}^{(n)}$ are the most powerful ones at the FvML(3) distribution, but their null size deviates quite much from the target size $\alpha = 5\%$ away from the FvML(3). The studentized parametric tests $\phi_{f_{\exp};\mathrm{Stud}}^{(n)}$, on the contrary, show a null behaviour that is satisfactory under all distributions considered. They also dominate their rank-based competitors under FvML densities, which is in line with the fact that they are optimal in the class of FvML densities. Outside this class, however, the proposed rank tests are more powerful than the studentized tests, which translates their better efficiency-robustness. The null behaviour of the proposed rank tests is very satisfactory, and their optimality under correctly specified angular densities is confirmed.

		ω			
Underlying density	Test	0	1	2	3
FvML(1)	$\phi^{(n)}_{f_{\exp,3}}$	.1162	.1572	.2673	.4493
	$\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$	.0486	.0733	.1502	.2867
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(1)}}$	.0494	.0732	.1516	.2891
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(3)}}$	.0525	.0720	.1467	.2743
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(1.1)}}$	.0504	.0662	.1237	.2248
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(2)}}$	.0527	.0738	.1505	.2890
FvML(3)	$\phi^{(n)}_{f_{\exp,3}}$	.0528	.2250	.7104	.9719
	$\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$	.0536	.2239	.7062	.9701
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(1)}}$	.0530	.2177	.6828	.9615
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(3)}}$	.0541	.2244	.7056	.9695
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(1.1)}}$	.0509	.2066	.6688	.9598
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(2)}}$	.0538	.2220	.7025	.9674
LIN(1.1)	$\phi^{(n)}_{f_{\exp,3}}$	.1290	.1661	.2729	.4556
	$\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$	.0498	.0668	.1372	.2749
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(1)}}$	.0493	.0676	.1396	.2747
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(3)}}$	.0496	.0723	.1588	.3245
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(1.1)}}$	.0496	.0711	.1608	.3359
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(2)}}$	.0501	.0698	.1499	.3001
LIN(2)	$\phi^{(n)}_{f_{\exp,3}}$	.1384	.1496	.1727	.2284
	$\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$	.0526	.0579	.0741	.1074
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(1)}}$	.0545	.0585	.0769	.1100
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(3)}}$	.0542	.0600	.0773	.1069
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(1.1)}}$	.0540	.0600	.0737	.0968
	$\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(2)}}$	.0540	.0610	.0781	.1110

Table 1: Rejection frequencies (out of $M = 10,000$ replications), under the null $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ ($\omega = 0$) and increasingly severe alternatives ($\omega = 1, 2, 3$), of the parametric FvML(3)-test ($\phi^{(n)}_{f_{\exp,3}}$), the studentized FvML-test ($\phi^{(n)}_{f_{\exp};\mathrm{Stud}}$), and of the rank tests achieving optimality at FvML(1), FvML(3), LIN(1.1), and LIN(2) densities ($\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(1)}}$, $\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{FvML}(3)}}$, $\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(1.1)}}$, and $\underset{\sim}{\phi}^{(n)}_{K_{\mathrm{LIN}(2)}}$, respectively). The sample size is $n = 250$ and the nominal level is 5%; see Section 6 for further details.

		ω			
Underlying density	Test	0	1	2	3
FvML(1)	$\phi^{(n)}_{f_{\exp,3}}$	.1027	.1391	.2902	.4998
	$\phi^{(n)}_{f_{\exp};\text{Stud}}$	.0543	.0794	.1922	.3783
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(1)}}$	.0541	.1943	.1516	.3786
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(3)}}$	.0537	.0789	.1835	.3568
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(1.1)}}$	.0524	.0712	.1560	.2997
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(2)}}$	.0532	.0792	.1914	.3695
FvML(3)	$\phi^{(n)}_{f_{\exp,3}}$	.0464	.2888	.7955	.9889
	$\phi^{(n)}_{f_{\exp};\text{Stud}}$	.0460	.2884	.7937	.9891
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(1)}}$	.0462	.2725	.7686	.9839
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(3)}}$	.0467	.2873	.7930	.9897
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(1.1)}}$	.0486	.2632	.7635	.9849
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(2)}}$	.0454	.2831	.7862	.9883
LIN(1.1)	$\phi^{(n)}_{f_{\exp,3}}$	.1080	.1598	.2922	.4763
	$\phi^{(n)}_{f_{\exp};\text{Stud}}$	.0474	.0798	.1808	.3363
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(1)}}$	.0490	.0793	.1801	.3345
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(3)}}$	.0480	.0876	.2114	.3920
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(1.1)}}$	.0476	.0896	.2173	.4075
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(2)}}$	.0498	.0841	.1968	.3653
LIN(2)	$\phi^{(n)}_{f_{\exp,3}}$	.1119	.1220	.1659	.2328
	$\phi^{(n)}_{f_{\exp};\text{Stud}}$	.0488	.0558	.0875	.1366
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(1)}}$	.0502	.0567	.0901	.1391
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(3)}}$	.0523	.0613	.0895	.1386
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(1.1)}}$	.0521	.0594	.0838	.1224
	$\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(2)}}$	.0516	.0599	.0915	.1422

Table 2: Rejection frequencies (out of $M = 10,000$ replications), under the null $\mathcal{H}_0 : \boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$ ($\omega = 0$) and increasingly severe alternatives ($\omega = 1, 2, 3$), of the parametric FvML(3)-test ($\phi^{(n)}_{f_{\exp,3}}$), the studentized FvML-test ($\phi^{(n)}_{f_{\exp};\text{Stud}}$), and of the rank tests achieving optimality at FvML(1), FvML(3), LIN(1.1), and LIN(2) densities ($\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(1)}}$, $\underset{\sim}{\phi}{}^{(n)}_{K_{\text{FvML}(3)}}$, $\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(1.1)}}$, and $\underset{\sim}{\phi}{}^{(n)}_{K_{\text{LIN}(2)}}$, respectively). The sample size is $n = 250$ and the nominal level is 5%; see Section 6 for further details.

Acknowledgments

Davy Paindaveine's research is supported by an A.R.C. contract from the Communauté Française de Belgique and by the IAP research network grant nr. P7/06 of the Belgian government (Belgian Science Policy). The authors are grateful to two anonymous referees and the Editor for their careful reading and insightful comments that allowed to improve the original manuscript.

A Appendix

In this Appendix, we prove Theorem 5.1 and Theorem 5.2.

Proof of Theorem 5.1. (i) Recalling that $\hat{\boldsymbol{\theta}}$ is an arbitrary consistent estimator of $\boldsymbol{\theta}$, we have that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$,

$$\hat{\mathcal{L}}_k(\hat{\boldsymbol{\theta}}) - \hat{\mathcal{L}}_k(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^n \big\{(\mathbf{X}_i'\boldsymbol{\theta})^2 - (\mathbf{X}_i'\hat{\boldsymbol{\theta}})^2\big\}$$

$$= \frac{1}{n}\sum_{i=1}^n \big\{\mathbf{X}_i'(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\mathbf{X}_i'(\boldsymbol{\theta}+\hat{\boldsymbol{\theta}})\big\} = (\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\Big\{\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i'\Big\}(\boldsymbol{\theta}+\hat{\boldsymbol{\theta}}) = o_{\mathrm{P}}(1)$$

as $n \to \infty$, so that $\hat{\mathcal{L}}_k(\hat{\boldsymbol{\theta}})$ is a consistent estimator of $\mathcal{L}_k(g) = 1 - \mathrm{E}^{(n)}_{\boldsymbol{\theta},g}[(\mathbf{X}_i'\boldsymbol{\theta})^2]$. Consequently, $Q^{(n)}_{f_{\exp};\mathrm{Stud}} - Q^{(n)}_{\boldsymbol{\theta},g}$ is $o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, with

$$Q^{(n)}_{\boldsymbol{\theta},g} := \frac{n(k-1)}{\mathcal{L}_k(g)}\,\bar{\mathbf{X}}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\bar{\mathbf{X}}$$

$$= \frac{n(k-1)}{\mathcal{L}_k(g)}\,\bar{\mathbf{X}}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\bar{\mathbf{X}} = \mathbf{Y}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\mathbf{Y},$$

where we used the fact that $\big(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\big)(\mathbf{I} - \boldsymbol{\theta}\boldsymbol{\theta}') = \mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'$ and where we let $\mathbf{Y} := \sqrt{n(k-1)/\mathcal{L}_k(g)}(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')\bar{\mathbf{X}}$. Under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$,

$$\sqrt{n}(\mathbf{I} - \boldsymbol{\theta}\boldsymbol{\theta}')\bar{\mathbf{X}} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \big\{\sqrt{1-(\mathbf{X}_i'\boldsymbol{\theta})^2}\,\mathbf{S}_i(\boldsymbol{\theta})\big\} \tag{A.21}$$

is asymptotically normal with mean zero and covariance matrix $\mathcal{L}_k(g)(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')/(k-1)$, so that $\mathbf{Y}$, under the same, is asymptotically normal with mean zero and covariance matrix $\mathbf{I}_k - \boldsymbol{\theta\theta}'$. By using again the fact that $\left(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right)(\mathbf{I} - \boldsymbol{\theta\theta}') = \mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'$ and by noting that $\mathrm{tr}[\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'] = k - s$, it is easy to check that Theorem 9.2.1 in Rao and Mitra (1971) provides the result.

(ii) Le Cam's third lemma implies that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$, the random vector in (A.21) is asymptotically normal with mean

$$\lim_{n\to\infty} \mathrm{Cov}_{\boldsymbol{\theta},g}\left[\sqrt{n}(\mathbf{I} - \boldsymbol{\theta\theta}')\bar{\mathbf{X}}, \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},g}\right]\boldsymbol{\tau}^{(n)} \tag{A.22}$$

and covariance matrix $\mathcal{L}_k(g)(\mathbf{I}_k - \boldsymbol{\theta\theta}')/(k-1)$. Using (2.3) and integrating by parts yields

$$\mathrm{E}_{\boldsymbol{\theta},g}\left[(1 - (\mathbf{X}_1'\boldsymbol{\theta})^2)\varphi_g(\mathbf{X}_1'\boldsymbol{\theta})\right] = \int_{-1}^{1}(1-t^2)\varphi_g(t)\tilde{g}(t)\,dt = k-1,$$

so that (A.22) can be rewritten as $\mathrm{E}_{\boldsymbol{\theta},g}\left[(1 - (\mathbf{X}_1'\boldsymbol{\theta})^2)\varphi_g(\mathbf{X}_1'\boldsymbol{\theta})\right]\mathrm{E}_{\boldsymbol{\theta},g}\left[S_1(\boldsymbol{\theta})(S_1(\boldsymbol{\theta}))'\right]\boldsymbol{\tau} = (\mathbf{I}_k - \boldsymbol{\theta\theta}')\boldsymbol{\tau} = \boldsymbol{\tau}$. Therefore, $\mathbf{Y}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$, is asymptotically normal with mean $\boldsymbol{\mu} := \sqrt{(k-1)/\mathcal{L}_k(g)}\,\boldsymbol{\tau}$ and covariance matrix $\mathbf{I}_k - \boldsymbol{\theta\theta}'$. From contiguity, we still have that $Q^{(n)}_{f_{\exp};\mathrm{Stud}} - \mathbf{Y}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\mathbf{Y}$ is $o_{\mathrm{P}}(1)$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$. Theorem 9.2.1 in Rao and Mitra (1971) then shows that, under this sequence of probability measures, $Q^{(n)}_{f_{\exp};\mathrm{Stud}}$ is asymptotically χ^2_{k-s} with non-centrality parameter $\boldsymbol{\mu}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\boldsymbol{\mu}$, which establishes the result.

(iii) This directly follows from the asymptotic null distribution given in (i) and the classical Helly-Bray theorem.

(iv) Fix $\kappa > 0$. Then, it follows from Part (i) of the proof that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},f_{\exp,\kappa}}$, with $\boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$, $Q^{(n)}_{f_{\exp};\mathrm{Stud}}$ is asymptotically equivalent in probability to

$$Q^{(n)}_{\boldsymbol{\theta};f_{\exp,\kappa}} = \frac{n(k-1)}{\mathcal{L}_k(f_{\exp,\kappa})}\,\bar{\mathbf{X}}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\bar{\mathbf{X}} = \frac{n\kappa^2(k-1)}{\mathcal{J}_k(f_{\exp,\kappa})}\,\bar{\mathbf{X}}'(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\bar{\mathbf{X}},$$

which is the FvML(κ)-most stringent statistic we derived in (5.18). □

In Theorem 5.1(ii), we assumed that $g \in \mathcal{F}_{\mathrm{ULAN}}$ to show, through Le Cam's third lemma, that $\sqrt{n}(\mathbf{I} - \boldsymbol{\theta\theta}')\bar{\mathbf{X}}$, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$, is asymptotically normal with mean $\boldsymbol{\tau}$ and covariance matrix $\mathcal{L}_k(g)(\mathbf{I}_k - \boldsymbol{\theta\theta}')/(k-1)$. Actually, the result still holds for $g \in \mathcal{F}$, as it can be shown that, as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},g}$,

$$\sqrt{n}(\mathbf{I} - \boldsymbol{\theta\theta}')\bar{\mathbf{X}} = \mathbf{M}^{(n)} + (\mathbf{I}_k + \boldsymbol{\theta\theta}')\boldsymbol{\tau} + o_{\mathrm{P}}(1) = \mathbf{M}^{(n)} + \boldsymbol{\tau} + o_{\mathrm{P}}(1),$$

where $\mathbf{M}^{(n)} := \sqrt{n}(\mathbf{I} - (\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}^{(n)})(\boldsymbol{\theta} + n^{-1/2}\boldsymbol{\tau}^{(n)})')\bar{\mathbf{X}}$, under the same, is clearly asymptotically normal with mean zero and covariance matrix $\mathcal{L}_k(g)(\mathbf{I}_k - \boldsymbol{\theta}\boldsymbol{\theta}')/(k-1)$.

Proof of Theorem 5.2. (i)-(ii) First note that, since $\boldsymbol{\theta}'\boldsymbol{\tau}^{(n)} = O(n^{-1/2})$, Proposition 4.1(iv) rewrites

$$\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta}+n^{-1/2}\boldsymbol{\tau}^{(n)},K} = \underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta},K} - \frac{\mathcal{J}(K,g)}{k-1}\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1) \tag{A.23}$$

as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$. Since Assumption (B) holds, Lemma 4.4 in Kreiss (1987) allows to replace in (A.23) the deterministic quantity $\boldsymbol{\tau}^{(n)}$ with the random one $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$, which yields

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

as $n \to \infty$, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$. This, jointly with Assumption (B)(iii) (which implies that $(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')\hat{\boldsymbol{\theta}} = 0$ almost surely), entails that, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, with $\boldsymbol{\theta} \in \mathcal{M}(\boldsymbol{\Upsilon})$,

$$\left(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right)\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\hat{\boldsymbol{\theta}},K} = \left(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right)\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{\boldsymbol{\theta},K} + o_{\mathrm{P}}(1),$$

as $n \to \infty$. It follows that $\underset{\sim}{Q}^{(n)}_{K} = \underset{\sim}{Q}^{(n)}_{\boldsymbol{\theta},K} + o_{\mathrm{P}}(1)$ as $n \to \infty$ under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},g}$, with $\boldsymbol{\theta} \in \mathcal{M}(\boldsymbol{\Upsilon})$, hence also under sequences of local alternatives. The results in (i)-(ii) then follow, as in the proof of Theorem 5.1(i)-(ii), from Theorem 9.2.1 in Rao and Mitra (1971) and Proposition 4.1(ii)-(iii) (recall that $(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}')(\mathbf{I} - \boldsymbol{\theta}\boldsymbol{\theta}') = \mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'$).

(iii) As in the proof of Theorem 5.1(iii), this is a direct consequence of Part (i) of the result and the classical Helly-Bray theorem.

(iv) Then, under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},f}$, with $\boldsymbol{\theta} \in \mathcal{S}^{k-1} \cap \mathcal{M}(\boldsymbol{\Upsilon})$, $\underset{\sim}{Q}^{(n)}_{K_f} = \underset{\sim}{Q}^{(n)}_{\boldsymbol{\theta},K_f} + o_{\mathrm{P}}(1)$ as $n \to \infty$. Now, Proposition 4.1(i) entails that, under the same sequence of hypotheses, $\underset{\sim}{Q}^{(n)}_{\boldsymbol{\theta},K_f}$ is asymptotically equivalent in probability to

$$Q^{(n)}_{\boldsymbol{\theta},K_f} = \frac{k-1}{\mathcal{J}_k(K_f)}(\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},K_f})'\left(\mathbf{I}_k - \boldsymbol{\Upsilon}(\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon})^{-1}\boldsymbol{\Upsilon}'\right)\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta},K_f},$$

which coincides with the f-most stringent statistic in (5.17). The result follows. □

References

Bowers, J. A., M. I. D. and Mould, G. I. (2000), "Directional statistics of the wind and waves," *Applied Ocean Research*, 22, 401–412.

Briggs, M. S. (1993), "Dipole and quadrupole tests of the isotropy of gamma-ray burst locations," *Astrophys. J.*, 407, 126–134.

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

Clark, R. M. (1983), "Estimation of parameters in the marginal Fisher distribution," *Austr. J. Statist.*, 25, 227–237.

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

Fisher, N. I. (1953), "Dispersion on a sphere," *Proceedings of the Royal Society of London A*, 217, 295–305.

— (1987), "Problems with the current definitions of the standard deviation of wind direction," *J. Clim. Appl. Meteorol.*, 26, 1522–1529.

— (1989), "Smoothing a sample of circular data," *J. Struct. Geol.*, 11, 775–778.

Fujikoshi, Y. and Watamori, Y. (1992), "Tests for the mean direction of the Langevin distribution with large concentration parameter," *J. Multivariate Anal.*, 42, 210–225.

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

Hallin, M. and Werker, B. J. (2003), "Semi-parametric efficiency, distribution-freeness and invariance," *Bernoulli*, 9, 137–165.

Jupp, P. and Mardia, K. (1989), "A unified view of the theory of directional statistics," *Internat. Statist. Rev.*, 57, 261–294.

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

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

Ley, C., Swan, Y., Thiam, B., and Verdebout, T. (2013), "Optimal R-estimation of a spherical location," *Statist. Sinica*, 23, 305–333.

Mardia, K. (1975), "Statistics of directional data (with discussion)," *J. Roy. Statist. Soc. Ser. B*, 37, 349–393.

Mardia, K. V. and Edwards, R. (1982), "Weighted distributions and rotating caps," *Biometrika*, 69, 323–330.

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

Merrifield, A. (2006), "An investigation of mathematical models for animal group movement, using classical and statistical approaches," Ph.D. thesis, The University of Sydney, Sydney, Australia.

Rao, C. and Mitra, S. (1971), *Generalized Inverses of Matrices and its Apllications*, New York: J. Wiley.

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

Storetvedt, K. M. and Scheidegger, A. E. (1992), "Orthogonal joint structures in the Bergen area, southwest Norway, and their regional significance," *Phys. Earth Planet. In.*, 73, 255–263.

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

Watamori, Y. (1992), "Tests for a given linear structure of the mean direction of the Langevin distribution," *Ann. Inst. Statist. Math.*, 44.

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