# On the power of axial tests of uniformity on spheres

**Christine Cutting[1], Davy Paindaveine[*1,2] and Thomas Verdebout[1]**

[1]*Université libre de Bruxelles*
*ECARES and Département de Mathématique*
*Avenue F.D. Roosevelt, 50*
*ECARES, CP114/04*
*B-1050, Brussels*
*Belgium*

[2]*Université Toulouse Capitole*
*Toulouse School of Economics*
*1, Esplanade de l'Université*
*31080 Toulouse Cedex 06*
*France*

**Abstract:** Testing uniformity on the $p$-dimensional unit sphere is arguably the most fundamental problem in directional statistics. In this paper, we consider this problem in the framework of *axial* data, that is, under the assumption that the $n$ observations at hand are randomly drawn from a distribution that charges antipodal regions equally. More precisely, we focus on axial, rotationally symmetric, alternatives and first address the problem under which the direction $\boldsymbol{\theta}$ of the corresponding symmetry axis is specified. In this setup, we obtain Le Cam optimal tests of uniformity, that are based on the sample covariance matrix (unlike their non-axial analogs, that are based on the sample average). For the more important unspecified-$\boldsymbol{\theta}$ problem, some classical tests are available in the literature, but virtually nothing is known on their non-null behavior. We therefore study the non-null behavior of the celebrated Bingham test and of other tests that exploit the single-spiked nature of the considered alternatives. We perform Monte Carlo exercises to investigate the finite-sample behavior of our tests and to show their agreement with our asymptotic results.

## 1. Introduction

Directional statistics are concerned with data taking values on the unit hypersphere $\mathcal{S}^{p-1} := \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\|^2 := \mathbf{x}'\mathbf{x} = 1\}$ of $\mathbb{R}^p$. Classical applications, that most often relate to the circular case ($p = 2$) or spherical one ($p = 3$), involve wind and animal migration data or belong to fields such as geology, paleomagnetism, or cosmology. We refer, e.g., to Fisher, Lewis and Embleton

---

*Corresponding author.

(1987), Mardia and Jupp (2000), and Ley and Verdebout (2017) for book-length treatments of the topic and for further applications.

Arguably the most fundamental problem in directional statistics is the problem of testing for uniformity, which, for a random sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ at hand, consists in testing the null hypothesis that the observations are sampled from the uniform distribution over $\mathcal{S}^{p-1}$. This is a very classical problem in multivariate analysis that can be traced back to Bernoulli (1735). As explained in the review paper García-Portugués and Verdebout (2018), the topic has recently received a lot of attention: to cite only a few contributions, Jupp (2008) proposed data-driven Sobolev tests, Cuesta-Albertos, Cuevas and Fraiman (2009) and García-Portugués, Navarro-Esteban and Cuesta-Albertos (2020) proposed tests based on random projections, Lacour and Pham Ngoc (2014) studied the problem for noisy data, Paindaveine and Verdebout (2016) obtained the high-dimensional limiting behavior of some classical test statistics under the null hypothesis while García-Portugués, Paindaveine and Verdebout (2020) transformed some uniformity tests into tests of rotational symmetry.

A classical uniformity test dates back to Rayleigh (1919) and rejects the null hypothesis for large values of $\|\bar{\mathbf{X}}_n\|$, where $\bar{\mathbf{X}}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i$ is the sample mean of the observations. This test will detect only alternatives whose mean vectors are non-zero, hence is therefore typically used when possible deviations from uniformity are suspected to be asymmetric. In particular, the Rayleigh test will show no power when the common distribution of the $\mathbf{X}_i$'s is an antipodally symmetric distribution on the sphere, that is, when this distribution attributes the same probability to antipodal regions. Far from being the exception, such antipodally symmetric distributions are actually those that need be considered when practitioners are facing *axial data*, that is, when one does not observe genuine locations on the sphere but rather axes (a typical example of axial data relates to the directions of optical axes in quartz crystals; see, e.g., Mardia and Jupp, 2000). Models and inference for axial data have been considered a lot in the literature: to cite a few, Tyler (1987) and more recently Paindaveine, Remy and Verdebout (2020) considered inference on the distribution of the spatial sign of a Gaussian vector, Watson (1965), Bijral, Breitenbach and Grudic (2007) and Sra and Karp (2013) considered inference for Watson distributions (see the next section), Dryden (2005) obtained distributions on high-dimensional spheres while Anderson and Stephens (1972), Bingham (1974) and Jupp (2001) considered uniformity tests against antipodally symmetric alternatives.

Now, while the literature offers both axial and non-axial tests of uniformity, axial procedures unfortunately remain much less well understood than their non-axial counterparts, particularly so when it comes to their non-null behaviors. Strong results have been obtained for non-axial tests of uniformity regarding their asymptotic power under suitable local alternatives to uniformity and even regarding their optimality (we refer to Cutting, Paindaveine and Verdebout, 2017 and to the references therein), but virtually nothing is known in that direction for axial tests of uniformity. This provides the main motivation for the present work, that intends to fill an important gap by studying the non-null behavior of some classical (and less classical) axial tests of uniformity. Quite natu-

rally, we will do so in the semiparametric distributional framework that has been classically considered for non-axial tests of uniformity, namely the framework of *rotationally symmetric distributions* indexed by a finite-dimensional parameter $(\kappa, \boldsymbol{\theta}) \in \mathbb{R}^+ \times \mathcal{S}^{p-1}$ and an infinite-dimensional parameter $f \in \mathcal{F}$ (a family of functions we define in the next section). Within this semiparametric model, the null hypothesis of uniformity takes the form $\mathcal{H}_0 : \kappa = 0$. In this paper, we first derive the shape of tests that are locally and asymptotically optimal under any $f \in \mathcal{F}$ for the specified-$\boldsymbol{\theta}$ problem. We then focus on the unspecified-$\boldsymbol{\theta}$ problem and discuss its connection with the specified-$\boldsymbol{\theta}$ problem. We derive the limiting behavior, under sequences of contiguous alternatives (for any $f \in \mathcal{F}$), of the tests provided in Anderson and Stephens (1972) and Bingham (1974). Doing so, we obtain in particular the limiting behavior, under local alternatives, of the extreme eigenvalues of the spatial sign covariance matrix, which is a result of independent interest; see Dürre, Tyler and Vogel (2016) and the references therein for a recent study of these eigenvalues.

The outline and contribution of the paper are as follows. In Section 2, we introduce the class of (rotationally symmetric) alternatives to uniformity that will be considered in this work, which deviate from uniformity along a direction $\boldsymbol{\theta}$ and have a severity controlled by a concentration parameter $\kappa$. In this framework, we identify the sequences $(\kappa_n)$ that make the corresponding sequences of alternatives contiguous to the null hypothesis of uniformity. In Section 3, we tackle the problem of testing uniformity under specified $\boldsymbol{\theta}$ and show that the resulting model is *locally asymptotically normal* (LAN). We define the resulting optimal tests of uniformity and determine their asymptotic powers under contiguous alternatives. In Section 4, we turn to the unspecified-$\boldsymbol{\theta}$ problem, we show that our LAN result naturally leads to the Bingham (1974) test of uniformity, and we study the limiting behavior of this test under contiguous alternatives. In Section 5, we turn our attention to tests that take into account the "single-spiked" structure of the considered alternatives. We characterize the asymptotic behavior of these tests both under the null hypothesis and under sequences of contiguous alternatives. While all results above are confirmed by suitable numerical exercises in Sections 3–5, we specifically conduct, in Section 6, Monte-Carlo simulations in order to compare the finite-sample powers of the various tests. We provide final comments and perspectives for future research in Section 7. Finally, an appendix contains all proofs.

## 2. Axial rotationally symmetric distributions

In this section, we set the notation and describe the class of axial distributions we will use as alternatives to uniformity. A celebrated class of axial distributions on the sphere is the one that collects the Watson distributions, which admit a density (throughout, densities over $\mathcal{S}^{p-1}$ are with respect to the surface area measure) of the form

$$\mathbf{x} \mapsto \frac{c_{p,\kappa} \Gamma(\frac{p-1}{2})}{2\pi^{(p-1)/2}} \exp(\kappa\,(\mathbf{x}'\boldsymbol{\theta})^2), \qquad \mathbf{x} \in \mathcal{S}^{p-1}, \tag{1}$$

where $\boldsymbol{\theta}$ belongs to $\mathcal{S}^{p-1}$, $\kappa$ is a real number, and where the value of the normalizing constant $c_{p,\kappa}$ can be obtained from (3) below; as usual, $\Gamma$ denotes Euler's Gamma function. Since this density is a symmetric function of $\mathbf{x}$, it attributes the same probability to antipodal regions on the sphere, hence is indeed suitable for axial data. The Watson distribution is rotationally symmetric about $\boldsymbol{\theta}$, in the sense that if $\mathbf{X}$ has density (1), then $\mathbf{OX}$ and $\mathbf{X}$ share the same distribution for any $p \times p$ orthogonal matrix $\mathbf{O}$ such that $\mathbf{O}\boldsymbol{\theta} = \boldsymbol{\theta}$; consequently, $\boldsymbol{\theta}$ will be considered a *location* parameter. The Watson distributions are the rotationally symmetric (or single-spiked) Bingham (1974) distributions. The parameter $\kappa$ is a *concentration* parameter: the larger $|\kappa|$, the more the probability mass will be concentrated—symmetrically about the poles $\pm\boldsymbol{\theta}$ for positive values of $\kappa$ (bipolar case) or symmetrically about the hyperspherical equator $\mathcal{S}_{\boldsymbol{\theta}}^{\perp} := \{\mathbf{x} \in \mathcal{S}^{p-1} : \mathbf{x}'\boldsymbol{\theta} = 0\}$ for negative values of $\kappa$ (girdle case). Of course, the value $\kappa = 0$ corresponds to the uniform distribution over the sphere.

In this paper, we consider a natural semiparametric extension of the class of Watson distributions, namely the class of axial distributions admitting a density of the form

$$\mathbf{x} \mapsto \frac{c_{p,\kappa,f}\Gamma(\frac{p-1}{2})}{2\pi^{(p-1)/2}} f(\kappa \, (\mathbf{x}'\boldsymbol{\theta})^2), \qquad \mathbf{x} \in \mathcal{S}^{p-1}, \tag{2}$$

where $\boldsymbol{\theta}$ and $\kappa$ are as in (1), $f$ belongs to the class of functions $\mathcal{F} := \{f : \mathbb{R} \to \mathbb{R}^{+} : f$ monotone increasing, twice differentiable at 0, with $f(0) = f'(0) = 1\}$, and where

$$c_{p,\kappa,f} = 1 \left/ \int_{-1}^{1} (1 - s^2)^{(p-3)/2} f(\kappa s^2) \, ds. \right. \tag{3}$$

The parameter $\kappa$ is still a concentration parameter, that shares the same interpretation as for Watson distributions. The restriction to $\mathcal{F}$ above is made for identifiability purposes. If $\kappa \neq 0$, then $f$ and the pair $\{\pm\boldsymbol{\theta}\}$ are identifiable, but $\boldsymbol{\theta}$ itself is not (which is natural for axial distributions). For $\kappa = 0$, the uniform distribution over the sphere is obtained (irrespective of $f$), in which case the location parameter $\boldsymbol{\theta}$ is unidentifiable, even up to a sign. The corresponding normalizing constant is $c_p := \lim_{\kappa \to 0} c_{p,\kappa,f}$. The distribution associated with (2) is rotationally symmetric about $\boldsymbol{\theta}$, and if $\mathbf{X}$ is a random vector with this distribution, then $\mathbf{X}'\boldsymbol{\theta}$ has density $s \mapsto c_{p,\kappa,f}(1 - s^2)^{(p-3)/2} f(\kappa s^2)\mathbb{I}[|s| \leq 1]$, which explains the expression (3). In the present axial case, this density is of course symmetric with respect to zero.

The semiparametric class of distributions just introduced will be used in the paper as alternatives to uniformity on the sphere. We will consider the following hypotheses and asymptotic scenarios. For $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$, a sequence $(\kappa_n)$ in $\mathbb{R}_0$ and $f \in \mathcal{F}$, we will denote as $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$ the hypothesis under which $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ form a random sample from the density $\mathbf{x} \mapsto c_{p,\kappa_n,f} f(\kappa_n \, (\mathbf{x}'\boldsymbol{\theta})^2)$ over $\mathcal{S}^{p-1}$. This triangular array framework will allow us to consider local alternatives, associated with suitable sequences $(\kappa_n)$ converging to zero. The null hypothesis of uniformity will be denoted as $\mathrm{P}_0^{(n)}$. The sequence of hypotheses $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$ determines a sequence of alternatives to uniformity: the larger $|\kappa_n|$, the more severe the corresponding alternative, whereas the sign of $\kappa_n$ determines the type of alternatives

considered, i.e., *bipolar* (for $\kappa_n > 0$) or *girdle-like* (for $\kappa_n < 0$). At places, it will be of interest to compare our results with those obtained in the non-axial case, that is, in the case where $\mathbf{X}_{n1}, \ldots, \mathbf{X}_{nn}$ have a common density proportional to $f(\kappa_n \mathbf{x}'\boldsymbol{\theta})$ still with $f$ monotone increasing (rather than $f(\kappa_n(\mathbf{x}'\boldsymbol{\theta})^2)$).

Our first result identifies the sequences of alternatives $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$ that are contiguous to the sequence of null hypotheses $\mathrm{P}_0^{(n)}$.

**Theorem 2.1.** *Fix $p \in \{2, 3, \ldots\}$, $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$, and $f \in \mathcal{F}$. Let $(\kappa_n)$ be a sequence in $\mathbb{R}_0$ that is $O(1/\sqrt{n})$. Then, the sequence of alternative hypotheses $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$ and the sequence of null hypotheses $\mathrm{P}_0^{(n)}$ are mutually contiguous.*

In other words, if $\kappa_n = O(1/\sqrt{n})$, then no test for $\mathcal{H}_{0n} : \{\mathrm{P}_0^{(n)}\}$ against $\mathcal{H}_{1n} : \{\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}\}$ can be consistent. Actually, it will follow from Theorem 3.1 in the next section that $1/\sqrt{n}$ is the *contiguity rate*, in the sense that if $\kappa_n = \tau/\sqrt{n}$ (for some non-zero real constant $\tau$), then there exist tests for $\mathcal{H}_{0n} : \{\mathrm{P}_0^{(n)}\}$ against $\mathcal{H}_{1n} : \{\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}\}$ showing non-trivial asymptotic powers (that is, asymptotic powers in $(\alpha, 1)$, where $\alpha$ denotes the nominal level). The contiguity rate in the axial case thus coincides with the one obtained in the non-axial case; see Theorems 2.1 and 3.1 in Cutting, Paindaveine and Verdebout (2017).

## 3. Tests of uniformity under specified location

In this section, we consider the problem of testing uniformity over $\mathcal{S}^{p-1}$ against the class of alternatives introduced in the previous section, in a situation where the location $\boldsymbol{\theta}$ is specified. In other words, this corresponds to cases where it is known in which direction the possible deviation from uniformity would materialize. Depending on the exact type of alternatives we want to focus on (bipolar, girdle-type, or both), we will then consider, for a fixed $\boldsymbol{\theta}$, the problem of testing $\mathcal{H}_0^{(n)} : \{\mathrm{P}_0^{(n)}\}$ against (i) $\mathcal{H}_1^{(n)} : \cup_{\kappa>0} \cup_{f \in \mathcal{F}} \{\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$, against (ii) $\mathcal{H}_1^{(n)} : \cup_{\kappa<0} \cup_{f \in \mathcal{F}} \{\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$, or against (iii) $\mathcal{H}_1^{(n)} : \cup_{\kappa\neq0} \cup_{f \in \mathcal{F}} \{\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$. Optimal testing may be based on the following Local Asymptotic Normality (LAN) result.

**Theorem 3.1.** *Fix $p \in \{2, 3, \ldots\}$, $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$, and $f \in \mathcal{F}$. Let $\kappa_n = \tau_n p/\sqrt{n}$, where the real sequence $(\tau_n)$ is $O(1)$ but not $o(1)$. Then, letting*

$$\Delta_{\boldsymbol{\theta}}^{(n)} := \frac{p}{\sqrt{n}} \sum_{i=1}^{n} \left\{ (\mathbf{X}_{ni}'\boldsymbol{\theta})^2 - \frac{1}{p} \right\} \quad and \quad \Gamma_p := \frac{2(p-1)}{p+2},$$

*we have that, as $n \to \infty$ under $\mathrm{P}_0^{(n)}$,*

$$\Lambda_n = \log \frac{d\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}}{d\mathrm{P}_0^{(n)}} = \tau_n \Delta_{\boldsymbol{\theta}}^{(n)} - \frac{\tau_n^2}{2}\Gamma_p + o_{\mathrm{P}}(1), \tag{4}$$

*where $\Delta_{\boldsymbol{\theta}}^{(n)}$ is asymptotically normal with mean zero and variance $\Gamma_p$. In other words, the sequence $(\{\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)} : \kappa \in \mathbb{R}\})$ (where we let $\mathrm{P}_{\boldsymbol{\theta},0,f}^{(n)} := \mathrm{P}_0^{(n)}$) is locally*

*asymptotically normal at* $\kappa = 0$ *with central sequence* $\Delta_{\boldsymbol{\theta}}^{(n)}$, *Fisher information* $\Gamma_p$, *and contiguity rate* $1/\sqrt{n}$.

This result confirms that $1/\sqrt{n}$ is the contiguity rate when testing uniformity against the considered axial alternatives. Note also that the central sequence rewrites

$$\Delta_{\boldsymbol{\theta}}^{(n)} = \sqrt{n}\big(p\,\boldsymbol{\theta}'\mathbf{S}_n\boldsymbol{\theta} - 1\big),$$

where $\mathbf{S}_n := n^{-1}\sum_{i=1}^{n}\mathbf{X}_{ni}\mathbf{X}_{ni}'$ is the sample covariance matrix of the observations (with respect to a fixed location, namely the origin of $\mathbb{R}^p$). Consequently, optimal testing of uniformity for axial data will be based on $\mathbf{S}_n$. This is to be compared with the non-axial case considered in Cutting, Paindaveine and Verdebout (2017), where optimal testing of uniformity is rather based on $\bar{\mathbf{X}}_n = n^{-1}\sum_{i=1}^{n}\mathbf{X}_{ni}$. This will have important consequences when considering the unspecified-$\boldsymbol{\theta}$ case we turn to in Sections 4–5.

More importantly, the optimal axial tests of uniformity in the specified location case directly result from the LAN property above. More precisely, Theorem 3.1 entails that, for the problem of testing $\mathcal{H}_0^{(n)} : \{\mathrm{P}_0^{(n)}\}$ against $\cup_{\kappa>0}\cup_{f\in\mathcal{F}}$ $\{\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$, the test $\phi_{\boldsymbol{\theta}+}^{(n)}$ rejecting the null hypothesis at asymptotic level $\alpha$ whenever

$$T_{\boldsymbol{\theta}}^{(n)} := \frac{\Delta_{\boldsymbol{\theta}}^{(n)}}{\sqrt{\Gamma_p}} > z_\alpha \tag{5}$$

is locally asymptotically most powerful; here, $z_\alpha = \Phi^{-1}(1-\alpha)$ denotes the upper $\alpha$-quantile of the standard normal distribution. A routine application of Le Cam's third lemma shows that, under $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$ with $\kappa_n = \tau p/\sqrt{n}$ $(\tau > 0)$, $T_{\boldsymbol{\theta}}^{(n)}$ converges weakly to a Gaussian random variable with mean $\Gamma_p^{1/2}\tau$ and variance one. Therefore, the corresponding asymptotic power of $\phi_{\boldsymbol{\theta}+}^{(n)}$ is

$$\lim_{n\to\infty}\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}[T_{\boldsymbol{\theta}}^{(n)} > z_\alpha] = 1 - \Phi\big(z_\alpha - \Gamma_p^{1/2}\tau\big). \tag{6}$$

Note that this asymptotic power does not converge to $\alpha$ as $p$ diverges to infinity. This may be surprising at first since departures from uniformity here are of a *single-spiked* nature, that is, only materialize in a single direction out of the $p$ directions in $\mathcal{S}^{p-1}$. The fact that this asymptotic power does not fade out for larger dimensions is actually explained by the fact that we did not consider local alternatives associated with $\kappa_n = \tau/\sqrt{n}$ but rather with $\kappa_n = \tau p/\sqrt{n}$, which properly scales local alternatives for different dimensions $p$ (note also that the behaviour of high-dimensional Watson distributions—see Chapter 8 in Chikuse (2003)—intuitively explains that, at least in the Watson case, the concentration should indeed scale linearly with the dimension).

Optimal tests for the other one-sided problem and for the two-sided problem are obtained in a similar way. More precisely, for the problem of testing $\mathcal{H}_0^{(n)} :$ $\{\mathrm{P}_0^{(n)}\}$ against $\cup_{\kappa<0}\cup_{f\in\mathcal{F}}\{\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$, the test $\phi_{\boldsymbol{\theta}-}^{(n)}$ rejecting the null hypothesis of uniformity at asymptotic level $\alpha$ whenever $T_{\boldsymbol{\theta}}^{(n)} < -z_\alpha$ is locally asymptotically

most powerful and has asymptotic power

$$\lim_{n\to\infty} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa_n,f}[T^{(n)}_{\boldsymbol{\theta}} < -z_\alpha] = \Phi(-z_\alpha - \Gamma_p^{1/2}\tau)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa_n,f}$ with $\kappa_n = \tau p/\sqrt{n}$ ($\tau < 0$). The corresponding two-sided test, $\phi^{(n)}_{\boldsymbol{\theta}\pm}$ say, rejects the null hypothesis at asymptotic level $\alpha$ whenever $|T^{(n)}_{\boldsymbol{\theta}}| > z_{\alpha/2}$. This test is locally asymptotically maximin for the two-sided problem and has asymptotic power

$$\lim_{n\to\infty} \mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa_n,f}[|T^{(n)}_{\boldsymbol{\theta}}| > z_{\alpha/2}] = 2 - \Phi(z_{\alpha/2} - \Gamma_p^{1/2}\tau) - \Phi(z_{\alpha/2} + \Gamma_p^{1/2}\tau)$$

under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa_n,f}$ with $\kappa_n = \tau p/\sqrt{n}$ ($\tau \neq 0$). Again, the local asymptotic powers of these tests do not fade out for larger dimensions $p$ but rather converge to a constant larger than $\alpha$.

We conducted the following Monte Carlo exercise in order to check the validity of our asymptotic results. For any combination $(n, p)$ of sample size $n \in \{100, 1000\}$ and dimension $p \in \{3, 10\}$, we generated collections of 5000 independent random samples of size $n$ from the Watson distribution with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa_n = \tau p/\sqrt{n}$, for $\tau = -2, -1, 0, 1, 2$; see Section 2. The value $\tau = 0$ corresponds to the null hypothesis of uniformity over $\mathcal{S}^{p-1}$, whereas the larger the non-zero value of $|\tau|$ is, the more severe the alternative is. Kernel density estimates of the resulting values of the test statistic $T^{(n)}_{\boldsymbol{\theta}}$ in (5) are provided in Figure 1, that further plots the densities of the corresponding asymptotic distributions (for the null case $\tau = 0$, histograms of the values of $T^{(n)}_{\boldsymbol{\theta}}$ are also shown). Clearly, these simulations confirm our asymptotic results (yet, unsurprisingly, larger dimensions require larger sample sizes for asymptotic results to materialize).

## 4. The unspecified location case: The Bingham test

We now turn to the unspecified-$\boldsymbol{\theta}$ version of the testing problems considered in the previous section. We focus first on the one-sided problem of testing $\mathcal{H}^{(n)}_0 : \{\mathrm{P}^{(n)}_0\}$ against $\mathcal{H}^{(n)}_1 : \cup_{\boldsymbol{\theta}\in\mathcal{S}^{p-1}} \cup_{\kappa>0} \cup_{f\in\mathcal{F}} \{\mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa,f}\}$. It is convenient to reparameterize the submodel associated with $\kappa \geq 0$ by writing $\boldsymbol{\vartheta} = \sqrt{\kappa}\boldsymbol{\theta}$. In this new parametrization (which, unlike the original curved one, is flat), the testing problem writes $\mathcal{H}^{(n)}_0 : \{\mathrm{P}^{(n)}_{\mathbf{0}} = \mathrm{P}^{(n)}_0\}$ against $\cup_{\boldsymbol{\vartheta}\in\mathbb{R}^p\backslash\{\mathbf{0}\}} \cup_{f\in\mathcal{F}} \{\mathrm{P}^{(n)}_{\boldsymbol{\vartheta},f}\}$, that is, simply consists in testing $\mathcal{H}^{(n)}_0 : \boldsymbol{\vartheta} = \mathbf{0}$ against $\mathcal{H}^{(n)}_1 : \boldsymbol{\vartheta} \neq \mathbf{0}$. Theorem 4.1 below then describes the asymptotic behavior of the corresponding local log-likelihood ratios.

To be able to state the result, we need to introduce the following notation. For a matrix $\mathbf{A}$, we will write $\mathrm{vec}\,\mathbf{A}$ for the vector obtained by stacking the columns of $\mathbf{A}$ on top of each other. We let $\mathbf{J}_p := (\mathrm{vec}\,\mathbf{I}_p)(\mathrm{vec}\,\mathbf{I}_p)'$, where $\mathbf{I}_\ell$ is the $\ell \times \ell$ identity matrix. Finally, with the usual Kronecker product $\otimes$, the $p^2 \times p^2$ commutation matrix is $\mathbf{K}_p := \sum_{i,j=1}^{p}(\mathbf{e}_i\mathbf{e}_j') \otimes (\mathbf{e}_j\mathbf{e}_i')$, where $e_\ell$ is the $\ell$th vector of the canonical basis of $\mathbb{R}^p$. We then have the following result.
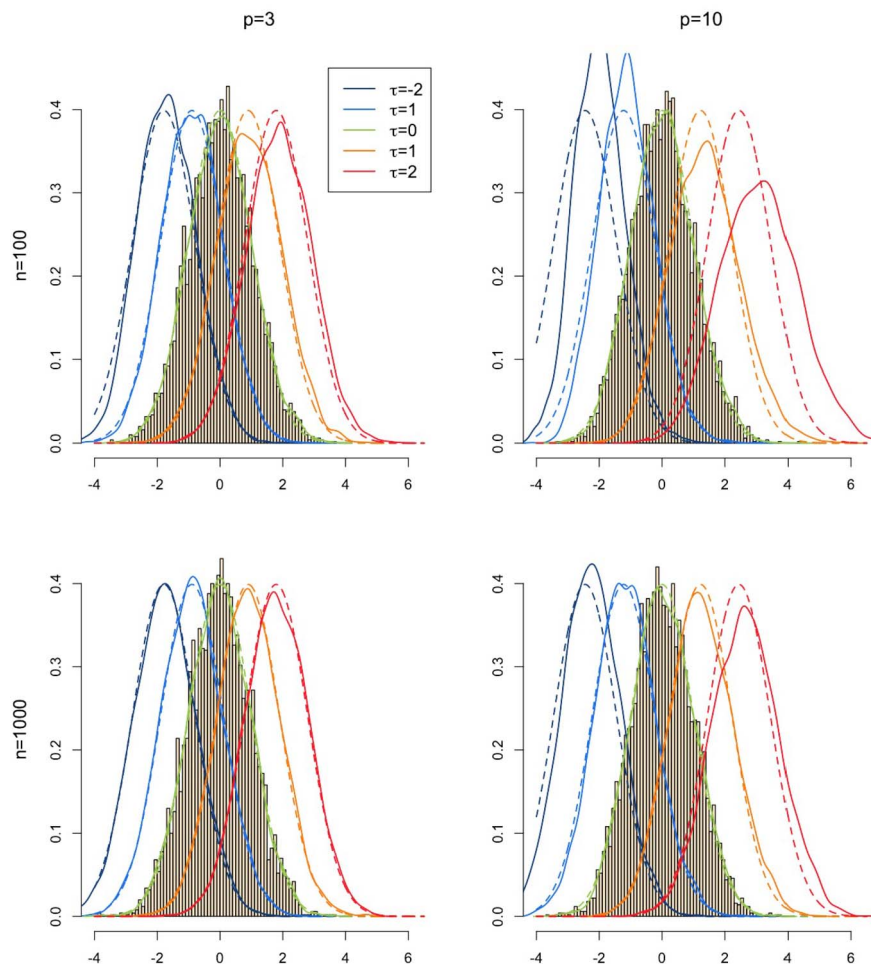
p=3    p=10



FIG 1. *Plots of the kernel density estimates (solid curves) of the values of the test statistic $T_{\boldsymbol{\theta}}^{(n)}$ in (5) obtained from $M = 5000$ independent random samples, of size $n = 100$ (top) or $1000$ (bottom), from the Watson distribution with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa = \tau p/\sqrt{n}$, with $\tau = -2, -1, 0, 1, 2$ and with $p = 3$ (left) or $p = 10$ (right); for $\tau = 0$, histograms of the values of $T_{\boldsymbol{\theta}}^{(n)}$ are shown. The densities of the corresponding asymptotic $(\mathcal{N}(\Gamma_p^{1/2}\tau, 1))$ distributions are also plotted (dashed curves). Throughout this paper, kernel density estimates are obtained from the R command* `density` *with default parameter values.*

**Theorem 4.1.** *Fix $p \in \{2, 3, \ldots\}$ and $f \in \mathcal{F}$. Let $\boldsymbol{\vartheta}_n = (p/\sqrt{n})^{1/2}\boldsymbol{\tau}_n$, where $(\boldsymbol{\tau}_n)$ is a sequence in $\mathbb{R}^p$ that is $O(1)$ but not $o(1)$. Then, letting*

$$\boldsymbol{\Delta}^{(n)} := p\sqrt{n}\,\mathrm{vec}\left(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\right) \quad and \quad \boldsymbol{\Gamma}_p := \frac{p}{p+2}\left(\mathbf{I}_{p^2} + \mathbf{K}_p - \frac{2}{p}\mathbf{J}_p\right),$$

*we have that, as $n \to \infty$ under $\mathrm{P}_0^{(n)}$,*

$$\log \frac{d\mathrm{P}_{\boldsymbol{\vartheta}_{n,f}}^{(n)}}{d\mathrm{P}_0^{(n)}} = (\mathrm{vec}(\boldsymbol{\tau}_n \boldsymbol{\tau}_n'))'\boldsymbol{\Delta}^{(n)} - \frac{1}{2}(\mathrm{vec}(\boldsymbol{\tau}_n \boldsymbol{\tau}_n'))'\boldsymbol{\Gamma}_p \mathrm{vec}(\boldsymbol{\tau}_n \boldsymbol{\tau}_n') + o_{\mathrm{P}}(1), \quad (7)$$

*where $\boldsymbol{\Delta}^{(n)}$ is, still under $\mathrm{P}_0^{(n)}$, asymptotically normal with mean vector zero and covariance matrix $\boldsymbol{\Gamma}_p$.*

Theorem 4.1 shows that the contiguity rate for $\boldsymbol{\vartheta}$ is $n^{-1/4}$, which corresponds to the contiguity rate $n^{-1/2}$ obtained for $\kappa$ in Theorem 3.1 (recall that $\boldsymbol{\vartheta} = \sqrt{\kappa}\boldsymbol{\theta}$); however, as we will explain below, the limiting experiment in Theorem 4.1 is non-standard. Writing $\mathbf{A}^-$ for the Moore-Penrose generalized inverse of $\mathbf{A}$, a natural test of uniformity is the test rejecting the null hypothesis at asymptotic level $\alpha$ whenever

$$Q^{(n)} := (\boldsymbol{\Delta}^{(n)})'\boldsymbol{\Gamma}_p^- \boldsymbol{\Delta}^{(n)} = \frac{np(p+2)}{2}\left(\mathrm{tr}[\mathbf{S}_n^2] - \frac{1}{p}\right) > \chi_{d_p,1-\alpha}^2, \quad (8)$$

where we denoted as $\chi_{d_p,1-\alpha}^2$ the upper $\alpha$-quantile of the chi-square distribution with $d_p := p(p+1)/2 - 1$ degrees of freedom. This test, which rejects the null hypothesis when the sample variance of the eigenvalues $\hat{\lambda}_{n1}, \ldots, \hat{\lambda}_{np}$ of $\mathbf{S}_n$ is too large, also addresses the problem of testing uniformity against the one-sided alternatives associated with $\kappa < 0$ or against the two-sided alternatives associated with $\kappa \neq 0$. This procedure, which is known as the Bingham (1974) test (hence will be denoted as $\phi_{\mathrm{Bing}}$ in the sequel), is often regarded as the simplest test of uniformity for axial data; see Section 10.7 in Mardia and Jupp (2000). When applied to the unit vectors $\mathbf{X}_{ni} = \mathbf{Z}_{ni}/\|\mathbf{Z}_{ni}\|$, $i = 1, \ldots, n$, obtained from Euclidean data $\mathbf{Z}_{ni}$, $i = 1, \ldots, n$, this test is also the sign test of sphericity from Hallin and Paindaveine (2006), and it follows from that paper that, as an axial test for uniformity on the sphere, the Bingham test is optimal against *angular Gaussian alternatives* (see Tyler, 1987), that is, against projections of elliptical distributions on the sphere.

Local asymptotic powers of the Bingham test can be obtained from the LAN result in Theorem 3.1 and Le Cam's third lemma. We have the following result.

**Theorem 4.2.** *Fix $p \in \{2, 3, \ldots\}$, $\boldsymbol{\theta} \in \mathcal{S}^{p-1}$, and $f \in \mathcal{F}$. Let $\kappa_n = \tau_n p/\sqrt{n}$, where the real sequence $(\tau_n)$ converges to $\tau$. Then, under $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$,*

$$Q^{(n)} \xrightarrow{\mathcal{D}} \chi_{d_p}^2\left(\frac{2(p-1)\tau^2}{p+2}\right), \quad (9)$$

*where $\chi_\ell^2(\delta)$ denotes the non-central chi-square distribution with $\ell$ degrees of freedom and non-centrality parameter $\delta$. Under the same sequence of alternatives, the asymptotic power of the Bingham test is therefore*

$$\lim_{n \to \infty} \mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}[Q^{(n)} > \chi_{d_p,1-\alpha}^2] = 1 - \Psi_{d_p}\left(\chi_{d_p,1-\alpha}^2; \frac{2(p-1)\tau^2}{p+2}\right), \quad (10)$$

*where $\Psi_\ell(\cdot; \delta)$ is the cumulative distribution function of the $\chi_\ell^2(\delta)$ distribution.*
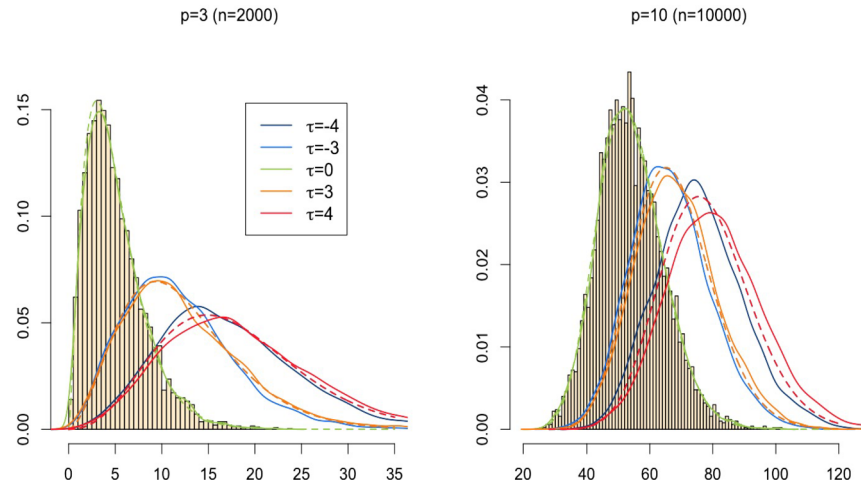
FIG 2. *(Left:) Plots of the kernel density estimates (solid curves) of the values of the Bingham test statistic $Q^{(n)}$ in (8) obtained from $M = 5000$ independent random samples of size $n = 2000$ from the Watson distribution with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa = \tau p / \sqrt{n}$, with $\tau = -4, -3, 0, 3, 4$ and with $p = 3$; for $\tau = 0$, histograms of the values of $Q^{(n)}$ are shown. The density of the corresponding asymptotic distributions in (9), which do not depend on the sign of $\tau$, are also plotted (dashed curves). (Right:) The corresponding results for $p = 10$ and $n = 10000$.*

This result in particular shows that the Bingham test is a two-sided procedure, as the asymptotic power in (10) exhibits a symmetric pattern with respect to girdle-type alternatives ($\tau < 0$) and bipolar alternatives ($\tau > 0$). This power, unlike the powers of the specified-$\boldsymbol{\theta}$ tests in the previous section, converges to $\alpha$ as $p$ diverges to infinity, which materializes the fact that, for larger dimensions, the Bingham test severely suffers (even asymptotically) from the unspecification of $\boldsymbol{\theta}$. Note also that since the Bingham test is invariant with respect to rotations, its limiting power naturally does not depend on the location parameter $\boldsymbol{\theta}$ under the alternative.

We conducted the following simulation exercise to check the validity of the asymptotic results of this section. In dimension $p = 3$, we generated 5000 mutually independent random samples of size $n = 2000$ from the Watson distribution with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa_n = \tau p / \sqrt{n}$, for $\tau = -4, -3, 0, 3, 4$; see Section 2. We did the same in dimension $p = 10$, with sample size $n = 10000$. For both dimensions $p$, Figure 2 reports kernel density estimates of the resulting values of the Bingham test statistic $Q^{(n)}$. They perfectly match with the corresponding asymptotic distribution in (9). The results also confirm the two-sided nature of the Bingham test, that, irrespective of $\tau_0$, asymptotically behaves in the exact same way under $\tau = \pm \tau_0$.

Our results indicate that the Bingham test shows non-trivial asymptotic powers under the contiguous alternatives identified in the previous section. However, a key point is the following: if $(\boldsymbol{\tau}_n) \to \boldsymbol{\tau}$ in the LAN result of Theorem 4.1,

then, under $P_{\boldsymbol{\vartheta}_{n},f}^{(n)}$ with $\boldsymbol{\vartheta}_n = (p/\sqrt{n})^{1/2}\boldsymbol{\tau}_n$, $\boldsymbol{\Delta}^{(n)}$ is asymptotically normal with mean $\mathbf{s}_{\boldsymbol{\tau}} = \boldsymbol{\Gamma}_p \mathrm{vec}(\boldsymbol{\tau}\boldsymbol{\tau}')$ and covariance matrix $\boldsymbol{\Gamma}_p$, so that the sequence of asymptotic experiments at hand does converge to a Gaussian shift experiment ($\boldsymbol{\Delta} \sim \mathcal{N}(\mathbf{s}_{\boldsymbol{\tau}}, \boldsymbol{\Gamma}_p)$) involving a *constrained* shift $\mathbf{s}_{\boldsymbol{\tau}}$. As a result, the Bingham test is not Le Cam optimal for the considered problem: this test, which would rather be Le Cam optimal (more precisely, locally asymptotically maximin) for an unconstrained shift $\mathbf{s} \in \mathbb{R}^{p^2}$, is here "wasting" power against multi-spiked alternatives that are incompatible with the present single-spiked axial model. This is in line with the fact that the Bingham test, which rejects the null hypothesis of uniformity when the sample variance of the eigenvalues $\hat{\lambda}_{n1}, \ldots, \hat{\lambda}_{np}$ of $\mathbf{S}_n$ is too large, uses these eigenvalues in a permutation-invariant way (in the considered single-spiked models, it would be more natural to consider specifically $\hat{\lambda}_{n1}$ and/or $\hat{\lambda}_{np}$ to detect possible deviations from uniformity).

## 5. The unspecified location case: Single-spiked tests

A natural question is then: how to construct a test that is more powerful than the Bingham test? We now describe two constructions that actually lead to the same test(s). Focusing again at first on the one-sided problem involving the bipolar alternatives, we saw in Section 3 that, in the specified location case, Le Cam optimal tests of uniformity reject $\mathcal{H}_0^{(n)} : \{P_0^{(n)}\}$ in favor of $\cup_{\kappa>0} \cup_{f \in \mathcal{F}} \{P_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$ for large values of $\Delta_{\boldsymbol{\theta}}^{(n)} = \sqrt{n}(p\,\boldsymbol{\theta}'\mathbf{S}_n\boldsymbol{\theta} - 1)$. In the unspecified location case, it is then natural, following Davies (1977, 1987, 2002), to consider the test, $\phi_+^{(n)}$ say, rejecting the null hypothesis of uniformity at asymptotic level $\alpha$ when

$$T_+^{(n)} := \sup_{\boldsymbol{\theta} \in \mathcal{S}^{p-1}} \Delta_{\boldsymbol{\theta}}^{(n)} = \sqrt{n}(p\hat{\lambda}_{n1} - 1) > c_{p,\alpha,+}, \tag{11}$$

where $\hat{\lambda}_{n1}$ still denotes the largest eigenvalue of $\mathbf{S}_n$ and $c_{p,\alpha,+}$ is such that this test has asymptotic size $\alpha$ under the null hypothesis.

A similar rationale yields natural tests for the other one-sided problem and for the two-sided problem: since Le Cam optimal tests of uniformity reject $\mathcal{H}_0^{(n)} : \{P_0^{(n)}\}$ in favor of $\cup_{\kappa<0} \cup_{f \in \mathcal{F}} \{P_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$ for large values of $-\Delta_{\boldsymbol{\theta}}^{(n)} = \sqrt{n}(p\,\boldsymbol{\theta}'\mathbf{S}_n\boldsymbol{\theta} - 1)$, the resulting unspecified-$\boldsymbol{\theta}$ test, $\phi_-^{(n)}$ say, will reject the null hypothesis of uniformity at asymptotic level $\alpha$ when

$$T_-^{(n)} := \sup_{\boldsymbol{\theta} \in \mathcal{S}^{p-1}} (-\Delta_{\boldsymbol{\theta}}^{(n)}) = -\sqrt{n}(p\hat{\lambda}_{np} - 1) > c_{p,\alpha,-}, \tag{12}$$

where $c_{p,\alpha,-}$ is such that this test has asymptotic size $\alpha$ under the null hypothesis. Finally, since Le Cam optimal tests of uniformity reject $\mathcal{H}_0^{(n)} : \{P_0^{(n)}\}$ in favor of $\cup_{\kappa\neq0} \cup_{f \in \mathcal{F}} \{P_{\boldsymbol{\theta},\kappa,f}^{(n)}\}$ for large values of $|\Delta_{\boldsymbol{\theta}}^{(n)}| = \sqrt{n}|p\,\boldsymbol{\theta}'\mathbf{S}_n\boldsymbol{\theta} - 1|$, the resulting unspecified-$\boldsymbol{\theta}$ test, $\phi_{\pm}^{(n)}$ say, will reject the null hypothesis of uniformity at asymptotic level $\alpha$ when

$$T_{\pm}^{(n)} := \sup_{\boldsymbol{\theta} \in \mathcal{S}^{p-1}} |\Delta_{\boldsymbol{\theta}}^{(n)}| = \sqrt{n}\max\left(|p\hat{\lambda}_{n1} - 1|, |p\hat{\lambda}_{np} - 1|\right) > c_{p,\alpha,\pm}, \tag{13}$$

where $c_{p,\alpha,\pm}$ is still such that this test has asymptotic size $\alpha$ under the null hypothesis of uniformity.

Another rationale for considering the above tests is the following. For the sake of brevity, let us focus on the one-sided problem involving the bipolar alternatives, that is, the ones associated with $\kappa > 0$. A natural idea to obtain an unspecified-$\boldsymbol{\theta}$ test is to replace $\boldsymbol{\theta}$ in the corresponding optimal specified-$\boldsymbol{\theta}$ test $\phi_{\boldsymbol{\theta}+}^{(n)}$ with an estimator $\hat{\boldsymbol{\theta}}_n$. Now, under $\mathrm{P}_{\boldsymbol{\theta},\kappa,f}^{(n)}$, we have $\mathrm{E}[\mathbf{X}_{n1}] = \mathbf{0}$ and

$$\mathrm{E}[\mathbf{X}_{n1}\mathbf{X}_{n1}'] = g_f(\kappa)\boldsymbol{\theta}\boldsymbol{\theta}' + \frac{1 - g_f(\kappa)}{p - 1}(\mathbf{I}_p - \boldsymbol{\theta}\boldsymbol{\theta}')$$

(see, e.g., Lemma B.3(i) in Cutting, Paindaveine and Verdebout, 2017), with

$$g_f(\kappa) := \mathrm{E}_{\boldsymbol{\theta},\kappa,f}[(\mathbf{X}_{n1}'\boldsymbol{\theta})^2] = c_{p,\kappa,f} \int_{-1}^{1}(1 - s^2)^{(p-3)/2}s^2 f(\kappa s^2)\,ds.$$

It is easy to check that, for any $f \in \mathcal{F}$, the function $\kappa \mapsto g_f(\kappa)$ is differentiable at 0, with derivative $g_f'(0) = \mathrm{Var}_0^{(n)}[(\mathbf{X}_{n1}'\boldsymbol{\theta})^2] > 0$, where $\mathrm{Var}_0^{(n)}$ still denotes variance under $\mathrm{P}_0^{(n)}$. Consequently, for $\kappa > 0$ small, we have $g_f(\kappa) > g_f(0) = 1/p$, so that $\boldsymbol{\theta}$ is, up to an unimportant sign (recall that only the pair $\{\pm\boldsymbol{\theta}\}$ is identifiable), the leading unit eigenvector of $\mathrm{E}[\mathbf{X}_{n1}\mathbf{X}_{n1}']$ (for many functions $f$, including the Watson one $f(z) = \exp(z)$, this remains true for any $\kappa > 0$). Therefore, a moment estimator of $\boldsymbol{\theta}$ is the leading eigenvector $\hat{\boldsymbol{\theta}}_n$ of $\mathbf{S}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{ni}\mathbf{X}_{ni}'$. Note that in the Watson parametric submodel $\{\mathrm{P}_{\boldsymbol{\theta},\kappa,\exp}^{(n)} : \kappa > 0\}$, this estimator $\hat{\boldsymbol{\theta}}_n$ is also the MLE of $\boldsymbol{\theta}$. The resulting test then rejects the null hypothesis of uniformity for large values of

$$\Delta_{\hat{\boldsymbol{\theta}}_n}^{(n)} := \sqrt{n}\big(p\hat{\boldsymbol{\theta}}_n'\mathbf{S}_n\hat{\boldsymbol{\theta}}_n - 1\big) = T_+^{(n)},$$

hence coincides with the test $\phi_+^{(n)}$ in (11). A similar reasoning for the other one-sided problem leads to the test $\phi_-^{(n)}$.

The critical values in (11)–(13) above can of course be obtained from the asymptotic distribution of the corresponding test statistics under the null hypothesis. For $p = 3$, the asymptotic null distributions of $T_+^{(n)}$ and $T_-^{(n)}$ were obtained in Anderson and Stephens (1972), where the corresponding one-sided tests $\phi_+^{(n)}$ and $\phi_-^{(n)}$ were first proposed. We extend their result to the two-sided test statistic $T_\pm^{(n)}$ and, more importantly, to the non-null case. The key to do so is the following result.

**Theorem 5.1.** *Fix $p \in \{2, 3, \ldots\}$ and $f \in \mathcal{F}$. Let $\mathbf{Z}$ be a $p \times p$ random matrix such that* $\mathrm{vec}\,\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_p)$*, with* $\mathbf{V}_p = (p/(p+2))(\mathbf{I}_{p^2} + \mathbf{K}_p) - (2/(p+2))\mathbf{J}_p$. *Then, (i) under $\mathrm{P}_0^{(n)}$,*

$$\begin{pmatrix} \sqrt{n}(p\hat{\lambda}_{n1} - 1) \\ \sqrt{n}(p\hat{\lambda}_{np} - 1) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} L_p^{\max} \\ L_p^{\min} \end{pmatrix},$$

where $L_{p,\max}$ *(resp., $L_{p,\min}$) is the largest (resp., smallest) eigenvalue of* $\mathbf{Z}$; *(ii) under* $\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$, *where* $\kappa_n = \tau_n p/\sqrt{n}$ *involves a real sequence* $(\tau_n)$ *converging to* $\tau$,

$$\begin{pmatrix} \sqrt{n}(p\hat{\lambda}_{n1} - 1) \\ \sqrt{n}(p\hat{\lambda}_{np} - 1) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} L_{p,\tau}^{\max} \\ L_{p,\tau}^{\min} \end{pmatrix},$$

*where $L_{p,\tau}^{\max}$ (resp., $L_{p,\tau}^{\min}$) is the largest (resp., smallest) eigenvalue of* $\mathbf{Z}_\tau :=$ $\mathbf{Z} + (2\tau/(p+2))\mathbf{W}_\tau$, *with* $\mathbf{W}_\tau := \mathrm{diag}(p-1,-1,\ldots,-1)$ *for* $\tau \geq 0$ *and* $\mathbf{W}_\tau :=$ $\mathrm{diag}(-1,\ldots,-1,p-1)$ *for* $\tau < 0$.

A direct consequence of Theorem 5.1(i) is that simulations can be used to obtain arbitrarily precise estimates of the asymptotic critical values needed to implement the tests $\phi_+^{(n)}$, $\phi_-^{(n)}$ and $\phi_\pm^{(n)}$. For instance, the test $\phi_+^{(n)}$ will reject the null hypothesis of uniformity at asymptotic level $\alpha$ whenever $T_+^{(n)} = \sqrt{n}(p\hat{\lambda}_{n1} - 1) > \hat{c}_{p,\alpha,+}^{(m)}$, where $\hat{c}_{p,\alpha,+}^{(m)}$ denotes the upper $\alpha$-quantile of $m$ independent realizations of $L_p^{\max}$. Interestingly, the following corollary shows that simulations can actually be avoided in dimensions $p = 2$ and $p = 3$, as the asymptotic null distribution of $T_+^{(n)}$, $T_-^{(n)}$ and $T_\pm^{(n)}$ can be explicitly determined for these values of $p$ (the result for $T_+^{(n)}$ and $T_-^{(n)}$ in dimension $p = 3$ in (15) below agrees with the one from Anderson and Stephens, 1972).

**Corollary 5.1.** *(i) Under the null hypothesis of uniformity over $\mathcal{S}^1$, the test statistics $T_+^{(n)}$, $T_-^{(n)}$, and $T_\pm^{(n)}$ converge weakly to $L_2^{\max}$, where $L_2^{\max}$ has cumulative distribution function*

$$\ell \mapsto (1 - \exp(-\ell^2))\,\mathbb{I}[\ell > 0]; \tag{14}$$

*(ii) under the null hypothesis of uniformity over $\mathcal{S}^2$, the test statistics $T_+^{(n)}$ and $T_-^{(n)}$ converge weakly to $L_3^{\max}$, where $L_3^{\max}$ has cumulative distribution function*

$$\ell \mapsto \left\{ \Phi(\sqrt{5}\ell) + \Phi(\tfrac{\sqrt{5}\ell}{2}) + 3\Phi''(\tfrac{\sqrt{5}\ell}{2}) - 1 \right\} \mathbb{I}[\ell > 0], \tag{15}$$

*whereas the test statistic $T_\pm^{(n)}$ converges weakly to $L_3 := \max(L_3^{\max}, -L_3^{\min})$, where $L_3$ has cumulative distribution function*

$$\ell \mapsto \left\{ 2\Phi(\tfrac{\sqrt{5}\ell}{2}) + 6\Phi''(\tfrac{\sqrt{5}\ell}{2}) - 2\sqrt{3}\Phi''(\tfrac{\sqrt{5}\ell}{\sqrt{3}}) - 1 \right\} \mathbb{I}[\ell > 0] \tag{16}$$

*(here, $\Phi''$ is the second derivative of the standard normal distribution function $\Phi$).*

Writing $\lambda_\ell(\mathbf{A})$ for the $\ell$th largest eigenvalue of the $p \times p$ matrix $\mathbf{A}$ and denoting as $\overset{\mathcal{D}}{=}$ equality in distribution, Theorem 5.1 entails that, under the null hypothesis,

$$T_+^{(n)} \xrightarrow{\mathcal{D}} L_p^{\max} = \lambda_1(\mathbf{Z}) \overset{\mathcal{D}}{=} \lambda_1(-\mathbf{Z}) = -\lambda_p(\mathbf{Z}) = -L_p^{\min} \overset{\mathcal{D}}{\leftarrow} T_-^{(n)}.$$

This shows that, for any dimension $p$, the test statistics $T_+^{(n)}$ and $T_-^{(n)}$ share the same weak limit under the null hypothesis, which is confirmed in dimensions $p = 2, 3$ by Corollary 5.1. Maybe surprisingly, this corollary further implies that, for $p = 2$, the two-sided test statistic $T_\pm^{(n)}$ has the same asymptotic null distribution as $T_+^{(n)}$ and $T_-^{(n)}$. This can be explained as follows: since $\mathbf{S}_n$ has trace one almost surely, its eigenvalues $\hat{\lambda}_{n\ell}$, $\ell = 1, \ldots, p$ sum up to one almost surely (incidentally, this implies that the quantities $\sqrt{n}(p\hat{\lambda}_{n\ell} - 1)$, $\ell = 1, \ldots, p$, do not admit a joint density, not even asymptotically so, which makes the proof of Theorem 5.1 rather challenging). For $p = 2$, it follows that $T_+^{(n)} = T_-^{(n)} = T_\pm^{(n)}$ almost surely, which of course entails that these three tests statistics share the same weak limit, not only under the null hypothesis but under *any* sequence of hypotheses. In line with this, $L_{2,\tau}^{\max} = -L_{2,\tau}^{\min} = \max(L_{2,\tau}^{\max}, -L_p^{\min})$ almost surely for any $\tau \in \mathbb{R}$.

To check the validity of Theorem 5.1 and Corollary 5.1, we conducted the following numerical exercises in dimensions $p = 3$ and $p = 10$. We generated 5000 mutually independent random samples of size $n = 2000$ (for $p = 3$) and $n = 10000$ (for $p = 10$) from the Watson distribution with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa_n = \tau p/\sqrt{n}$, for $\tau = -4, -3, 0, 3, 4$. Figure 3 plots kernel density estimates of the resulting values of $T_+^{(n)}$, $T_-^{(n)}$ and $T_\pm^{(n)}$, along with the densities of the corresponding asymptotic distributions; for $p = 3$ and $\tau = 0$, these densities are those associated with the distribution functions in (15)–(16), whereas, in all other cases, they are kernel density estimates obtained from $10^6$ independent realizations of $L_{p,\tau}^{\max}$, $-L_{p,\tau}^{\min}$, and $\max(L_{p,\tau}^{\max}, -L_{p,\tau}^{\min})$, respectively; see Theorem 5.1. Clearly, the results support our asymptotic findings. It is seen that the one-sided test $\phi_+^{(n)}$ not only shows power against the bipolar alternatives it is designed for (those associated with $\tau > 0$) but also against girdle-type ones (those associated with $\tau < 0$), which is actually desirable. The same can be said about the one-sided test $\phi_-^{(n)}$, but each of these tests, of course, shows higher powers against the alternatives it was designed for. In contrast, the two-sided test $\phi_\pm^{(n)}$ shows a symmetric power pattern for positive and negative values of $\tau$.

## 6. Finite-sample comparisons

In the previous sections, we conducted Monte-Carlo exercises in order to check the correctness of our null and non-null asymptotic results, but it is of course of primary importance to compare the power behaviors of the various tests considered in this work. In this section, we therefore study the finite-sample powers of the Bingham test $\phi_{\text{Bing}}^{(n)}$ and of the test $\phi_+^{(n)}$ (we could similarly consider the tests $\phi_-^{(n)}$ and $\phi_\pm^{(n)}$), and we compare them with those of the optimal specified-$\boldsymbol{\theta}$ test $\phi_{\boldsymbol{\theta}+}^{(n)}$. Our asymptotic results further allow us to complement these finite-sample comparisons with comparisons of the corresponding asymptotic powers.

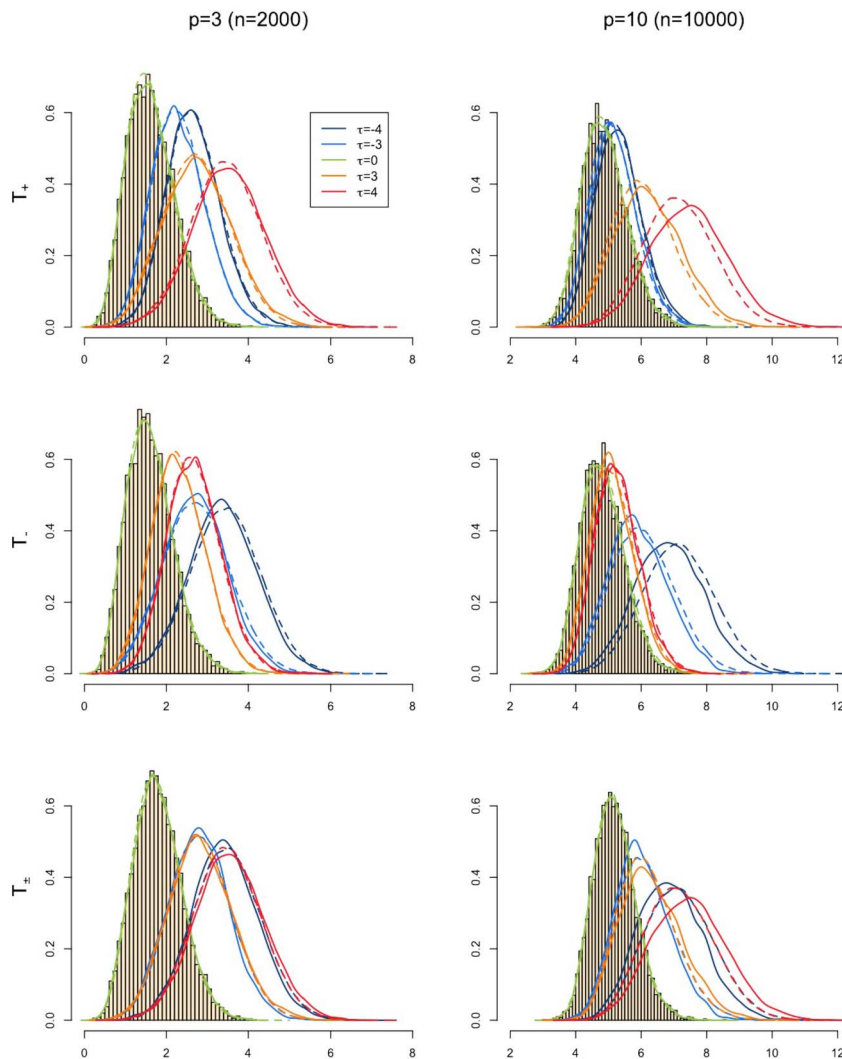We conducted the following Monte Carlo experiment. For any combina-

p=3 (n=2000)                    p=10 (n=10000)



FIG 3. *(Left:) Plots of the kernel density estimates (solid curves) of the values of $T_+^{(n)} = \sqrt{n}(p\hat{\lambda}_{n1}-1)$ (top), $T_-^{(n)} = -\sqrt{n}(p\hat{\lambda}_{np}-1)$ (middle) or $T_{\pm}^{(n)} = \max(T_+^{(n)}, T_-^{(n)})$ (bottom), obtained from 5000 independent random samples of size $n = 2000$ from the Watson distribution with location $\boldsymbol{\theta} = (1,0,\ldots,0)' \in \mathbb{R}^p$ and concentration $\kappa = \tau p/\sqrt{n}$, with $\tau = -4,-3,0,3,4$ and with $p = 3$; for $\tau = 0$, histograms of the corresponding test statistics are shown. The density of the corresponding asymptotic distributions are also plotted (dashed curves). (Right:) The corresponding results for $p = 10$ and $n = 10000$. See Section 5 for details.*

tion $(n,p)$ of sample size $n \in \{200, 20000\}$ and dimension $p \in \{3, 10\}$, we generated collections of 2000 independent random samples of size $n$ from the Watson distribution on $\mathcal{S}^{p-1}$ with location $\boldsymbol{\theta} = (1,0,\ldots,0)' \in \mathbb{R}^p$ and concentration $\kappa_n = \tau_\ell p/\sqrt{n}$, with $\tau_\ell = 0.8\ell$, $\ell = 0,1,\ldots,5$. The value $\ell = 0$

corresponds to the null hypothesis of uniformity, whereas $\ell = 1, \ldots, 5$ provide increasingly severe bipolar alternatives. In each sample, we performed three tests at asymptotic level $\alpha = 5\%$, namely the specified-$\boldsymbol{\theta}$ test $\phi_{\boldsymbol{\theta}+}^{(n)}$ in (5), the Bingham test $\phi_{\mathrm{Bing}}^{(n)}$ in (8), and the test $\phi_+^{(n)}$ in (11); for $p = 3$, the asymptotic critical value for $\phi_+^{(n)}$ was obtained from Corollary 5.1(ii), whereas, for $p = 10$, an approximation of the corresponding critical value was obtained from 10000 independent realizations of $L_p^{\max}$ in Theorem 5.1.

Figure 4 shows the resulting empirical powers along with their theoretical asymptotic counterparts (for any given $p$ and $\tau_\ell$, the asymptotic power of $\phi_+^{(n)}$ was obtained from 10000 independent copies of the random variable $L_{p,\tau_\ell}^{\max}$ in Theorem 5.1). The results show that, as expected, the optimal specified-$\boldsymbol{\theta}$ test outperforms both unspecified-$\boldsymbol{\theta}$ tests. The test $\phi_+^{(n)}$ dominates the Bingham test $\phi_{\mathrm{Bing}}^{(n)}$ and this dominance, quite intuitively, increases with the dimension $p$. Clearly, rejection frequencies agree very well with our asymptotic results for large sample sizes.

## 7. Final comments and perspectives for future research

Axial data are commonly met in applications, which explains that statistical procedures for such data are presented in most directional statistics monographs and have been the topic of numerous research papers; we refer to Fisher, Lewis and Embleton (1987), Mardia and Jupp (2000), Ley and Verdebout (2017), and to the references therein. However, the non-null properties of axial tests of hypotheses are barely known, compared to those of their non-axial counterparts. Since this is in particular the case for axial tests of uniformity, we systematically studied in this work the asymptotic power of several axial tests of uniformity. In particular, we derived the asymptotic powers of the Bingham and Anderson tests under contiguous rotationally symmetric alternatives. Our results identify the underlying contiguity rate and allow for theoretical power comparisons. Throughout, our asymptotic findings were confirmed via simulations.

Far from being of academic interest only, our results will be useful to practitioners. First, our results guarantee that the tests we study in this work indeed show power against axial alternatives to uniformity. Second, the specific limiting powers we provide allow one to determine the sample size that should be used to achieve a given target power. Last but not least, our results suggest that a natural approach in practice is to combine the outcome of the various tests of uniformity we studied, as this will provide some insight on why a possible rejection occurs. For instance, if the Bingham test does not lead to rejection but the Anderson tests do, then we may infer that the departure from uniformity does not materialize in a single direction but is rather spread over several spikes. As another example, comparing the outcomes of the various Anderson tests will inform practitioners on the bipolar or girdle nature of the distribution.

The non-null asymptotic analysis conducted in this work essentially settles the low-dimensional case. Perspectives for future research therefore mainly re-
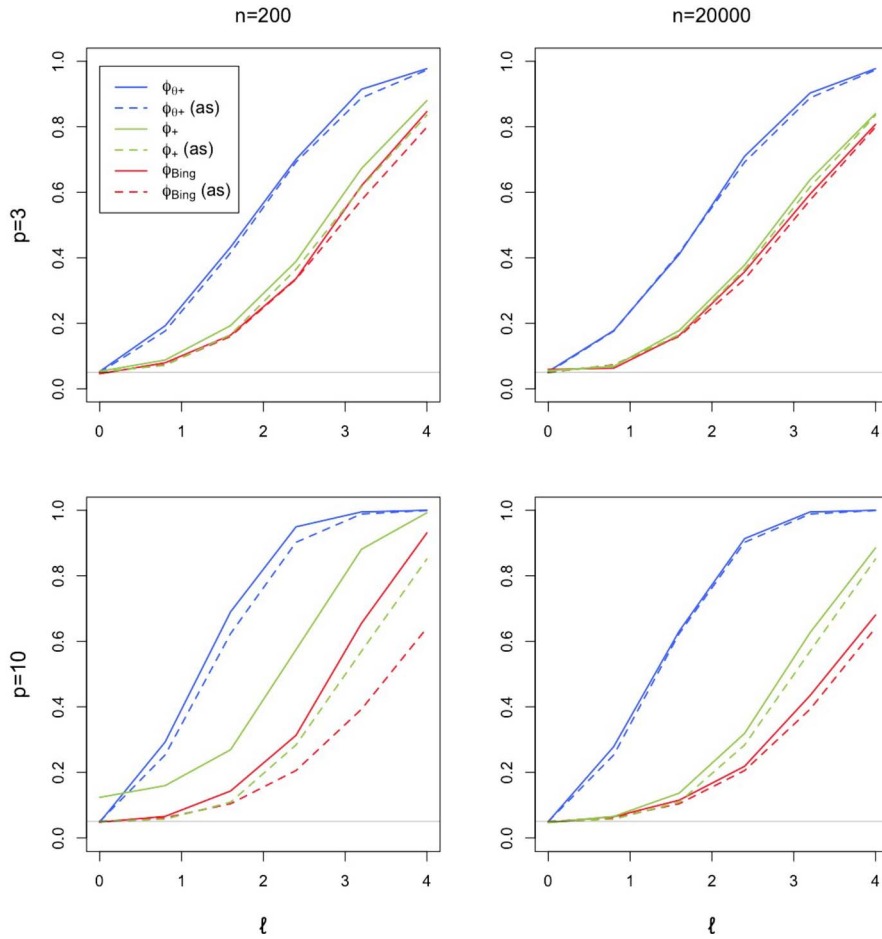
FIG 4. *Rejection frequencies (solid curves) of three axial tests of uniformity over $\mathcal{S}^{p-1}$ obtained from* 2000 *mutually independent random samples of size $n$ from the Watson distribution with location $\boldsymbol{\theta} = (1, 0, \ldots, 0)' \in \mathbb{R}^p$ and concentration $\kappa = \tau_\ell p / \sqrt{n}$, with $\ell = 0$ (null hypothesis of uniformity) and $\tau_\ell = 0.8, 1.6, 2.4, 3.2, 4.0$ (increasingly severe bipolar alternatives). The tests considered are the test $\phi_{\boldsymbol{\theta}+}^{(n)}$ in ([5]), the Bingham test $\phi_{\mathrm{Bing}}^{(n)}$ in ([8]), and the test $\phi_+^{(n)}$ in ([11]). The corresponding asymptotic powers (dashed curves) are also provided; see Section [6] for details.*

late to the high-dimensional framework. It is rather straightforward to extend the contiguity and LAN results in Theorems [2.1]–[3.1] to the case where the dimension $p = p_n$ diverges to infinity with $n$ at an arbitrary rate. However, in high dimensions, it is extremely challenging to derive the non-null asymptotic powers of the Bingham and Anderson tests under suitable local alternatives. For the Anderson test, for instance, this is due to the fact that eigenvalues of sample covariance matrices suffer complicated phase transition phenomenons which, close

to uniformity, result in a lack of consistency. These challenging questions are of course beyond the scope of the present work, hence are left for future research.

## Appendix A: Proofs of Theorems 2.1 to 4.2

These proofs require the following preliminary result.

**Lemma A.1.** *Let* $g : \mathbb{R} \to \mathbb{R}$ *be twice differentiable at 0,* $(\kappa_n)$ *be a sequence in* $\mathbb{R}_0$*, and* $(p_n)$ *be a sequence in* $\{2, 3, \ldots\}$*. Assume that* $\kappa_n$ *is* $o(p_n)$ *as* $n \to \infty$*. Then,*

$$R_n(g) := c_{p_n} \int_{-1}^{1} (1 - s^2)^{(p_n-3)/2} g(\kappa_n s^2) ds$$

$$= g(0) + \frac{\kappa_n}{p_n} g'(0) + \frac{3\kappa_n^2}{2p_n(p_n + 2)} g''(0) + o\left(\frac{\kappa_n^2}{p_n^2}\right),$$

*with* $c_p = 1 / \int_{-1}^{1} (1 - s^2)^{(p-3)/2} \, dt$.

PROOF OF LEMMA A.1. Recall that if $\mathbf{X}$ is uniformly distributed over $\mathcal{S}^{p_n-1}$, then, for any $\boldsymbol{\theta} \in \mathcal{S}^{p_n-1}$, we have that $\boldsymbol{\theta}'\mathbf{X}$ has density $c_{p_n}(1-s^2)^{(p_n-3)/2}\mathbb{I}[|s| \leq 1]$, that is, the distribution of $\boldsymbol{\theta}'\mathbf{X}$ is symmetric about zero and such that $(\boldsymbol{\theta}'\mathbf{X})^2 \sim \text{Beta}(1/2, (p_n-1)/2)$. Consequently,

$$c_{p_n} \int_{-1}^{1} s^2 (1 - s^2)^{(p_n-3)/2} \, ds = \mathrm{E}[(\boldsymbol{\theta}'\mathbf{X})^2] = \frac{1}{p_n}, \tag{17}$$

and

$$c_{p_n} \int_{-1}^{1} s^4 (1 - s^2)^{(p_n-3)/2} \, ds = \mathrm{E}[(\boldsymbol{\theta}'\mathbf{X})^4] = \frac{3}{p_n(p_n + 2)}. \tag{18}$$

From (17), we can write

$$R_n(g) - g(0) - \frac{\kappa_n}{p_n} g'(0) = c_{p_n} \int_{-1}^{1} (1-s^2)^{(p_n-3)/2} \left( g(\kappa_n s^2) - g(0) - \kappa_n s^2 g'(0) \right) ds.$$

Letting $t = |\kappa_n|^{1/2} s$ and using (17)–(18) then provides

$$R_n(g) - g(0) - \frac{\kappa_n}{p_n} g'(0)$$

$$= \frac{3\kappa_n^2}{p_n(p_n + 2)} \int_{-\infty}^{\infty} h_n(t) \left( \frac{g\left(\frac{\kappa_n}{|\kappa_n|} t^2\right) - g(0) - \frac{\kappa_n}{|\kappa_n|} t^2 g'(0)}{t^4} \right) dt,$$

or, equivalently,

$$\frac{R_n(g) - g(0) - \frac{\kappa_n}{p_n} g'(0) - \frac{3\kappa_n^2}{2p_n(p_n+2)} g''(0)}{\frac{3\kappa_n^2}{p_n(p_n+2)}}$$

$$= \int_{-\infty}^{\infty} h_n(t) \left( \frac{g\left(\frac{\kappa_n}{|\kappa_n|}t^2\right) - g(0) - \frac{\kappa_n}{|\kappa_n|}t^2 g'(0)}{t^4} \right) dt - \frac{1}{2}g''(0), \quad (19)$$

where $h_n$ is defined through

$$t \mapsto h_n(t) = \frac{\left(\frac{t}{|\kappa_n|^{1/2}}\right)^4 \left(1 - \left(\frac{t}{|\kappa_n|^{1/2}}\right)^2\right)^{(p_n-3)/2} \mathbb{I}\big[|t| \le |\kappa_n|^{1/2}\big]}{\int_{-|\kappa_n|^{1/2}}^{|\kappa_n|^{1/2}} \left(\frac{t}{|\kappa_n|^{1/2}}\right)^4 \left(1 - \left(\frac{t}{|\kappa_n|^{1/2}}\right)^2\right)^{(p_n-3)/2} dt}.$$

Since $\kappa_n$ is $o(p_n)$, it can be checked that the sequence $(h_n)$ is an *approximate δ-sequence*, in the sense that $\int_{-\infty}^{\infty} h_n(t)\, dt = 1$ for any $n$ and $\int_{-\varepsilon}^{\varepsilon} h_n(t)\, dt \to 1$ for any $\varepsilon > 0$. Hence,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} h_n(t) \left( \frac{g\left(\frac{\kappa_n}{|\kappa_n|}t^2\right) - g(0) - \frac{\kappa_n}{|\kappa_n|}t^2 g'(0)}{t^4} \right) dt$$

$$= \lim_{t\to 0} \frac{g\left(\frac{\kappa_n}{|\kappa_n|}t^2\right) - g(0) - \frac{\kappa_n}{|\kappa_n|}t^2 g'(0)}{t^4},$$

which, by using L'Hôpital's rule, is equal to

$$\lim_{t\to 0} \frac{\frac{2\kappa_n}{|\kappa_n|}tg'\left(\frac{\kappa_n}{|\kappa_n|}t^2\right) - \frac{2\kappa_n}{|\kappa_n|}tg'(0)}{4t^3} = \frac{1}{2}\lim_{t\to 0} \frac{g'\left(\frac{\kappa_n}{|\kappa_n|}t^2\right) - g'(0)}{\frac{\kappa_n}{|\kappa_n|}t^2} = \frac{1}{2}g''(0).$$

The result thus follows from (19). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

The proof of Theorems 2.1–3.1 actually only requires the particular case of lemma A.1 corresponding to $p_n \equiv p$. We still presented this more general version of the lemma as it allows one to extend Theorems 2.1–3.1 to high-dimensional asymptotic scenarios where $p_n$ would diverge to infinity with $n$ (this would prove the claims in high dimensions provided in the last paragraph of Section 2).

PROOF OF THEOREM 2.1. In this proof, all expectations and variances are taken under the null of uniformity $\mathrm{P}_0^{(n)}$ and all stochastic convergences and $o_\mathrm{P}$'s are as $n \to \infty$ under $\mathrm{P}_0^{(n)}$. Throughout, we write $\ell_{f,k} := (\log f)^k$ and $E_{nk} := \mathrm{E}[\ell_{f,k}(\kappa_n(\mathbf{X}_{ni}'\boldsymbol{\theta})^2)]$.

Consider the local log-likelihood ratio

$$\Lambda_n := \log \frac{d\mathrm{P}_{\boldsymbol{\theta},\kappa_n,f}^{(n)}}{d\mathrm{P}_0^{(n)}} = \sum_{i=1}^{n} \log \frac{c_{p,\kappa_n,f} f(\kappa_n(\mathbf{X}_{ni}'\boldsymbol{\theta})^2)}{c_p}$$

$$= n\left(\log \frac{c_{p,\kappa_n,f}}{c_p} + E_{n1}\right) + \sum_{i=1}^{n} \left(\log f(\kappa_n(\mathbf{X}_{ni}'\boldsymbol{\theta})^2) - E_{n1}\right)$$

$$=: L_{n1} + L_{n2}.$$

Lemma A.1 readily yields

$$\log \frac{c_{p,\kappa_n,f}}{c_p} = -\log\left(c_p \int_{-1}^{1} (1-s^2)^{(p-3)/2} f(\kappa_n s^2)\, ds\right)$$

$$= -\log\left(1 + \frac{\kappa_n}{p} + \frac{3\kappa_n^2}{2p(p+2)}f''(0) + o(\kappa_n^2)\right)$$

$$= -\frac{\kappa_n}{p} - \frac{3\kappa_n^2}{2p(p+2)}f''(0) + \frac{\kappa_n^2}{2p^2} + o(\kappa_n^2). \tag{20}$$

Similarly,

$$E_{n1} = c_p \int_{-1}^{1} (1-s^2)^{(p-3)/2}\ell_{f,1}(\kappa_n s^2)\,ds$$

$$= \frac{\kappa_n}{p}\ell'_{f,1}(0) + \frac{3\kappa_n^2}{2p(p+2)}\ell''_{f,1}(0) + o(\kappa_n^2)$$

$$= \frac{\kappa_n}{p} + \frac{3\kappa_n^2}{2p(p+2)}(f''(0) - 1) + o(\kappa_n^2). \tag{21}$$

Combining (20) and (21) provides

$$L_{n1} = \frac{n\kappa_n^2}{2p^2} - \frac{3n\kappa_n^2}{2p(p+2)} + o(n\kappa_n^2) = -\frac{n\kappa_n^2(p-1)}{p^2(p+2)} + o(n\kappa_n^2).$$

Turning to $L_{n2}$, write

$$L_{n2} = \sqrt{nV_n}\sum_{i=1}^{n} W_{ni} := \sqrt{nV_n}\sum_{i=1}^{n}\frac{\log f(\kappa_n(\mathbf{X}'_{ni}\boldsymbol{\theta})^2) - E_{n1}}{\sqrt{nV_n}},$$

where we let $V_n := \mathrm{Var}\big[\log f(\kappa_n(\mathbf{X}'_{ni}\boldsymbol{\theta})^2)\big]$. First note that, since

$$E_{n2} = c_p \int_{-1}^{1} (1-s^2)^{(p-3)/2}\ell_{f,2}(\kappa_n s^2)\,ds$$

$$= \frac{\kappa_n}{p}\ell'_{f,2}(0) + \frac{3\kappa_n^2}{2p(p+2)}\ell''_{f,2}(0) + o(\kappa_n^2)$$

$$= \frac{3\kappa_n^2}{p(p+2)} + o(\kappa_n^2), \tag{22}$$

we have

$$nV_n = n\big(E_{n2} - E_{n1}^2\big) = \frac{2n\kappa_n^2(p-1)}{p^2(p+2)} + o(n\kappa_n^2), \tag{23}$$

which leads to

$$\Lambda_n = -\frac{n\kappa_n^2(p-1)}{p^2(p+2)} + \sqrt{\frac{2n\kappa_n^2(p-1)}{p^2(p+2)} + o(n\kappa_n^2)}\sum_{i=1}^{n}W_{ni} + o(n\kappa_n^2). \tag{24}$$

Since $W_{ni}$, $i = 1,\ldots,n$, are mutually independent with mean zero and variance $1/n$, we obtain that

$$\mathrm{E}\big[\Lambda_n^2\big] = (\mathrm{E}[\Lambda_n])^2 + \mathrm{Var}[\Lambda_n] = \frac{n^2\kappa_n^4(p-1)^2}{p^4(p+2)^2} + o(n^2\kappa_n^4) + \frac{2n\kappa_n^2(p-1)}{p^2(p+2)} + o(n\kappa_n^2). \tag{25}$$

If $\kappa_n = o(1/\sqrt{n})$, then (25) implies that $\exp(\Lambda_n) = 1 + o_{\mathrm{P}}(1)$, so that Le Cam's first lemma yields that $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa_n,f}$ and $\mathrm{P}^{(n)}_0$ are mutually contiguous.

We may therefore assume that $\kappa_n = \tau_n/\sqrt{n}$, where $(\tau_n)$ is $O(1)$ but not $o(1)$, or, equivalently, that $\kappa_n = \tau_n p/\sqrt{n}$ with the same sequence $(\tau_n)$. Then, (24) rewrites

$$\Lambda_n = -\frac{(p-1)\tau_n^2}{p+2} + \sqrt{\frac{2(p-1)\tau_n^2}{p+2} + o(1)} \sum_{i=1}^{n} W_{ni} + o(1).$$

Applying the Cauchy-Schwarz inequality and the Chebychev inequality, then using Lemma A.1 and (23), yields that, for some positive constant $C$ and any $\varepsilon > 0$,

$$\sum_{i=1}^{n} \mathrm{E}[W_{ni}^2 \mathbb{I}[|W_{ni}| > \varepsilon]]$$

$$\leq n\sqrt{\mathrm{E}[W_{ni}^4]\mathrm{P}[|W_{ni}| > \varepsilon]} \leq \frac{n}{\varepsilon}\sqrt{\mathrm{E}[W_{ni}^4]\mathrm{Var}[W_{ni}]} = \frac{1}{\varepsilon}\sqrt{n\mathrm{E}[W_{ni}^4]} \leq \frac{C\sqrt{nE_{n4}}}{\varepsilon n V_n}$$

$$= \frac{C\left(\frac{n\kappa_n}{p}\ell'_{f,4}(0) + \frac{3n\kappa_n^2}{2p(p+2)}\ell''_{f,4}(0) + o(n\kappa_n^2)\right)^{1/2}}{\varepsilon\left(\frac{2n\kappa_n^2(p-1)}{p^2(p+2)} + o(n\kappa_n^2)\right)} = \frac{o(\tau_n)}{\varepsilon\left(\frac{2(p-1)\tau_n^2}{p+2} + o(\tau_n^2)\right)} = o(1),$$

where we have used the fact that $\ell'_{f,4}(0) = \ell''_{f,4}(0) = 0$. This shows that $\sum_{i=1}^{n} W_{ni}$ satisfies the classical Levy-Lindeberg condition, hence is asymptotically standard normal (as already mentioned, $W_{ni}$, $i = 1,\ldots,n$, are mutually independent with mean zero and variance $1/n$). For any subsequence $(\exp(\Lambda_{n_m}))$ converging in distribution, the weak limit must then be $\exp(Y)$, with $Y \sim \mathcal{N}(-\eta, 2\eta)$, where we let $\eta := ((p-1)/(p+2))\lim_{m\to\infty} \tau_{n_m}^2$. Mutual contiguity of $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\kappa_n,f}$ and $\mathrm{P}^{(n)}_0$ then follows from the fact that $\mathrm{P}[\exp(Y) = 0] = 0$ and $\mathrm{E}[\exp(Y)] = 1$. $\square$

PROOF OF THEOREM 3.1. As in the proof of Theorem 2.1, all expectations and variances in this proof are taken under the null of uniformity $\mathrm{P}^{(n)}_0$ and all stochastic convergences and $o_{\mathrm{P}}$'s are as $n \to \infty$ under $\mathrm{P}^{(n)}_0$. Recall that we have obtained in the proof of Theorem 2.1 that

$$\Lambda_n = -\frac{(p-1)\tau_n^2}{p+2} + \sqrt{\frac{2(p-1)\tau_n^2}{p+2} + o(1)} \sum_{i=1}^{n} W_{ni} + o(1)$$

$$= -\frac{\tau_n^2}{2}\Gamma_p + |\tau_n|\Gamma_p^{1/2} \sum_{i=1}^{n} W_{ni} + o_{\mathrm{P}}(1),$$

where $\sum_{i=1}^{n} W_{ni} = (1/\sqrt{nV_n}) \sum_{i=1}^{n} (\log f(\kappa_n(\mathbf{X}'_{ni}\boldsymbol{\theta})^2) - E_{n1})$ is asymptotically

standard normal. Since $(\tau_n\Gamma_p^{1/2})$ is $O(1)$, it is therefore sufficient to show that

$$d_n := \mathrm{E}\left[\left(\Gamma_p^{-1/2}\Delta_{\boldsymbol{\theta}}^{(n)} - \frac{\tau_n}{|\tau_n|}\sum_{i=1}^{n}W_{ni}\right)^2\right] = o(1). \tag{26}$$

To do so, write

$$\Gamma_p^{-1/2}\Delta_{\boldsymbol{\theta}}^{(n)} - \frac{\tau_n}{|\tau_n|}\sum_{i=1}^{n}W_{ni}$$

$$= \frac{1}{\sqrt{nV_n}}\sum_{i=1}^{n}\left(p\sqrt{V_n}\Gamma_p^{-1/2}\left((\mathbf{X}_{ni}'\boldsymbol{\theta})^2 - \frac{1}{p}\right) - \frac{\tau_n}{|\tau_n|}\left(\log f(\kappa_n(\mathbf{X}_{ni}'\boldsymbol{\theta})^2) - E_{n1}\right)\right)$$

$$=: \frac{M_n}{\sqrt{nV_n}}. \tag{27}$$

Then using the fact that $\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta})^2] = 1/p$ and $\mathrm{E}[(\mathbf{X}_{n1}'\boldsymbol{\theta})^4] = 3/(p(p+2))$ (see the proof of Lemma A.1) provides $\mathrm{Var}[(\mathbf{X}_{n1}'\boldsymbol{\theta})^2] = \Gamma_p/p^2$, we obtain

$$\mathrm{E}\big[M_n^2\big] = n\mathrm{E}\left[\left(p\sqrt{V_n}\Gamma_p^{-1/2}\left((\mathbf{X}_{n1}'\boldsymbol{\theta})^2 - \frac{1}{p}\right) - \frac{\tau_n}{|\tau_n|}\left(\log f(\kappa_n(\mathbf{X}_{n1}'\boldsymbol{\theta})^2) - E_{n1}\right)\right)^2\right]$$

$$= 2nV_n - 2np\sqrt{V_n}\Gamma_p^{-1/2}\frac{\tau_n}{|\tau_n|}\mathrm{E}\left[\left((\mathbf{X}_{n1}'\boldsymbol{\theta})^2 - \frac{1}{p}\right)\left(\log f(\kappa_n(\mathbf{X}_{n1}'\boldsymbol{\theta})^2) - E_{n1}\right)\right],$$

which, letting $g(x) := x\log f(x)$, rewrites

$$\mathrm{E}\big[M_n^2\big] = 2nV_n - 2np\sqrt{V_n}\Gamma_p^{-1/2}\frac{\tau_n}{|\tau_n|}\left(\mathrm{E}\big[(\mathbf{X}_{ni}'\boldsymbol{\theta})^2\log f(\kappa_n(\mathbf{X}_{n1}'\boldsymbol{\theta})^2)\big] - \frac{E_{n1}}{p}\right)$$

$$= 2nV_n - 2\sqrt{nV_n}\Gamma_p^{-1/2}\frac{\tau_n}{|\tau_n|}\left(\frac{\sqrt{np}}{\kappa_n}\mathrm{E}\big[g(\kappa_n(\mathbf{X}_{n1}'\boldsymbol{\theta})^2)\big] - \sqrt{n}E_{n1}\right). \tag{28}$$

Lemma A.1 provides

$$\mathrm{E}[g(\kappa_n(\mathbf{X}_{n1}'\boldsymbol{\theta})^2)] = c_p\int_{-1}^{1}(1-s^2)^{(p-3)/2}g(\kappa_n s^2)\,ds = \frac{3\kappa_n^2}{p(p+2)} + o(\kappa_n^2).$$

Using this jointly with (21) and (23), it follows from (27)–(28) that

$$d_n = 2 - \frac{2\tau_n\left(\frac{3\sqrt{n}\kappa_n}{p+2} + o\left(\frac{\sqrt{n}\kappa_n}{p}\right)\right) - \left\{\frac{\sqrt{n}\kappa_n}{p} + \frac{3\sqrt{n}\kappa_n^2}{2p(p+2)}(f''(0)-1) + o(\sqrt{n}\kappa_n^2)\right\}}{\sqrt{\Gamma_p}|\tau_n|\left(\frac{n\kappa_n^2\Gamma_p}{p^2} + o(n\kappa_n^2)\right)^{1/2}}$$

$$= 2 - \frac{2\tau_n\left(\Gamma_p\tau_n + o(1)\right)}{\sqrt{\Gamma_p}|\tau_n|\left(\Gamma_p\tau_n^2 + o(1)\right)^{1/2}} = o(1),$$

as was to be shown. □

PROOF OF THEOREM 4.1. First note that by proceeding strictly along the same lines as in the proof of Theorem 3.1, one can show that, for any sequence $(\boldsymbol{\theta}_n)$ in $\mathcal{S}^{p-1}$, and any real sequence $(\tau_n)$ that is $O(1)$ but not $o(1)$,

$$\log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}}{d\mathrm{P}^{(n)}_0} = \tau_n \Delta^{(n)}_{\boldsymbol{\theta}_n} - \frac{\tau_n^2}{2}\Gamma_p + o_\mathrm{P}(1), \tag{29}$$

under $\mathrm{P}^{(n)}_0$, with $\kappa_n = \tau_n p/\sqrt{n}$ (that is, the stochastic second-order expansion in (4) remains true if one replaces the fixed location $\boldsymbol{\theta}$ with an $n$-dependent value $\boldsymbol{\theta}_n$). Therefore, noting that the parameter value $\boldsymbol{\vartheta}_n = (p/\sqrt{n})^{1/2}\boldsymbol{\tau}_n$ in the statement of the theorem corresponds to $\boldsymbol{\theta}_n = \boldsymbol{\tau}_n/\|\boldsymbol{\tau}_n\|$ and $\kappa_n = p\|\boldsymbol{\tau}_n\|^2/\sqrt{n}$, we have

$$\log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}}{d\mathrm{P}^{(n)}_0} = \log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\tau}_n/\|\boldsymbol{\tau}_n\|,p\|\boldsymbol{\tau}_n\|^2/\sqrt{n},f}}{d\mathrm{P}^{(n)}_0} = \|\boldsymbol{\tau}_n\|^2 \Delta^{(n)}_{\boldsymbol{\tau}_n/\|\boldsymbol{\tau}_n\|} - \frac{\|\boldsymbol{\tau}_n\|^4}{2}\Gamma_p + o_\mathrm{P}(1)$$

under $\mathrm{P}^{(n)}_0$. Now, since the central sequence in Theorem 3.1 rewrites

$$\Delta^{(n)}_{\boldsymbol{\theta}} = \frac{p}{\sqrt{n}}\sum_{i=1}^n \left\{ (\mathbf{X}'_{ni}\boldsymbol{\theta})^2 - \frac{1}{p} \right\} = p\sqrt{n}\,\boldsymbol{\theta}'\left(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\right)\boldsymbol{\theta}$$

$$= p\sqrt{n}\,(\boldsymbol{\theta}\otimes\boldsymbol{\theta})'\mathrm{vec}\left(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\right) = (\mathrm{vec}(\boldsymbol{\theta}\boldsymbol{\theta}'))'\boldsymbol{\Delta}^{(n)}, \tag{30}$$

this rewrites

$$\log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\vartheta}_n,f}}{d\mathrm{P}^{(n)}_0} = (\mathrm{vec}(\boldsymbol{\tau}_n\boldsymbol{\tau}'_n))'\boldsymbol{\Delta}^{(n)} - \frac{\|\boldsymbol{\tau}_n\|^4}{2}\Gamma_p + o_\mathrm{P}(1)$$

under $\mathrm{P}^{(n)}_0$. By using the identities $(\mathrm{vec}\,\mathbf{A})'(\mathrm{vec}\,\mathbf{B}) = \mathrm{tr}[\mathbf{A}'\mathbf{B}]$ and $\mathbf{K}_p(\mathrm{vec}\,\mathbf{A}) = \mathrm{vec}(\mathbf{A}')$, straightforward calculations yield $\|\boldsymbol{\tau}_n\|^4\Gamma_p = (\mathrm{vec}(\boldsymbol{\tau}_n\boldsymbol{\tau}'_n))'\mathbf{T}_p\mathrm{vec}(\boldsymbol{\tau}_n\boldsymbol{\tau}'_n)$, which establishes (7). Since the asymptotic normality result readily follows from the multivariate central limit theorem, the theorem is proved. □

PROOF OF THEOREM 4.2. Denoting as $\mathrm{E}^{(n)}_0$ and $\mathrm{Var}^{(n)}_0$ expectation and variance under $\mathrm{P}^{(n)}_0$, one has (see (30))

$$\lim_{n\to\infty} \mathrm{E}^{(n)}_0\big[\Delta^{(n)}_{\boldsymbol{\theta}}\boldsymbol{\Delta}^{(n)}\big] = \lim_{n\to\infty} \mathrm{E}^{(n)}_0\left[\boldsymbol{\Delta}^{(n)}\left(p\sqrt{n}\,(\boldsymbol{\theta}\otimes\boldsymbol{\theta})'\mathrm{vec}\left(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\right)\right)'\right]$$

$$= \lim_{n\to\infty} \mathrm{Var}^{(n)}_0\big[\boldsymbol{\Delta}^{(n)}\big](\boldsymbol{\theta}\otimes\boldsymbol{\theta}) = \boldsymbol{\Gamma}_p(\boldsymbol{\theta}\otimes\boldsymbol{\theta}).$$

Therefore, Le Cam's third lemma implies that, under $P_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$ with $\kappa_n = \tau p/\sqrt{n}$, $(\tau \neq 0)$, $\boldsymbol{\Delta}^{(n)}$ is asymptotically normal with mean vector $\tau \boldsymbol{\Gamma}_p(\boldsymbol{\theta} \otimes \boldsymbol{\theta})$ and covariance matrix $\boldsymbol{\Gamma}_p$, so that, under the same sequence of hypotheses,

$$Q^{(n)} = (\boldsymbol{\Delta}^{(n)})'\boldsymbol{\Gamma}_p^{-}\boldsymbol{\Delta}^{(n)} \xrightarrow{\mathcal{D}} \chi_{d_p}^2(\delta),$$

with

$$\delta = \tau^2(\boldsymbol{\theta} \otimes \boldsymbol{\theta})'\boldsymbol{\Gamma}_p\boldsymbol{\Gamma}_p^{-}\boldsymbol{\Gamma}_p(\boldsymbol{\theta} \otimes \boldsymbol{\theta}) = \frac{p\tau^2}{p+2}\Big(2 - \frac{2}{p}\Big) = \frac{2(p-1)\tau^2}{p+2}.$$

The asymptotic power in (10) readily follows. $\qquad\square$

## Appendix B: Proofs of Theorem 5.1 and Corollary 5.1

PROOF OF THEOREM 5.1. Since $\mathrm{E}[\mathbf{X}_{n1}\mathbf{X}_{n1}'] = (1/p)\mathbf{I}_p$ and $\mathrm{E}[\mathrm{vec}(\mathbf{X}_{n1}\mathbf{X}_{n1}')(\mathrm{vec}(\mathbf{X}_{n1}\mathbf{X}_{n1}'))'] = 1/(p(p+2))(\mathbf{I}_{p^2} + \mathbf{K}_p + \mathbf{J}_p)$ under $P_0^{(n)}$ (see, e.g., Lemma A.2 in Paindaveine and Verdebout, 2016), the multivariate central limit theorem yields

$$\sqrt{n}\,\mathrm{vec}\Big(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\Big) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\mathbf{0}, \frac{1}{p^2}\mathbf{V}_p\Big).$$

Now, by using (30), we obtain that, under $P_0^{(n)}$,

$$\mathrm{E}\Big[\sqrt{n}\,\mathrm{vec}\Big(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\Big)\Delta_{\boldsymbol{\theta}}^{(n)}\Big] = np\,\mathrm{E}\Big[\mathrm{vec}\Big(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\Big)\mathrm{vec}'\Big(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\Big)\Big]\mathrm{vec}(\boldsymbol{\theta}\boldsymbol{\theta}')$$

$$= p\Big(\frac{1}{p^2}\mathbf{V}_p\Big)\mathrm{vec}(\boldsymbol{\theta}\boldsymbol{\theta}') = \frac{2}{p+2}\mathrm{vec}(\boldsymbol{\theta}\boldsymbol{\theta}') - \frac{2}{p(p+2)}\mathrm{vec}(\mathbf{I}_p),$$

so that Le Cam's third lemma shows that, under $P_{\boldsymbol{\theta},\kappa_n,f}^{(n)}$, where $\kappa_n = \tau_n p/\sqrt{n}$ is based on a sequence $(\tau_n)$ converging to $\tau$,

$$\sqrt{n}\,\mathrm{vec}\Big(\mathbf{S}_n - \frac{1}{p}\mathbf{I}_p\Big) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\frac{2\tau}{p+2}\mathrm{vec}(\boldsymbol{\theta}\boldsymbol{\theta}') - \frac{2\tau}{p(p+2)}\mathrm{vec}(\mathbf{I}_p), \frac{1}{p^2}\mathbf{V}_p\Big),$$

which rewrites

$$\sqrt{n}\,\mathrm{vec}\big(\mathbf{S}_n - \boldsymbol{\Sigma}_n\big) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\mathbf{0}, \frac{1}{p^2}\mathbf{V}_p\Big), \tag{31}$$

where

$$\boldsymbol{\Sigma}_n := \Big(\frac{1}{p} - \frac{2\tau}{\sqrt{n}p(p+2)}\Big)\mathbf{I}_p + \frac{2\tau}{\sqrt{n}(p+2)}\boldsymbol{\theta}\boldsymbol{\theta}'$$

$$= \Big(\frac{1}{p} + \frac{2(p-1)\tau}{\sqrt{n}p(p+2)}\Big)\boldsymbol{\theta}\boldsymbol{\theta}' + \Big(\frac{1}{p} - \frac{2\tau}{\sqrt{n}p(p+2)}\Big)(\mathbf{I}_p - \boldsymbol{\theta}\boldsymbol{\theta}').$$

We need to consider the cases (a) $\tau \geq 0$ and (b) $\tau < 0$ separately.

In Case (a), $\mathbf{\Sigma}_n$ has eigenvalues

$$\lambda_{n1} = \frac{1}{p} + \frac{2(p-1)\tau}{\sqrt{n}p(p+2)} \quad \text{and} \quad \lambda_{n2} = \ldots = \lambda_{np} = \frac{1}{p} - \frac{2\tau}{\sqrt{n}p(p+2)}. \quad (32)$$

Fix then arbitrarily $\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p$ such that the $p \times p$ matrix $\mathbf{G}_p := (\boldsymbol{\theta}, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_p)$ is orthogonal. Letting $\mathbf{\Lambda}_n := \mathrm{diag}(\lambda_{n1}, \lambda_{n2}, \ldots, \lambda_{np})$, we have $\mathbf{\Sigma}_n \mathbf{G}_p = \mathbf{G}_p \mathbf{\Lambda}_n$, so that $\mathbf{G}_p$ is an eigenvectors matrix for $\mathbf{\Sigma}_n$. Clearly, $\xi_{n1} := \sqrt{n}p(\hat{\lambda}_{n1} - \lambda_{n1})$ is the largest root of the polynomial $P_{n1}(h) := \det(\sqrt{n}p(\mathbf{S}_n - \lambda_{n1}\mathbf{I}_p) - h\mathbf{I}_p)$, whereas $\xi_{np} := \sqrt{n}p(\hat{\lambda}_{np} - \lambda_{np})$ is the smallest root of $P_{np}(h) := \det(\sqrt{n}p(\mathbf{S}_n - \lambda_{np}\mathbf{I}_p) - h\mathbf{I}_p)$. Letting

$$\mathbf{Z}_n := \sqrt{n}p(\mathbf{G}_p' \mathbf{S}_n \mathbf{G}_p - \mathbf{\Lambda}_n) = \mathbf{G}_p' \sqrt{n}p(\mathbf{S}_n - \mathbf{\Sigma}_n)\mathbf{G}_p, \quad (33)$$

rewrite these polynomials as $P_{nj}(h) = \det(\sqrt{n}p\mathbf{G}_p' \mathbf{S}_n \mathbf{G}_p - \sqrt{n}p\lambda_{nj}\mathbf{I}_p - h\mathbf{I}_p) = \det(\mathbf{Z}_n + \sqrt{n}p(\mathbf{\Lambda}_n - \lambda_{nj}\mathbf{I}_p) - h\mathbf{I}_p)$, $j = 1, p$, which gives

$$P_{n1}(h) = \det(\mathbf{Z}_n + \mathrm{diag}(0, -v_\tau, \ldots, -v_\tau) - h\mathbf{I}_p)$$

and

$$P_{np}(h) = \det(\mathbf{Z}_n + \mathrm{diag}(v_\tau, 0, \ldots, 0) - h\mathbf{I}_p),$$

where we wrote

$$v_\tau := \sqrt{n}p\left[\frac{2(p-1)\tau}{\sqrt{n}p(p+2)} + \frac{2\tau}{\sqrt{n}p(p+2)}\right] = \frac{2p\tau}{p+2}.$$

Note that (31) readily implies that $\mathrm{vec}\,\mathbf{Z}_n = \sqrt{n}p(\mathbf{G}_p \otimes \mathbf{G}_p)'\,\mathrm{vec}(\mathbf{S}_n - \mathbf{\Sigma}_n)$ converges weakly to $\mathrm{vec}\,\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_p)$. It readily follows that $(\xi_{n1}, \xi_{np})'$ converges weakly to $(\xi_1, \xi_p)'$, where $\xi_1$ is the largest root of the polynomial $\det(\mathbf{Z} + \mathrm{diag}(0, -v_\tau, \ldots, -v_\tau) - h\mathbf{I}_p)$ and $\xi_p$ is the smallest root of the polynomial $\det(\mathbf{Z} + \mathrm{diag}(v_\tau, 0, \ldots, 0) - h\mathbf{I}_p)$. This implies that

$$\left( \begin{array}{c} \sqrt{n}(p\hat{\lambda}_{n1} - 1) \\ \sqrt{n}(p\hat{\lambda}_{np} - 1) \end{array} \right) = \left( \begin{array}{c} \xi_{n1} \\ \xi_{np} \end{array} \right) + \frac{2\tau}{p+2}\left( \begin{array}{c} p-1 \\ -1 \end{array} \right)$$

$$\xrightarrow{\mathcal{D}} \left( \begin{array}{c} \xi_1 \\ \xi_p \end{array} \right) + \frac{2\tau}{p+2}\left( \begin{array}{c} p-1 \\ -1 \end{array} \right) =: \left( \begin{array}{c} \eta_1 \\ \eta_p \end{array} \right).$$

Clearly, $\eta_1$ is the largest root of the polynomial $\det(\mathbf{Z} + \mathrm{diag}(0, -v_\tau, \ldots, -v_\tau) + (2(p-1)\tau/(p+2))\mathbf{I}_p - h\mathbf{I}_p)$, that is the largest eigenvalue of $\mathbf{Z} + (2\tau/(p+2))\mathrm{diag}(p-1, -1, \ldots, -1)$, whereas $\eta_p$ is the smallest root of the polynomial $\det(\mathbf{Z} + \mathrm{diag}(v_\tau, 0, \ldots, 0) - (2\tau/(p+2))\mathbf{I}_p - h\mathbf{I}_p)$, that is, the smallest eigenvalue of $\mathbf{Z} + (2\tau/(p+2))\mathrm{diag}(p-1, -1, \ldots, -1)$. This proves the result for $\tau \geq 0$.

We turn to Case (b), for which $\mathbf{\Sigma}_n$ has eigenvalues

$$\lambda_{n1} = \ldots = \lambda_{n,p-1} = \frac{1}{p} - \frac{2\tau}{\sqrt{n}p(p+2)} \quad \text{and} \quad \lambda_{np} = \frac{1}{p} + \frac{2(p-1)\tau}{\sqrt{n}p(p+2)}. \quad (34)$$

Here, we accordingly fix arbitrarily $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{p-1}$ such that the $p \times p$ matrix $\tilde{\mathbf{G}}_p :=$ $(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{p-1}, \boldsymbol{\theta})$ is orthogonal. Still with $\boldsymbol{\Lambda}_n := \text{diag}(\lambda_{n1}, \lambda_{n2}, \ldots, \lambda_{np})$, we have $\boldsymbol{\Sigma}_n \tilde{\mathbf{G}}_p = \tilde{\mathbf{G}}_p \boldsymbol{\Lambda}_n$, so that $\tilde{\mathbf{G}}_p$ is an eigenvectors matrix for $\boldsymbol{\Sigma}_n$. As above, $\xi_{n1} :=$ $\sqrt{n}p(\hat{\lambda}_{n1} - \lambda_{n1})$ is the largest root of the polynomial $P_{n1}(h) := \det(\sqrt{n}p(\mathbf{S}_n - \lambda_{n1}\mathbf{I}_p) - h\mathbf{I}_p)$, whereas $\xi_{np} := \sqrt{n}p(\hat{\lambda}_{np} - \lambda_{np})$ is the smallest root of $P_{np}(h) :=$ $\det(\sqrt{n}p(\mathbf{S}_n - \lambda_{np}\mathbf{I}_p) - h\mathbf{I}_p)$. Letting

$$\mathbf{Z}_n := \sqrt{n}p(\tilde{\mathbf{G}}_p' \mathbf{S}_n \tilde{\mathbf{G}}_p - \boldsymbol{\Lambda}_n) = \tilde{\mathbf{G}}_p' \sqrt{n}p(\mathbf{S}_n - \boldsymbol{\Sigma}_n)\tilde{\mathbf{G}}_p, \tag{35}$$

rewrite these polynomials as $P_{nj}(h) = \det(\sqrt{n}p\tilde{\mathbf{G}}_p' \mathbf{S}_n \tilde{\mathbf{G}}_p - \sqrt{n}p\lambda_{nj}\mathbf{I}_p - h\mathbf{I}_p) =$ $\det(\mathbf{Z}_n + \sqrt{n}p(\boldsymbol{\Lambda}_n - \lambda_{nj}\mathbf{I}_p) - h\mathbf{I}_p)$, $j = 1, p$, which here gives

$$P_{n1}(h) = \det(\mathbf{Z}_n + \text{diag}(0, \ldots, 0, v_\tau) - h\mathbf{I}_p)$$

and

$$P_{np}(h) = \det(\mathbf{Z}_n + \text{diag}(-v_\tau, \ldots, -v_\tau, 0) - h\mathbf{I}_p),$$

with the same $v_\tau$ as above. Of course, we still have that $\text{vec}\,\mathbf{Z}_n = \sqrt{n}p(\tilde{\mathbf{G}}_p \otimes \tilde{\mathbf{G}}_p)' \text{vec}(\mathbf{S}_n - \boldsymbol{\Sigma}_n)$ converges weakly to $\text{vec}\,\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_p)$. It readily follows that $(\xi_{n1}, \xi_{np})'$ converges weakly to $(\xi_1, \xi_p)'$, where $\xi_1$ is the largest root of the polynomial $\det(\mathbf{Z} + \text{diag}(0, \ldots, 0, v_\tau) - h\mathbf{I}_p)$ and $\xi_p$ is the smallest root of the polynomial $\det(\mathbf{Z} + \text{diag}(-v_\tau, \ldots, -v_\tau, 0) - h\mathbf{I}_p)$. This still implies that

$$\begin{pmatrix} \sqrt{n}(p\hat{\lambda}_{n1} - 1) \\ \sqrt{n}(p\hat{\lambda}_{np} - 1) \end{pmatrix} = \begin{pmatrix} \xi_{n1} \\ \xi_{np} \end{pmatrix} + \frac{2\tau}{p+2}\begin{pmatrix} -1 \\ p-1 \end{pmatrix}$$

$$\xrightarrow{\mathcal{D}} \begin{pmatrix} \xi_1 \\ \xi_p \end{pmatrix} + \frac{2\tau}{p+2}\begin{pmatrix} -1 \\ p-1 \end{pmatrix} =: \begin{pmatrix} \eta_1 \\ \eta_p \end{pmatrix}.$$

Clearly, $\eta_1$ is the largest root of the polynomial $\det(\mathbf{Z} + \text{diag}(0, \ldots, 0, v_\tau) - (2\tau/(p+2))\mathbf{I}_p - h\mathbf{I}_p)$, that is the largest eigenvalue of $\mathbf{Z} + (2\tau/(p+2))\text{diag}(-1, \ldots, -1, p-1)$, whereas $\eta_p$ is the smallest root of the polynomial $\det(\mathbf{Z} + \text{diag}(-v_\tau, \ldots, -v_\tau, 0) + (2(p-1)\tau/(p+2))\mathbf{I}_p - h\mathbf{I}_p)$, that is, the smallest eigenvalue of $\mathbf{Z} + (2\tau/(p+2))\text{diag}(-1, \ldots, -1, p-1)$. This establishes the result for $\tau < 0$.    $\square$

PROOF OF COROLLARY 5.1. According to Theorem 5.1, $L_p^{\max}$ is equal in distribution to the first marginal of $\boldsymbol{\ell}^{(p)} = (\ell_1^{(p)}, \ldots, \ell_p^{(p)})'$, where $\ell_1^{(p)} \geq \ldots \geq \ell_p^{(p)}$ are the eigenvalues of $\mathbf{Z} = (Z_{ij})$, with $\text{vec}\,\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_p)$. Note that $Z_{ij} = Z_{ji}$ almost surely for any $1 \leq i < j \leq p$, so $\text{vec}\,\mathbf{Z}$ may not have a density with respect to the Lebesgue measure on $\mathbb{R}^{p^2}$, but $\text{vech}\,\mathbf{Z}$ might in principle have a density with respect to the Lebesgue measure on $\mathbb{R}^{p(p+1)/2}$. If $\text{vech}\,\mathbf{Z}$ indeed has such a density, then Theorem 13.3.1 of Anderson (2003) can be used to obtain the density of $\boldsymbol{\ell}^{(p)}$ from that of $\text{vech}\,\mathbf{Z}$. However, since $\text{tr}[\mathbf{Z}] = (\text{vec}\,\mathbf{I}_p)'(\text{vec}\,\mathbf{Z}) = 0$ almost surely, $\text{vech}\,\mathbf{Z}$ does not admit a density and the eigenvalues $\ell_1^{(p)}, \ldots, \ell_p^{(p)}$ add up to zero almost surely (and thus do not admit a joint density over $\mathbb{R}^p$ either).

We solve this issue by considering a sequence of $p \times p$ random matrices $\mathbf{Z}_{\delta_k}$ ($k = 1, 2, \ldots$), with $\delta_k(> 0)$ converging to zero as $k$ goes to infinity, and such that $\operatorname{vec} \mathbf{Z}_\delta \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_{p,\delta})$, with

$$\mathbf{V}_{p,\delta} := \frac{p}{p+2}(\mathbf{I}_{p^2} + \mathbf{K}_p) - \frac{2-\delta}{p+2}\mathbf{J}_p.$$

Of course, $\mathbf{Z}_{\delta_k}$ converges weakly to $\mathbf{Z}$, so that the continuous mapping theorem ensures that $\boldsymbol{\ell}_{\delta_k}^{(p)} = (\ell_{1\delta_k}^{(p)}, \ldots, \ell_{p\delta_k}^{(p)})'$ (here, $\ell_{1\delta}^{(p)} \geq \ldots \geq \ell_{p\delta}^{(p)}$ are the eigenvalues of $\mathbf{Z}_\delta$) also converges weakly to $\boldsymbol{\ell}^{(p)} = (\ell_1^{(p)}, \ldots, \ell_p^{(p)})'$. Let then $\mathbf{D}_p$ be the $p$-dimensional duplication matrix, that is such that $\mathbf{D}_p(\operatorname{vech} \mathbf{A}) = \operatorname{vec} \mathbf{A}$ for any $p \times p$ symmetric matrix $\mathbf{A}$. Write $\mathbf{W}_\delta := \operatorname{vech} \mathbf{Z}_\delta = \mathbf{D}_p^-(\operatorname{vec} \mathbf{Z}_\delta)$, where $\mathbf{D}_p^- = (\mathbf{D}_p'\mathbf{D}_p)^{-1}\mathbf{D}_p'$ is the Moore-Penrose inverse of $\mathbf{D}_p$. By definition of $\mathbf{Z}_\delta$, the random vector $\mathbf{W}_\delta$ has density

$$\mathbf{w} \mapsto h_\delta(\mathbf{w}) = \frac{a_p}{\sqrt{\delta}} \exp\left(-\frac{1}{2}\mathbf{w}'\left(\mathbf{D}_p^-\left\{\frac{p}{p+2}(\mathbf{I}_{p^2} + \mathbf{K}_p) - \frac{2-\delta}{p+2}\mathbf{J}_p\right\}(\mathbf{D}_p^-)'\right)^{-1}\mathbf{w}\right),$$

where

$$a_p := \frac{(p+2)^{p(p+1)/4}}{2^{(p^2+3p-2)/4}(\pi p)^{p(p+1)/4}}$$

is a normalizing constant. By using the identities $\mathbf{K}_p\mathbf{D}_p = \mathbf{D}_p$ and $\mathbf{D}_p^-(\operatorname{vec} \mathbf{I}_p) = \mathbf{D}_p'(\operatorname{vec} \mathbf{I}_p) = \operatorname{vech} \mathbf{I}_p$, it is easy to check that

$$\left(\mathbf{D}_p^-\left\{\frac{p}{p+2}(\mathbf{I}_{p^2} + \mathbf{K}_p) - \frac{2-\delta}{p+2}\mathbf{J}_p\right\}(\mathbf{D}_p^-)'\right)^{-1}$$

$$= \frac{p+2}{2p}\mathbf{D}_p'\left\{\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_p) + \frac{\frac{2-\delta}{p+2}}{\frac{2p}{p+2} - \frac{p(2-\delta)}{p+2}}\mathbf{J}_p\right\}\mathbf{D}_p$$

$$= \frac{p+2}{2p}\mathbf{D}_p'\left\{\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_p) + \frac{2-\delta}{\delta p}\mathbf{J}_p\right\}\mathbf{D}_p.$$

Using the identities $(\operatorname{vec} \mathbf{A})'(\operatorname{vec} \mathbf{B}) = \operatorname{tr}[\mathbf{A}'\mathbf{B}]$ and $\mathbf{K}_p(\operatorname{vec} \mathbf{A}) = \operatorname{vec}(\mathbf{A}')$, the resulting density for $\mathbf{Z}_\delta$ is therefore

$$\mathbf{z} \mapsto f(\mathbf{z}) = h_\delta(\operatorname{vech} \mathbf{z})$$

$$= \frac{a_p}{\sqrt{\delta}} \exp\left(-\frac{p+2}{4p}(\operatorname{vech} \mathbf{z})'\mathbf{D}_p'\left\{\frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_p) + \frac{2-\delta}{\delta p}\mathbf{J}_p\right\}\mathbf{D}_p(\operatorname{vech} \mathbf{z})\right)$$

$$= \frac{a_p}{\sqrt{\delta}} \exp\left(-\frac{p+2}{4p}(\operatorname{vec} \mathbf{z})'\left\{\mathbf{I}_{p^2} + \frac{2-\delta}{\delta p}\mathbf{J}_p\right\}\operatorname{vec} \mathbf{z}\right)$$

$$= \frac{a_p}{\sqrt{\delta}} \exp\left(-\frac{p+2}{4p}\left\{(\operatorname{tr} \mathbf{z}^2) + \frac{2-\delta}{\delta p}(\operatorname{tr} \mathbf{z})^2\right\}\right).$$

Theorem 13.3.1 from Anderson ([2003](#)) then implies that $\boldsymbol{\ell}_\delta^{(p)} = (\ell_{1\delta}^{(p)}, \ldots, \ell_{p\delta}^{(p)})'$ has density

$$(\ell_1, \ldots, \ell_p)' \mapsto \frac{b_p}{\sqrt{\delta}} \exp\left( -\frac{p+2}{4p} \left\{ \left( \sum_{j=1}^p \ell_j^2 \right) + \frac{2-\delta}{\delta p} \left( \sum_{j=1}^p \ell_j \right)^2 \right\} \right)$$

$$\times \left( \prod_{1 \le k < j \le p} (\ell_k - \ell_j) \right) \mathbb{I}[\ell_1 \ge \ldots \ge \ell_p],$$

with

$$b_p := \frac{(p+2)^{p(p+1)/4}}{2^{(p^2+3p-2)/4} p^{p(p+1)/4} \prod_{j=1}^p \Gamma(\frac{j}{2})}.$$

We now turn to the particular cases (i) $p = 2$ and (ii) $p = 3$ considered in the statement of the corollary. (i) Our general derivation above states that $\boldsymbol{\ell}_\delta^{(2)} = (\ell_{1\delta}^{(2)}, \ell_{2\delta}^{(2)})'$ has density $(\ell_1, \ell_2)' \mapsto \mathcal{I}(\ell_1, \ell_2)\mathbb{I}[\ell_1 \ge \ell_2]$, with

$$\mathcal{I}(\ell_1, \ell_2) := \frac{1}{\sqrt{2\pi\delta}}(\ell_1 - \ell_2)e^{-\frac{1}{2}\{(\ell_1^2+\ell_2^2)+\frac{2-\delta}{2\delta}(\ell_1+\ell_2)^2\}}.$$

Direct computations allow checking that $\mathcal{I}(\ell_1, \ell_2)$ is the derivative of the function

$$\ell_2 \mapsto \frac{\sqrt{2\delta}}{\sqrt{\pi}(2+\delta)} e^{-\frac{1}{2}\{(\ell_1^2+\ell_2^2)+\frac{2-\delta}{2\delta}(\ell_1+\ell_2)^2\}}$$

$$+ \frac{2^{5/2}\ell_1}{(2+\delta)^{3/2}} e^{-\frac{2\ell_1^2}{2+\delta}} \Phi\left( \frac{(2-\delta)\ell_1 + (2+\delta)\ell_2}{\sqrt{2\delta(\delta+2)}} \right).$$

It follows that $\ell_{1\delta}^{(2)}$ has density

$$\ell_1 \mapsto \int_{-\infty}^{\ell_1} \mathcal{I}(\ell_1, \ell_2)\, d\ell_2$$

$$= \left[ \frac{\sqrt{2\delta}}{\sqrt{\pi}(2+\delta)} e^{-\frac{1}{2}\{(\ell_1^2+\ell_2^2)+\frac{2-\delta}{2\delta}(\ell_1+\ell_2)^2\}} \right.$$

$$\left. + \frac{2^{5/2}\ell_1}{(2+\delta)^{3/2}} e^{-\frac{2\ell_1^2}{2+\delta}} \Phi\left( \frac{(2-\delta)\ell_1 + (2+\delta)\ell_2}{\sqrt{2\delta(\delta+2)}} \right) \right]_{-\infty}^{\ell_1}$$

$$= \frac{\sqrt{2\delta}}{\sqrt{\pi}(2+\delta)} e^{-\frac{2\ell_1^2}{\delta}} + \frac{2^{5/2}\ell_1}{(2+\delta)^{3/2}} e^{-\frac{2\ell_1^2}{2+\delta}} \Phi\left( \frac{4\ell_1}{\sqrt{2\delta(\delta+2)}} \right).$$

Taking the limit as $\delta \to 0$, we obtain that $\ell_1^{(2)} (\overset{\mathcal{D}}{=} L_2^{\max})$ has density $\ell_1 \mapsto 2\ell_1 \exp(-\ell_1^2)\mathbb{I}[\ell_1 > 0]$. From Theorem [5.1](#), this is the density of the asymptotic distribution of $T_+^{(n)} = \sqrt{n}(p\hat{\lambda}_{n1} - 1)$. Part (i) of the result then follows from the fact that $T_+^{(n)} = T_-^{(n)} = T_\pm^{(n)}$ almost surely for $p = 2$ (see the discussion below the corollary).

(ii) For $p = 3$, the density of $\boldsymbol{\ell}_\delta^{(3)} = (\ell_{1\delta}^{(3)}, \ell_{2\delta}^{(3)}, \ell_{3\delta}^{(3)})'$ is

$$(\ell_1, \ell_2, \ell_3)' \mapsto \mathcal{J}(\ell_1, \ell_2, \ell_3)\mathbb{I}[\ell_1 \geq \ell_2 \geq \ell_3],$$

with

$$\mathcal{J}(\ell_1, \ell_2, \ell_3) := \frac{125}{216\pi\sqrt{\delta}}(\ell_1 - \ell_2)(\ell_1 - \ell_3)(\ell_2 - \ell_3)e^{-\frac{5}{12}\{(\ell_1^2+\ell_2^2+\ell_3^2)+\frac{2-\delta}{3\delta}(\ell_1+\ell_2+\ell_3)^2\}}.$$

Lengthy yet straightforward computations show that $\mathcal{J}(\ell_1, \ell_2, \ell_3)$ is the derivative of the function $\ell_3 \mapsto \mathcal{K}(\ell_1, \ell_2, \ell_3)$, with

$$\mathcal{K}(\ell_1, \ell_2, \ell_3) :=$$

$$-\frac{25\sqrt{\delta}(\ell_1 - \ell_3)}{48\pi(1+\delta)^2}$$

$$\times\{2(2\ell_1 + 2\ell_3 - \ell_2) + \delta(\ell_1 + \ell_3 - 2\ell_2)\}e^{-\frac{5}{12}\{(\ell_1^2+\ell_2^2+\ell_3^2)+\frac{2-\delta}{3\delta}(\ell_1+\ell_2+\ell_3)^2\}}$$

$$-\frac{5\sqrt{10}(\ell_1 - \ell_3)}{288\sqrt{\pi}(1+\delta)^{5/2}}$$

$$\times\{20(2\ell_1^1 + 5\ell_1\ell_3 + 2\ell_3^2) - 2\delta(5(\ell_1 - \ell_3)^2 - 18) - \delta^2(5(\ell_1 - \ell_3)^2 - 36)\}$$

$$\times e^{-\frac{20(\ell_1^2+\ell_1\ell_3+\ell_3^2)+5\delta(\ell_1-\ell_3)^2}{24(1+\delta)}}\Phi\left(\frac{\sqrt{5}(2(\ell_1 + \ell_2 + \ell_3) - \delta(\ell_1 + \ell_3 - 2\ell_2))}{6\sqrt{\delta(1+\delta)}}\right).$$

Therefore, $(\ell_{1\delta}^{(3)}, \ell_{3\delta}^{(3)})'$ has density

$$(\ell_1, \ell_3) \mapsto \left(\int_{\ell_3}^{\ell_1} \mathcal{J}(\ell_1, \ell_2, \ell_3)\, d\ell_2\right)\mathbb{I}[\ell_1 \geq \ell_3]$$

$$= (\mathcal{K}(\ell_1, \ell_1, \ell_3) - \mathcal{K}(\ell_1, \ell_3, \ell_3))\mathbb{I}[\ell_1 \geq \ell_3],$$

with

$$\mathcal{K}(\ell_1, \ell_1, \ell_3) - \mathcal{K}(\ell_1, \ell_3, \ell_3) =$$

$$-\frac{25\sqrt{\delta}(\ell_1 - \ell_3)}{48\pi(1+\delta)^2}\{2(\ell_1 + 2\ell_3) + \delta(\ell_3 - \ell_1)\}e^{-\frac{5}{12}\{(2\ell_1^2+\ell_3^2)+\frac{2-\delta}{3\delta}(2\ell_1+\ell_3)^2\}}$$

$$+\frac{25\sqrt{\delta}(\ell_1 - \ell_3)}{48\pi(1+\delta)^2}\{2(2\ell_1 + \ell_3) + \delta(\ell_1 - \ell_3)\}e^{-\frac{5}{12}\{(\ell_1^2+2\ell_3^2)+\frac{2-\delta}{3\delta}(\ell_1+2\ell_3)^2\}}$$

$$-\frac{5\sqrt{10}(\ell_1 - \ell_3)}{288\sqrt{\pi}(1+\delta)^{5/2}}e^{-\frac{20(\ell_1^2+\ell_1\ell_3+\ell_3^2)+5\delta(\ell_1-\ell_3)^2}{24(1+\delta)}}$$

$$\times\{20(2\ell_1^1 + 5\ell_1\ell_3 + 2\ell_3^2) - 2\delta(5(\ell_1 - \ell_3)^2 - 18) - \delta^2(5(\ell_1 - \ell_3)^2 - 36)\}$$

$$\times\left[\Phi\left(\frac{\sqrt{5}(2(2\ell_1+\ell_3)-\delta(\ell_3-\ell_1))}{6\sqrt{\delta(1+\delta)}}\right)-\Phi\left(\frac{\sqrt{5}(2(\ell_1+2\ell_3)-\delta(\ell_1-\ell_3))}{6\sqrt{\delta(1+\delta)}}\right)\right].$$

Taking the limit as $\delta \to 0$ shows that the density of $(\ell_1^{(3)}, \ell_3^{(3)})'$ is

$$(\ell_1,\ell_3)\mapsto-\frac{100\sqrt{10}}{288\sqrt{\pi}}(\ell_1-\ell_3)(2\ell_1^2+5\ell_1\ell_3+2\ell_3^2)e^{-\frac{5(\ell_1^2+\ell_1\ell_3+\ell_3^2)}{6}}\mathbb{I}[-2\ell_1<\ell_3\le-\ell_1/2].$$
(36)

From Theorem 5.1, the density of the asymptotic null distribution of $T_+^{(n)} = \sqrt{n}(p\hat{\lambda}_{n1}-1)$ coincides with the density of $\ell_1^{(3)} (\stackrel{\mathcal{D}}{=} L_3^{\max})$. After marginalization in (36), this last density is seen to be

$$\ell_1\mapsto\left\{\sqrt{\frac{5}{2\pi}}e^{-\frac{5\ell_1^2}{2}}+\sqrt{\frac{5}{8\pi}}e^{-\frac{5\ell_1^2}{8}}+\frac{3}{4}\sqrt{\frac{5}{8\pi}}(5\ell_1^2-4)e^{-\frac{5\ell_1^2}{8}}\right\}\mathbb{I}[\ell_1\ge0]$$

$$=\left\{\frac{d}{d\ell_1}\Phi(\sqrt{5}\ell_1)+\frac{d}{d\ell_1}\Phi(\tfrac{\sqrt{5}\ell_1}{2})+3\frac{d}{d\ell_1}\Phi''(\tfrac{\sqrt{5}\ell_1}{2})\right\}\mathbb{I}[\ell_1\ge0],$$

which proves the result for $T_+^{(n)}$, hence also for $T_-^{(n)}$ (recall from the discussion below the corollary that $T_+^{(n)}$ and $T_-^{(n)}$ share the same weak limit in any dimension $p$). Finally, the result for $T_{\pm}^{(n)}$ follows from (36) by using the fact that $T_{\pm}^{(n)}$ converges weakly to $\max(\ell_1^{(3)}, -\ell_3^{(3)})$. $\qquad\square$

## Acknowledgement

## References

ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley, New York. MR0091588

ANDERSON, T. W. and STEPHENS, M. A. (1972). Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika* **43** 613–621. MR0334367

BERNOULLI, D. (1735). Quelle est la cause physique de l'inclinaison des plans des orbites des planètes? In *Recueil des pièces qui ont remporté le prix de*

*l'Académie Royale des Sciences de Paris 1734* 93–122. Académie Royale des Sciences de Paris, Paris. Reprinted in *Daniel Bernoulli, Werke*, Vol. 3, 226–303, Birkhäuser, Basel (1982).

BIJRAL, A. S., BREITENBACH, M. and GRUDIC, G. (2007). Mixture of Watson distributions: a generative model for hyperspherical embeddings. In *Artificial Intelligence and Statistics* 35–42.

BINGHAM, C. (1974). An antipodally symmetric distribution on the sphere. *Ann. Statist.* **2** 1201–1225. MR0397988

CHIKUSE, Y. (2003). *Statistics on Special Manifolds. Lecture Notes in Statistics* **174**. Springer, New York. MR1960435

CUESTA-ALBERTOS, J. A., CUEVAS, A. and FRAIMAN, R. (2009). On projection-based tests for directional and compositional data. *Stat. Comput.* **19** 367–380. MR2565311

CUTTING, C., PAINDAVEINE, D. and VERDEBOUT, T. (2017). Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. *Ann. Statist.* **45** 1024–1058. MR3662447

DAVIES, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **64** 247–254. MR0501523

DAVIES, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **74** 33-43. MR0885917

DAVIES, R. B. (2002). Hypothesis testing when a nuisance parameter is present only under the alternative: Linear model case. *Biometrika* **89** 484–489. MR1913976

DRYDEN, I. L. (2005). Statistical analysis on high-dimensional spheres and shape spaces. *Ann. Statist.* **33** 1643–1665. MR2166558

DÜRRE, A., TYLER, D. E. and VOGEL, D. (2016). On the eigenvalues of the spatial sign covariance matrix in more than two dimensions. *Statist. Probab. Lett.* **111** 80–85. MR3474786

FISHER, N. I., LEWIS, T. and EMBLETON, B. J. (1987). *Statistical analysis of spherical data*. Cambridge Univ. Press press, Cambridge. MR0899958

GARCÍA-PORTUGUÉS, E., NAVARRO-ESTEBAN, P. and CUESTA-ALBERTOS, J. A. (2020). On a projection-based class of uniformity tests on the hypersphere. *Submitted*.

GARCÍA-PORTUGUÉS, E., PAINDAVEINE, D. and VERDEBOUT, T. (2020). On optimal tests for rotational symmetry against new classes of hyperspherical distributions. *J. Amer. Statist. Assoc. To appear*.

GARCÍA-PORTUGUÉS, E. and VERDEBOUT, T. (2018). An overview of uniformity tests on the hypersphere. *arXiv preprint arXiv:1804.00286*.

HALLIN, M. and PAINDAVEINE, D. (2006). Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756. MR2329465

JUPP, P. (2001). Modifications of the Rayleigh and Bingham tests for uniformity of directions. *J. Multivariate Anal.* **77** 1–20. MR1838711

JUPP, P. E. (2008). Data-driven Sobolev tests of uniformity on compact Riemannian manifolds. *Ann. Statist.* **36** 1246–1260. MR2418656

LACOUR, C. and PHAM NGOC, T. M. (2014). Goodness-of-fit test for noisy

directional data. *Bernoulli* **20** 2131–2168. MR3263101

LEY, C. and VERDEBOUT, T. (2017). *Modern Directional Statistics*. CRC Press, Boca Raton. MR3752655

MARDIA, K. V. and JUPP, P. E. (2000). *Directional Statistics*. John Wiley & Sons, Chichester. MR1828667

PAINDAVEINE, D., REMY, J. and VERDEBOUT, T. (2020). Sign tests for weak principal directions. *Bernoulli. To appear.*

PAINDAVEINE, D. and VERDEBOUT, T. (2016). On high-dimensional sign tests. *Bernoulli* **22** 1745–1769. MR3474832

RAYLEIGH, L. (1919). On the problem of random vibrations and random flights in one, two and three dimensions. *Phil. Mag.* **37** 321–346.

SRA, S. and KARP, D. (2013). The multivariate Watson distribution: Maximum-likelihood estimation and other aspects. *J. Multivariate Anal.* **114** 256–269. MR2993885

TYLER, D. E. (1987). Statistical analysis for the angular central Gaussian distribution on the sphere. *Biometrika* **74** 579–589. MR0909362

WATSON, G. S. (1965). Equatorial distributions on a sphere. *Biometrika* **52** 193–201. MR0207115