Statistical Papers (2024) 65:3447–3475
https://doi.org/10.1007/s00362-024-01529-1

REGULAR ARTICLE

A unified approach to goodness-of-fit testing for spherical and hyperspherical data

Bruno Ebner[1] · **Norbert Henze**[1] · **Simos Meintanis**[2,3]

Received: 27 June 2023 / Revised: 5 January 2024 / Published online: 26 February 2024
© The Author(s) 2024

Abstract
We propose a general and relatively simple method to construct goodness-of-fit tests on the sphere and the hypersphere. The method is based on the characterization of probability distributions via their characteristic function, and it leads to test criteria that are convenient regarding applications and consistent against arbitrary deviations from the model under test. We emphasize goodness-of-fit tests for spherical distributions due to their importance in applications and the relative scarcity of available methods.

Keywords Goodness-of-fit test · Characteristic function · Resampling methods · Spherical distribution

Mathematics Subject Classification 62H15 · 62H11 · 62G20

1 Introduction

In this paper, we consider the problem of testing goodness-of-fit for distributions defined on the surface of the sphere $\mathcal{S}^2$ or of the hypersphere $\mathcal{S}^{d-1}$, where $d \geq 3$. In this respect, there is a plethora of such tests for distributions defined on $\mathbb{R}^d$, even in

Norbert Henze and Simos Meintanis have contributed equally to this work.

✉ Bruno Ebner
bruno.ebner@kit.edu

Norbert Henze
norbert.henze@kit.edu

Simos Meintanis
simosmei@econ.uoa.gr

1 Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Englerstr. 2, Karlsruhe 76131, Baden-Württemberg, Germany

2 Department of Economics, National and Kapodistrian University of Athens, Athens, Greece

3 Pure and Applied Analytics, North-West University, Potchefstroom, South Africa

the multivariate case $d > 1$. Besides, also goodness-of-fit tests on the circular domain $\mathcal{S}^1$ is a relatively well-explored area. For the latter case we refer, e.g., to García-Portugués and Verdebout (2018, Chaps. 6 and 7), Jammalamadaka et al. (2019), and Jammalamadaka et al. (2020).

On the other hand, the same problem for data taking values on $\mathcal{S}^{d-1}$, where $d \geq 3$, has been mostly confined to testing for uniformity. Nevertheless, and while the notion of a "non-preferred direction" (Provide proper quotation mark) and hence of testing for uniformity is certainly central to (hyper)spherical data analysis, there are several more flexible distributions which, in fact, often have the uniform as a special case. The reader is referred to the monographs (Ley and Verdebout 2017, Sect. 2.3; Mardia and Jupp 2000, Sect. 9.3), for such non-uniform models for (hyper)spherical data. At the same time, it seems that goodness-of-fit tests specifically tailored to (hyper)spherical laws are scarce, certainly in the case of a composite null hypothesis, where parameters need to be estimated from the data at hand, but also for a completely specified hypothesis with fixed (known) parameter values. For the latter case, the test based on nearest neighbors proposed in Ebner et al. (2018) seems to be one of the few tests available, while to the best of our knowledge, there is much need for research in the case of a composite hypothesis.

In view of these lines, we suggest a procedure for testing the goodness-of-fit of distributions defined on $\mathcal{S}^{d-1}$, where $d \geq 3$.[1] The suggested test is innovative in that it is general-purpose suitable for arbitrary (hyper)spherical distributions, either with fixed or with estimated parameters, and it is straightforwardly applicable, provided that we can easily draw Monte Carlo samples from the distribution under test.

To be specific, let $\|\cdot\|$ be the Euclidean norm in $\mathbb{R}^d$, $d \geq 2$, and write $\mathcal{S}^{d-1} := \{x \in \mathbb{R}^d : \|x\| = 1\}$ for the surface of the unit sphere in $\mathbb{R}^d$. Suppose X is a random (column) vector in $\mathbb{R}^d$ taking on values in $\mathcal{S}^{d-1}$ with a density f with respect to surface measure.

We start our exposition with the simple null hypothesis

$$H_0 : f = f_0, \tag{1}$$

where f_0 is some given density on $\mathcal{S}^{d-1}$, which should be tested against the general alternative H_A that the distributions pertaining to f and f_0 are different.

Most of the approaches for testing goodness-of-fit depend on some discrepancy measure between a given distributional quantity, such as the density, the distribution function or the characteristic function (CF), and a corresponding empirical counterpart, and thereby typically assume that the functional form of this distributional quantity is known under the null hypothesis. Here, we focus on using the CF rather than the density or distribution function in constructing our test. The functional form of the CF, however, is available only for distributions on the real line $\mathbb{R}^1$ and for a few selected cases of multivariate distributions, such as the multivariate normal and the multivariate stable distributions; see Ebner and Henze (2020) and Meintanis et al. (2015).

In order to circumvent this obstacle, which is even more challenging for distributions taking on values on $\mathcal{S}^{d-1}$, in addition to the data at hand we also consider a Monte

[1] The new test also applies to circular distributions ($d = 2$), but herein we emphasize the higher dimensional cases.

Carlo sample from the distribution under test and thereafter formulate our test as a two-sample test that incorporates a pair of empirical quantities obtained from the two samples, i.e., one quantity from the data at hand and the other from the aforementioned Monte Carlo sample.

This idea also applies to the problem of testing the composite null hypothesis

$$H_{0,\vartheta} : f(\cdot) = f_0(\cdot, \vartheta) \quad \text{for some } \vartheta \in \Theta, \tag{2}$$

against general alternatives. Here, $\{f_0(\cdot, \vartheta) : \vartheta \in \Theta\}$ is a given family of densities on $\mathcal{S}^{d-1}$ that is parameterized in terms of $\vartheta \in \Theta$, where $\Theta \subset \mathbb{R}^s$ for some $s \geq 1$. In this case, the Monte Carlo sample should be drawn from the specific member of the family under test that corresponds to an estimate of the parameter ϑ, obtained on the basis of the data at hand.

In this connection, we note that the idea of a goodness-of-fit method that makes use of an artificial sample from the distribution under test seems to date back to Friedman (2003), at least for independent data and simple hypotheses. Recently, Chen and Xia (2023) employ artificial samples for a test of multivariate normality in high dimensions using nearest neighbors, while Arboretti et al. (2021) applies a test procedure for mixed data by means of artificial samples.

The remainder of this work unfolds as follows: In Sect. 2, we formulate our test statistic, and in Sect. 3 we obtain its limit null distribution as well as the corresponding law under fixed alternatives to H_0. In Sect. 4, the validity of a bootstrap resampling scheme necessary for actually carrying out the test for simple hypotheses with fixed parameters is established, while in Sect. 5 a corresponding bootstrap resampling scheme is suggested for the composite hypothesis test statistic. Section 6 contains an extensive Monte Carlo study of the finite-sample behavior of the new tests, including comparisons, while Sect. 7 illustrates real-data applications. The final Sect. 8 provides some discussion. All proofs of the theorems of Sects. 3 and 4 are provided in Appendix 1.

2 Test statistic

Let $\varphi(t) = \mathbb{E}(\mathrm{e}^{\mathrm{i}t^\top X})$, $t \in \mathbb{R}^d$, denote the CF of X, where $\top$ denotes transpose, and $\mathrm{i} = \sqrt{-1}$ stands for the imaginary unit.

We start our exposition with the simple null hypothesis (1) and note that, if X_0 has density f_0 and CF $\varphi_0(t) = \mathbb{E}(\mathrm{e}^{\mathrm{i}t^\top X_0})$, $t \in \mathbb{R}^d$, say, the standard CF-based statistic for testing H_0 versus H_A is given by

$$D_{n,w} = \int |\varphi_n(t) - \varphi_0(t)|^2 w(t)\, \mathrm{d}t. \tag{3}$$

Here,

$$\varphi_n(t) = \frac{1}{n} \sum_{j=1}^{n} \exp(\mathrm{i}t^\top X_j), \ t \in \mathbb{R}^d, \tag{4}$$

is the empirical CF of $X_1, \ldots, X_n$, and $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) copies of X. In (3), the domain of integration as well as the non-negative weight function $w(\cdot)$ will be specified below in a way that $D_{n,w}$ is amenable to computation and so that a test that rejects the null hypothesis H_0 for large values of $D_{n,w}$ is consistent against each alternative. Notice that $D_{n,w}$ is an estimator of the population Fourier-type discrepancy measure

$$D_w(f, f_0) = \int |\varphi(t) - \varphi_0(t)|^2 w(t)\, \mathrm{d}t \tag{5}$$

between f and f_0.

As already mentioned in the Introduction, a test statistic formulated by (3) is only feasible if the null CF $\varphi_0(\cdot)$ is known and if integration, as indicated in $D_{n,w}$, can be performed in general dimension for some suitable weight function $w(\cdot)$. These prerequisites, however, are rarely fulfilled, and in order to circumvent this obstacle we suggest to replace $\varphi_0(t)$ occurring in (3) by

$$\psi_m(t) = \frac{1}{m} \sum_{j=1}^{m} \exp(\mathrm{i}t^\top Y_j), \quad t \in \mathbb{R}^d. \tag{6}$$

Here, $Y_1, \ldots, Y_m$ are i.i.d. random vectors, which are independent of $X_1, \ldots, X_n$ and have the same distribution as X_0. Notice that $\psi_m(t)$ is the empirical CF of $Y_1, \ldots, Y_m$ and thus an estimator of $\varphi_0(t)$. Of course, realizations of $Y_1, \ldots, Y_m$ are generated via Monte Carlo. In the spirit of meanwhile time-honored weighted L^2-statistics (see, e.g., Ebner and Henze (2020) for an overview of weighted L^2-statistics in the context of testing for multivariate normality), we therefore replace $\varphi_0(t)$ in (3) by $\psi_m(t)$ and thus arrive at the test statistic

$$T_{n,m,w} = \frac{mn}{m+n} \int |\varphi_n(t) - \psi_m(t)|^2 w(t)\, \mathrm{d}t. \tag{7}$$

Notice that $T_{n,m,w}$ is reminiscent of a CF-based test for a two-sample problem, one sample being the data $X_1, \ldots, X_n$ at hand, while the other consists of artificial data generated under the null hypothesis H_0. For more details on weighted L^2-statistics for the two-sample problem which are based on the empirical CF, the reader is referred to Meintanis (2005) and Alba Fernández et al. (2008).

In the setting of the composite null hypothesis $H_{0,\vartheta}$ figuring in (2), the test statistic in (7) should be modified accordingly in order to take into account the extra variability introduced by the presence of the unknown parameter ϑ and its estimator. Specifically for the composite null hypothesis we suggest the test statistic

$$\widehat{T}_{n,m,w} = \frac{mn}{m+n} \int |\varphi_n(t) - \widehat{\psi}_m(t)|^2 w(t)\, \mathrm{d}t, \tag{8}$$

with $\varphi_n(t)$ defined in (4) and

$$\widehat{\psi}_m(t) = \frac{1}{m}\sum_{j=1}^{m} \exp(\mathrm{i}t^\top \widehat{Y}_j), \quad t \in \mathbb{R}^d, \tag{9}$$

where $\widehat{Y}_j$, $j \in \{1, \ldots, m\}$, are i.i.d. copies of a random vector having density $f_0(\cdot; \widehat{\vartheta}_n)$. Here, $\widehat{\vartheta}_n := \widehat{\vartheta}_n(X_1, \ldots, X_n)$ is some estimator of ϑ computed from $X_1, \ldots, X_n$.

We close this section by noting that CF-based methods for goodness-of-fit using the notion of artificial samples have been recently proposed in Chen et al. (2022) and Karling et al. (2023), the first within the family of multivariate elliptical distributions and the latter for skewed families of distributions.

3 Asymptotics

In this section, we provide the limit distribution of $T_{n,m,w}$ defined in (7). To be flexible with respect to both the region of integration and to the weight function w, let M be some nonempty Borel set in $\mathbb{R}^d$, and let μ be some finite measure on (the Borel subsets of) M. Thus, M could be $\mathbb{R}^d$ itself, and μ could be absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$, or M could be $\mathcal{S}^{d-1}$, and μ could be absolutely continuous with respect to the spherical measure. Notably, M could also be some countable subset T of $\mathbb{R}^d$, with μ having a probability mass with respect to the counting measure on T.[2]

In this setting, let $X, X_1, X_2, \ldots$ and $Y, Y_1, Y_2, \ldots$ be independent M-valued random vectors that are defined on some common probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Moreover, let $X, X_1, X_2, \ldots$ be i.i.d. with density f with respect to μ and CF $\varphi(t) = \mathbb{E}[\exp(\mathrm{i}t^\top X)]$, $t \in \mathbb{R}^d$. Furthermore, let $Y, Y_1, Y_2, \ldots$ be i.i.d. with density g with respect to μ and CF $\psi(t) = \mathbb{E}[\exp(\mathrm{i}t^\top Y)]$, $t \in \mathbb{R}^d$. Recall that

$$\varphi_n(t) := \frac{1}{n}\sum_{j=1}^{n} \exp(\mathrm{i}t^\top X_j), \qquad \psi_m(t) = \frac{1}{m}\sum_{j=1}^{m} \exp(\mathrm{i}t^\top Y_j), \quad t \in \mathbb{R}^d,$$

are the empirical CF's of $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, respectively. This section tackles the limit distribution of

$$T_{n,m} := \frac{nm}{n+m}\int_M \left|\varphi_n(t) - \psi_m(t)\right|^2 \mu(\mathrm{d}t)$$

as $m, n \to \infty$, under each of the conditions $\varphi = \psi$ and $\varphi \neq \psi$. Putting

$$\mathrm{C}(x) := \int_M \cos\left(t^\top x\right) \mu(\mathrm{d}t), \quad \mathrm{S}(x) := \int_M \sin\left(t^\top x\right) \mu(\mathrm{d}t), \quad x \in \mathbb{R}^d, \tag{10}$$

[2] A counting measure would be sufficient for circular distributions; see for instance Jammalamadaka et al. (2019).

and noting that $\left|\varphi_n(t)-\psi_m(t)\right|^2 = \big(\varphi_n(t)-\psi_m(t)\big)\big(\overline{\varphi_n(t)}-\overline{\psi_m(t)}\big)$, where $\overline{z}$ stands for the complex conjugate of a complex number z, straightforward algebra yields

$$\begin{aligned}\frac{m+n}{mn}T_{n,m} = \frac{1}{n^2}\sum_{j,k=1}^{n}\big(\mathrm{C}(X_k-X_j)+\mathrm{i}\,\mathrm{S}(X_k-X_j)\big)\\ -\frac{1}{mn}\sum_{k=1}^{m}\sum_{j=1}^{n}\big(\mathrm{C}(X_j-Y_k)+\mathrm{i}\,\mathrm{S}(X_j-Y_k)\big)\\ -\frac{1}{mn}\sum_{k=1}^{m}\sum_{j=1}^{n}\big(\mathrm{C}(Y_k-X_j)+\mathrm{i}\,\mathrm{S}(Y_k-X_j)\big)\\ +\frac{1}{m^2}\sum_{j,k=1}^{m}\big(\mathrm{C}(Y_k-Y_j)+\mathrm{i}\,\mathrm{S}(Y_k-Y_j)\big).\end{aligned}$$

Since $\mathrm{S}(-x) = -\mathrm{S}(x)$ and $\mathrm{S}(0_d) = 0$, where 0_d is the origin in $\mathbb{R}^d$, the sum of the imaginary parts vanishes, and we obtain

$$\begin{aligned}\frac{m+n}{mn}T_{n,m} = \frac{1}{n^2}\sum_{j,k=1}^{n}\mathrm{C}(X_j-X_k)-\frac{2}{mn}\sum_{j=1}^{n}\sum_{k=1}^{m}\mathrm{C}(X_j-Y_k)\\ +\frac{1}{m^2}\sum_{j,k=1}^{m}\mathrm{C}(Y_j-Y_k).\end{aligned}\tag{11}$$

Since the computation of the test statistic $T_{m,n}$ requires the evaluation of the integrals $\mathrm{C}(x)$ and $\mathrm{S}(x)$ occurring in (10), it seems to be indispensable to impose some restrictions on the set M and the measure μ in order to render the test feasible. To this end, we assume that M – like $\mathbb{R}^d$, $\mathcal{S}^{d-1}$ or the grid $\mathbb{Z}^d$ – is symmetric with respect to the origin 0_d, i.e., we have $-M = M$, where $-M := \{-x : x \in M\}$. Moreover, we suppose that μ is invariant with respect to the reflection $T(x) := -x$, $x \in \mathbb{R}^d$, i.e., we have $\mu = \mu^T$, where μ^T is the image of μ under T. The condition that the set M should be symmetric with respect to the origin is directly related to the uniqueness of CF's, since this uniqueness only holds if two such functions coincide for each possible argument t. Thus typically we take $M = \mathbb{R}^d$, although simplifications may occur, e.g., when testing for a distribution on the real line, in which case we may take $M = (0,\infty)$ or for a circular distribution for which $M = \mathbb{Z}^1$ suffices. Invariance of the measure μ is another natural requirement since the discrepancy measure $\left|\varphi_n(t)-\psi_m(t)\right|^2$, which is integrated with respect to μ, also remains invariant under sign changes.

By transformation of integrals, we then obtain

$$\mathrm{S}(x) = \int_M \sin(t^\top x)\,\mu(\mathrm{d}t) = -\mathrm{S}(x), \quad x \in \mathbb{R}^d,$$

which, together with $\mathrm{S}(-x) = -\mathrm{S}(x)$, $x \in \mathbb{R}^d$, implies

$$\mathrm{S}(x) = 0, \qquad x \in \mathbb{R}^d. \tag{12}$$

Putting

$$\mathrm{CS}(\xi) := \cos\xi + \sin\xi, \quad \xi \in \mathbb{R},$$

we now use the addition theorems $\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$ and $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$. In view of (12), we thus have

$$\int_M \mathrm{CS}(t^\top X_j)\mathrm{CS}(t^\top X_k)\,\mu(\mathrm{d}t) = \int_M \cos(t^\top(X_j - X_k))\,\mu(\mathrm{d}t) = \mathrm{C}(X_k - X_j),$$

($j, k \in \{1, \ldots, n\}$). In this way, some algebra yields

$$T_{n,m} = \frac{mn}{m+n}\int_M \left(\frac{1}{n}\sum_{j=1}^n \mathrm{CS}(t^\top X_j) - \frac{1}{m}\sum_{k=1}^m \mathrm{CS}(t^\top Y_k)\right)^2 \mu(\mathrm{d}t). \tag{13}$$

Now, writing $\mathcal{B}(M)$ for the σ-field of Borel sets on M, let $\mathbb{H} := \mathrm{L}^2(M, \mathcal{B}(M), \mu)$ be the separable Hilbert space of (equivalence classes of) measurable functions $u : M \to \mathbb{R}$ satisfying $\int_M u^2\,\mathrm{d}\mu < \infty$, equipped with the inner product $\langle u, v\rangle = \int_M uv\,\mathrm{d}\mu$ and the norm $\|u\|_{\mathbb{H}} = \langle u, u\rangle^{1/2}$, $u \in \mathbb{H}$.

Theorem 3.1 *Suppose that $\varphi = \psi$. If M is symmetric with respect to 0_d and μ is invariant with respect to reflections at 0_d, then there is a centered Gaussian random element W of $\mathbb{H}$ having covariance kernel $K(s,t) = \mathbb{E}\big[W(s)W(t)\big]$, where*

$$K(s,t) = \mathrm{Cov}\big(\mathrm{CS}(s^\top X), \mathrm{CS}(t^\top X)\big), \quad s, t \in M, \tag{14}$$

such that $T_{n,m} \xrightarrow{\mathcal{D}} \|W\|_{\mathbb{H}}^2$ as $n, m \to \infty$.

The next result gives the almost sure limit of $(m+n)T_{n,m}/(mn)$ as $m, n \to \infty$.

Theorem 3.2 *Let*

$$\varrho(t) := \mathbb{E}[\mathrm{CS}(t^\top X)], \quad b(t) := \mathbb{E}[\mathrm{CS}(t^\top Y)], \quad t \in M. \tag{15}$$

Under the conditions on M and μ stated in Theorem 3.1, we have

$$\lim_{n,m\to\infty} \frac{m+n}{mn} T_{n,m} = \int_M \big(\varrho(t) - b(t)\big)^2\,\mu(\mathrm{d}t) \quad \mathbb{P}\text{-almost surely}.$$

Let

$$\Delta := \int_M |\varphi(t) - \psi(t)|^2 \, \mu(\mathrm{d}t).$$

It is readily seen that Δ equals the almost sure limit figuring in Theorem 3.2. Thus, Δ is the measure of deviation between the distributions of X and Y, expressed in form of a weighted L^2-distance of the corresponding CFs, and this measure of deviation is estimated by $T_{n,m}$. As the next result shows, the statistic $T_{n,m}$, when suitably normalized, has a normal limit distribution as $n, m \to \infty$ if $\Delta > 0$. To prove this result, we need a technical condition which sometimes is called the *usual limiting regime in two-sample problems* (see, e.g., Henze and Penrose 1999), namely

$$\lim_{m,n\to\infty} \frac{m}{m+n} = \tau \tag{16}$$

for some $\tau \in [0, 1]$.

In contrast to Theorems 3.1 and 3.2, this condition is needed now to assess the asymptotic proportions of the X- sample and the Y-sample. In what follows, put

$$K_1(s,t) = \mathrm{Cov}(\mathrm{CS}(s^\top X), \mathrm{CS}(t^\top X)), \quad K_2(s,t) = \mathrm{Cov}(\mathrm{CS}(s^\top Y), \mathrm{CS}(t^\top Y)), \tag{17}$$

and let

$$K^*(s,t) = \tau K_1(s,t) + (1-\tau) K_2(s,t), \quad s,t \in M. \tag{18}$$

Furthermore, let

$$z(t) = \varrho(t) - b(t), \tag{19}$$

where $\varrho(t)$ and $b(t)$ are given in (15).

Theorem 3.3 *Suppose the standing assumptions on M and μ hold. If $\Delta > 0$, then*

$$\sqrt{\frac{mn}{m+n}} \left(\frac{T_{n,m}}{\frac{mn}{m+n}} - \Delta \right) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \sigma^2)$$

under the limiting regime (16)*, where*

$$\sigma^2 = 4 \int_M \int_M K^*(s,t) z(s) z(t) \, \mu(\mathrm{d}s) \mu(\mathrm{d}t).$$

Remark 3.4 Compared to Chen et al. (2022), who address the problem of composite hypotheses, the limit results of this section are obtained for simple hypotheses without estimated parameters. However, the results obtained herein hold for artificial sample

size $m \neq n$, which is much more general and thus flexible than the case $m = n$ treated in Chen et al. (2022). Moreover, our setting is different from that of elliptical distributions on the classical Euclidean space $\mathbb{R}^d$. In the following section, we suggest a resampling version of the test, and we prove its asymptotic validity.

4 Resampling under a simple hypothesis

Since, under H_0, both the finite-sample and the limit distribution of $T_{n,m}$ as $n, m \to \infty$ depend on the unknown underlying distribution of X_1, we use a bootstrap procedure in order to carry out a test that rejects H_0 for large values of $T_{n,m}$. The bootstrap distribution of $T_{n,m}$ is the conditional distribution of $T_{n,m}$ given the pooled sample $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, and a test of H_0 at nominal level α rejects H_0 if $T_{n,m}$ exceeds the $(1-\alpha)$-quantile of this bootstrap distribution. Since the bootstrap distribution is difficult to compute, it is estimated by a Monte Carlo procedure that repeatedly samples from the empirical distribution of the pooled sample. To be specific, one first computes the observed value $t_{n,m}$ of $T_{n,m}$ based on realizations $x_1, \ldots, x_n, y_1, \ldots, y_m$ of $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, respectively. In a second step, one generates b independent samples by Monte Carlo simulation. Here, for each $j \in \{1, \ldots, b\}$, the j^{th} sample consists of $x_1(j), \ldots, x_n(j), y_1(j), \ldots, y_m(j)$, where these values have been chosen independently of each other with a uniform distribution over $\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ (which means, in particular, that $x_1(j)$ can also be taken from any of $y_1, \ldots, y_m$). For each $j \in \{1, \ldots, b\}$, one then computes the value $t_{n,m}(j) = T_{n,m}(x_1(j), \ldots, x_n(j), y_1(j), \ldots, y_m(j))$ of the test statistic $T_{n,m}$. Letting $c_{n,m;1-\alpha}$ denote the $(1-\alpha)$-quantile of the b values $t_{n,m}(1), \ldots, t_{n,m}(b)$, the hypothesis H_0 is rejected if $t_{n,m} > c_{n,m;1-\alpha}$.

To prove that this bootstrap procedure yields a test of H_0 of asymptotic level α, we use a Hilbert space central limit theorem for triangular arrays (see Kundu et al. 2000). This theorem reads as follows.

Theorem 4.1 *Let $\{e_j : j \geq 1\}$ be a complete orthonormal basis of the separable Hilbert space $\mathbb{H}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|_{\mathbb{H}}$. For each $m \geq 1$, let $X_{m1}, X_{m2}, \ldots, X_{mm}$ be independent $\mathbb{H}$-valued random elements such that $\mathbb{E}(\langle X_{mj}, e_\ell \rangle) = 0$ and $\mathbb{E}\|X_{mj}\|_{\mathbb{H}}^2 < \infty$ for each $j \in \{1, \ldots, m\}$ and each $\ell \geq 1$. Put $S_m = \sum_{j=1}^m X_{mj}$, and let C_m be the covariance operator of S_m. Assume that the following conditions hold:*

(i) $\lim_{m\to\infty} \langle C_m e_k, e_\ell \rangle = a_{k\ell}$ *(say) exists for each $k \geq 1$ and $\ell \geq 1$,*
(ii) $\lim_{m\to\infty} \sum_{k=1}^\infty \langle C_m e_k e_k \rangle = \sum_{k=1}^\infty a_{kk} < \infty$,
(iii) $\lim_{m\to\infty} L_m(\varepsilon, e_k) = 0$ *for each $\varepsilon > 0$ and each $k \geq 1$, where*

$$L_m(\varepsilon, h) := \sum_{j=1}^m \mathbb{E}\big(\langle X_{mj}, h \rangle^2 \mathbf{1}\{|\langle X_{mj}, h \rangle| > \varepsilon\}\big), \quad h \in \mathbb{H}.$$

Then $S_m \xrightarrow{\mathcal{D}} G$ *as* $m \to \infty$, *where* G *is a centered random element of* $\mathbb{H}$ *with covariance operator* C *characterized by* $\langle Ch, e_\ell \rangle = \sum_{j=1}^{\infty} \langle h, e_j \rangle a_{j\ell}$ *for each* $h \in \mathbb{H}$ *and each* $\ell \geq 1$.

To apply Theorem 4.1 in our situation of the Hilbert space $\mathbb{H} := \mathrm{L}^2(M, \mathcal{B}(M), \mu)$ let, in greater generality than considered so far, $X_1^{m,n}, \ldots, X_n^{m,n}, Y_1^{m,n}, \ldots, Y_m^{m,n}$ be i.i.d. M-valued random vectors with common distribution, and put

$$\varrho_{m,n}(t) = \mathbb{E}\big[\mathrm{CS}\big(t^\top X_1^{m,n}\big)\big], \quad t \in M.$$

Moreover, let

$$U_{m,n}(t) := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \big\{\mathrm{CS}\big(t^\top X_j^{m,n}\big) - \varrho_{m,n}(t)\big\},$$

$$V_{m,n}(t) := \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \big\{\mathrm{CS}\big(t^\top Y_k^{m,n}\big) - \varrho_{m,n}(t)\big\}, \quad t \in M$$

(cf. (27)), and write $U_{m,n} = U_{m,n}(\cdot)$, $V_{m,n} = V_{m,n}(\cdot)$ as well as

$$\widetilde{W}_{n,m} := a_{m,n} U_{m,n} - b_{m,n} V_{m,n}, \tag{20}$$

where $a_{m,n}$ and $b_{m,n}$ are defined in (26).

Theorem 4.2 *In the setting given above suppose that, under the limiting regime* (16), *we have* $X_1^{m,n} \xrightarrow{\mathcal{D}} X_\infty$ *for some* M*-valued random vector* X_∞ *with distribution* H_∞. *Put* $\varrho_\infty(t) := \mathbb{E}[\mathrm{CS}(t^\top X_\infty)]$, $t \in M$, *and* $W_\infty(t) := \mathrm{CS}(t^\top X_\infty) - \varrho_\infty(t)$, $t \in M$. *Moreover, let* $W_\infty := W_\infty(\cdot)$ *be the random element of* $\mathbb{H} = L^2(M, \mathcal{B}(M), \mu)$ *with covariance operator* C_∞ *that is associated with the covariance function*

$$c_\infty(s,t) = \mathbb{E}\big[\mathrm{CS}(s^\top X_\infty)\mathrm{CS}(t^\top X_\infty)\big] - \varrho_\infty(s)\varrho_\infty(t), \quad s,t \in M,$$

via

$$\langle C_\infty g, h \rangle = \int_M \int_M c_\infty(s,t)\, g(s) h(t)\, \mu(ds)\mu(dt).$$

We then have $\widetilde{W}_{n,m} \xrightarrow{\mathcal{D}} W_\infty$.

Notice that the test statistic $T_{n,m}$ occurring in (13), computed on the random variables $X_1^{m,n}, \ldots, X_n^{m,n}, Y_1^{m,n}, \ldots, Y_n^{m,n}$, equals $\|\widetilde{W}_{n,m}\|_{\mathbb{H}}^2$, where $\widetilde{W}_{n,m}$ is given in (20). From Theorem 4.2, we thus have the following corollary.

Corollary 4.3 *The limit distribution of the test statistic* $T_{n,m}$ *under the limiting regime* (16) *is that of* $\|W_\infty\|_{\mathbb{H}}{}^2$.

Let F_n and G_m denote the empirical distributions of $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, respectively, and write $H_{n,m} := \frac{n}{m+n} F_n + \frac{m}{m+n} G_m$ for the empirical distribution of the pooled sample $X_1, \ldots, X_n, Y_1, \ldots, Y_m$. By the Glivenko–Cantelli theorem, $H_{n,m}$ converges weakly to $H_\infty := (1-\tau)F + \tau F = F$ with probability one under the limiting regime (16), and thus the bootstrap distribution of $T_{n,m}$ converges almost surely to the distribution of $\|W_\infty\|^2_{\mathbb{H}}$. The latter distribution coincides with the distribution of $\|W\|^2_{\mathbb{H}}$, where W is given in Theorem 3.1. This shows the asymptotic validity of the bootstrap.

Remark 4.4 The resampling bootstrap procedure applied herein may also be replaced by a permutation procedure. The validity of the exhaustive permutation (that includes all possible permutations) may be directly obtained by observing that, under the null hypothesis H_0, the observations $(x_1, \ldots, x_n, y_1, \ldots, y_m)$ are exchangeable. Another potential resampling scheme may be that of weighted bootstrap; see Alba-Fernándes et al. (2017).

5 Resampling under a composite hypothesis

Analogously to $T_{n,m,w}$, the limit null distribution of $\widehat{T}_{n,m,w} = \widehat{T}_{n,w}(x_1, \ldots, x_n, \widehat{y}_1, \ldots, \widehat{y}_m)$ depends (in a very complicated way) on unknown quantities, and hence it cannot be used to compute critical values and actually carry out the test. To this end, we consider a parametric bootstrap procedure involving the test statistic in (8) computed on the basis of bootstrap observations from $f_0(\cdot; \vartheta)$, where the parameter ϑ is replaced by estimators.

More precisely, let $\widehat{T}_{n,m,w,\mathrm{obs}}$ denote the observed value of the test statistic. For given $\alpha \in (0, 1)$, write $t^*_{n,m,\alpha}$ for the upper α-percentile of the bootstrap distribution of $T_{n,m,w}$. We then define the test function as

$$\Xi^*_{n,m} = \begin{cases} 1, \text{ if } \quad T_{n,m,w,\mathrm{obs}} \geq t^*_{n,m,\alpha}, \\ 0, \text{ otherwise.} \end{cases} \tag{21}$$

In practice, the bootstrap distribution of $T_{n,m,w,\mathrm{obs}}$ is approximated as follows:

1. Generate a bootstrap sample $x^*_1, \ldots, x^*_n$ from $f_0(\cdot; \widehat{\vartheta}_n)$.
2. Calculate the estimator $\vartheta^*_n = \vartheta_n(x^*_1, \ldots, x^*_n)$
3. Generate a bootstrap sample $y^*_1, \ldots, y^*_m$ from $f_0(\cdot; \vartheta^*_n)$.
4. Compute $T^*_{n,m,w} = T_{n,m,w}(x^*_1, \ldots, x^*_n, y^*_1, \ldots, y^*_m)$.
5. Repeat steps 1–4 a number of times, say b, and thus obtain $T^*_{n,m,w,1}, \ldots, T^*_{n,m,w,\mathrm{b}}$.

Then we approximate the upper α–percentile $t^*_{n,m,\alpha}$ in (21) of the null distribution of $T_{n,m,w}$ by the upper α-percentile of the empirical distribution of $T^*_{n,m,w,1}, \ldots, T^*_{n,m,w,\mathrm{b}}$.

Although we provide no asymptotic theory for the resampling under a composite hypothesis, our simulations show that the above method works well. Nevertheless, it remains an open problem to formally prove that this bootstrap is asymptotically valid.

6 Simulations

In this section we provide results of competitive Monte Carlo simulations for the case of both a simple and a composite hypothesis. We throughout restrict the simulation to the spherical setting for the dimension $d = 3$ and, for computational feasibility, to the sample size $n = 50$. All simulations are performed using the statistical programming language `R`, see R Core Team (2022). We implement the test statistic by fixing the measure $\mu(\mathrm{d}t)$ to be the density of the zero-mean spherical stable distribution. Then the test may be computed as in Eq. (11) with $\mathrm{C}(x) = \mathrm{e}^{-\gamma\|x\|^{\xi}}$. Specifically, we have

$$T_{n,m,\gamma}^{\{\xi\}} = \frac{1}{m+n} \times \left(\frac{m}{n} \sum_{j,k=1}^{n} \mathrm{e}^{-\gamma\|X_j - X_k\|^{\xi}} - 2 \sum_{j=1}^{n} \sum_{k=1}^{m} \mathrm{e}^{-\gamma\|X_j - Y_k\|^{\xi}} + \frac{n}{m} \sum_{j,k=1}^{m} \mathrm{e}^{-\gamma\|Y_j - Y_k\|^{\xi}} \right), \tag{22}$$

where the notation $T_{n,m,\gamma}^{\{\xi\}}$ is used to emphasize the flexibility of the test with respect to the tuning parameters $(\xi, \gamma) \in (0, 2] \times (0, \infty)$, which are at the disposal of the practitioner and may be employed in order to obtain better power against different alternatives.

Another option, although not yielding a proper measure, is to apply the so–called energy statistic suggested by Székely and Rizzo (2013), which again results from (11) with $\mathrm{C}(x) = -\|x\|^{\xi}$, $\xi \in (0, 2)$. The explicit formula for this statistic is obtained from (22) by replacing $\mathrm{e}^{-\gamma\|\cdot\|^{\xi}}$ with $-\|\cdot\|^{\xi}$. This test statistic will be denoted by $SR_{n,m}^{\{\xi\}}$.

The spherical distributions were generated using the package `Directional`, see Tsagris et al. (2021), and the uniformity tests by the package `sphunif`, see García-Portugués and Verdebout (2020).

6.1 Testing the simple hypothesis of uniformity

We test the hypothesis

$$H_0 : f(\cdot) \equiv 1/|\mathcal{S}^{d-1}|,$$

where $|\mathcal{S}^{d-1}| = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of the $(d-1)$-sphere. Hence we test whether f is the density of the uniform law $\mathcal{U}(\mathcal{S}^{d-1})$, which is a classical testing problem in directional statistics. For an overview of existing procedures we refer to García-Portugués and Verdebout (2018). As competing tests we consider the following procedures:

- The modified Rayleigh test R_n, see Mardia and Jupp (2000, Sect. 10.4.1) based on the mean of the directions,
- The Ebner et al. (2018) test NN_a^J based on volumes of the J nearest neighbor balls with power a,

- The Bingham test B_n, see Bingham (1974), based on the empirical scatter matrix of the sample,
- The Sobolev test, see Giné (1975), G_n based on Sobolev norms, and
- The test of Cuesta-Albertos et al. (2009) CA_n, which is based on random projections that characterize the uniform law,
- The test XM_n in Xu and Matsuda (2020), based on a U-statistic representation of a directional kernel Stein discrepancy with fixed tuning parameter $\kappa = 1$.

Empirical critical values for each testing procedure have been obtained by a Monte Carlo simulation study under H_0 with 100,000 replications each.

We considered the following alternatives to the uniform distribution on $\mathcal{S}^{d-1}$. These alternatives are chosen to simulate different uni-, bi- and trimodal models. For details on the hyperspherical von Mises–Fisher distribution, see Sect. 9.3 in Mardia and Jupp (2000).

- The density of the von Mises–Fisher distribution in dependence of the mean direction $\theta \in \mathcal{S}^{d-1}$ and concentration parameter $\kappa \geq 0$ is given by

$$f(x) = \frac{(\kappa/2)^{d/2-1}}{\Gamma(d/2) I_{d/2-1}(\kappa)} \exp(\kappa x^\top \theta), \quad x \in \mathcal{S}^{d-1}.$$

  Here, $I_{d/2-1}$ is the modified Bessel function of the first kind and order $d/2 - 1$. This class is denoted with $\mathrm{vMF}(\theta, \kappa)$.
- We simulate a mixture of two von Mises–Fisher distributions with different centers by the following procedure. Simulate $U \sim \mathcal{U}(0, 1)$ and, independently, $Y_i \sim \mathrm{vMF}(\theta_i, \kappa_i)$ with corresponding location and concentration parameters for $i = 1, 2$, and choose $p \in (0, 1)$. Then we generate a member X of the random sample according to

$$X = Y_1 \mathbf{1}\{U < p\} + Y_2 \mathbf{1}\{U \geq p\}.$$

  We denote this alternative class by $\mathrm{MMF}((p, 1-p), (\theta_1, \theta_2), (\kappa_1, \kappa_2))$.
- In a similar manner as for the mixture of two vMF distributions, we simulate a mixture of three von Mises–Fisher distributions with different centers, by additionally simulating independently a third random vector $Y_3 \sim \mathrm{vMF}(\theta_3, \kappa_3)$ and generating the member X by

$$X = Y_1 \mathbf{1}\{U < p\} + Y_2 \mathbf{1}\{p \leq U < 2p\} + Y_3 \mathbf{1}\{U \geq 2p\}, \quad p \in (0, 1/2).$$

  We denote this class with $\mathrm{MMF}((p, p, 1-2p), (\theta_1, \theta_2, \theta_3), (\kappa_1, \kappa_2, \kappa_3))$.

In each of the alternatives, we put $\mathbb{1} = (1, \ldots, 1)/\sqrt{d}$, $\mu_1 = (1, 0, \ldots, 0)$, and $\mu_2 = (-1, 1, \ldots, 1)/\sqrt{d}$. The results of the simulation study for the new test applied at $\xi = 2$, as well as the stated competitors are displayed in Tables 1 and 2. In this section, for the sake of simplicity, we suppress the sample size n and the Monte Carlo size m in the notation of the new test and write $T_\gamma^{\{\xi\}}$. As can be seen, the suggested tests perform well in comparison, although they are never the best performing procedures.

This behavior might be explained by the approximation of the true CF under the null hypothesis. To investigate the impact of the sample size m of the simulated data set $Y_1, \ldots, Y_m$, we simulated the empirical power of the test for four vMF distributions and for different values of m, see Fig. 1. Clearly, the choice of m has an impact on the estimation, and larger values of m are desirable, but increasing m leads to longer computation time. Table 3 exhibits the impact of the weighting measure μ and hence of the choice of the function $C(\cdot)$. In this regard, we compare our test with the energy test of Székely and Rizzo (2013), denoted by $SR^{\{\xi\}}$, for different combinations of the tuning parameters (γ, ξ). In terms of power for the uni- and bimodal alternatives considered, the choice of $C(\cdot)$ has nearly no influence on the empirical power, with the exception of the MMF$((0.5, 0.5), (-\mu_1, \mu_1), (2, 2))$ alternative, where $T_\gamma^{\{1\}}$, $T_\gamma^{\{1.5\}}$ and $T_\gamma^{\{2\}}$ outperform the $SR^{\{\xi\}}$-procedures for some values of the tuning parameter γ.

6.2 Testing the fit to the von Mises–Fisher distribution

For the case of a composite hypothesis we consider the hypothesis that the underlying density belongs to the family of von Mises–Fisher distributions vMF(κ, θ), i.e., we test the hypothesis

$$H_0 : f(\cdot) = \frac{(\kappa/2)^{d/2-1}}{\Gamma(d/2) I_{d/2-1}(\kappa)} \exp(\kappa \cdot^\top \theta), \quad \text{for some } \kappa \geq 0 \text{ and } \theta \in \mathcal{S}^{d-1}, \tag{23}$$

against general alternatives. The main difference to subsection 6.1 is that we consider a test to a *family* of distributions, where the parameters are unknown and hence have to be estimated. To test the hypothesis we chose $T_\gamma^{\{2\}}$ in Eq. (22), for different values of the tuning parameter γ, and we implemented the parametric bootstrap procedure from Sect. 5. To approximate the unknown parameters we calculated the maximum likelihood estimates for κ and θ, as proposed in Sect. 10.3.1 of Mardia and Jupp (2000). As far as we know, testing composite hypotheses for spherical or hyperspherical distributions with estimated parameters has not been considered before in the literature. As alternative models we chose the same distributions as described in Subsection 6.1.

In view of the extensive computation time due to the parametric bootstrap procedure, we considered the simulation setting $n = 50$, $m = 200$, a sample size of 500 in the bootstrap algorithm, and 5000 Monte Carlo replications. Throughout the study, we fixed the significance level to 0.05. The results are reported in Table 4. Notably, the novel test maintains the nominal significance level very closely, and its power with respect to bimodal alternatives increases the more these two modes are pronounced.

6.3 Testing the fit to the angular central Gaussian distribution

In this subsection, we consider testing the fit to an angular central Gaussian model, i.e., we test the hypothesis

$$H_0 : f(\cdot) = |\Sigma|^{-1/2} (\cdot^\top \Sigma^{-1} \cdot)^{-d/2}, \quad \text{for some } \Sigma, \tag{24}$$

Table 1 Empirical rejection rates for testing uniformity for the test $T_\gamma^{\{2\}}$ and competitors ($n = 50$, $m = 500$, $\alpha = 0.05$, 10,000 replications)

Alternative	$T_5^{\{2\}}$	$T_2^{\{2\}}$	$T_1^{\{2\}}$	$T_{.5}^{\{2\}}$	$T_{.25}^{\{2\}}$	$T_{.17}^{\{2\}}$	R_n	$NN_{.5}^{15}$	B_n	G_n	CA_n	XM_n
vMF($\mu_1, 0$) = $\mathcal{U}(\mathcal{S}^2)$	5	5	5	5	5	4	5	4	5	5	5	4
vMF($\mu_1, 0.25$)	8	11	11	11	11	11	12	8	5	12	11	7
vMF($\mu_1, 0.5$)	20	31	32	32	33	33	37	21	5	36	32	11
vMF($\mu_1, 0.75$)	44	63	64	66	67	66	72	49	7	71	64	22
vMF($\mu_1, 1$)	71	88	89	89	90	89	93	77	9	92	89	40
MMF((0.25, 0.75), (μ_1, μ_1), (0, 2))	95	99	99	99	99	99	99	96	39	99	99	60
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (2, 2))	78	83	81	80	77	75	77	78	64	85	81	7
MMF((0.5, 0.5), (μ_1, μ_1), (0, 2))	60	75	76	76	76	75	80	61	18	81	75	11
MMF((0.5, 0.5), ($-\mu_1, \mu_1$), (2, 2))	37	23	16	12	8	7	6	48	63	18	13	0
MMF((0.75, 0.25), (μ_1, μ_1), (0, 2))	17	23	23	24	24	23	26	16	8	26	23	3
MMF((0.75, 0.25), ($-\mu_1, \mu_1$), (2, 2))	76	83	82	80	76	75	77	78	63	85	82	8
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 0))	48	54	55	52	50	49	52	48	32	57	51	1
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 1))	35	21	15	12	9	7	7	44	53	17	13	0
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 2))	88	83	76	70	59	53	51	93	93	81	74	4
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 3))	100	100	99	98	91	88	83	100	100	99	98	30
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 4))	100	100	100	100	98	97	94	100	100	100	100	63
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (0, 3))	100	100	100	100	100	100	100	100	86	100	100	87
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (1, 3))	99	100	100	100	99	99	99	99	90	100	100	53
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (2, 3))	98	99	98	98	97	96	96	99	96	99	98	33
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (3, 3))	99	99	98	97	94	92	91	99	99	99	98	28
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (4, 3))	100	100	100	100	100	100	100	100	76	100	100	21
MMF((0.5, 0.5), ($-\mu_1, \mu_1$), (5, 0))	98	99	99	99	99	98	99	98	89	99	99	28

Table 1 continued

Alternative	$T_5^{\{2\}}$	$T_2^{\{2\}}$	$T_1^{\{2\}}$	$T_{.5}^{\{2\}}$	$T_{.25}^{\{2\}}$	$T_{.17}^{\{2\}}$	R_n	$NN_{.5}^{15}$	B_n	G_n	CA_n	XM_n
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (5, 1))	93	89	85	81	72	68	66	94	95	88	84	2
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (5, 2))	96	90	81	68	42	34	25	99	100	85	71	1
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (5, 3))	99	98	93	82	36	22	11	100	100	96	76	1
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (5, 4))	100	100	99	95	46	23	8	100	100	100	89	3
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (0, 3))	85	93	93	93	93	92	94	86	48	95	92	19
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (1, 3))	60	58	53	51	45	42	43	63	64	58	52	0
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (2, 3))	69	53	38	29	17	14	11	80	90	44	31	0
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (3, 3))	89	77	59	42	17	11	6	96	99	65	41	0
MMF((0.5, 0.5), $(-\mu_1, \mu_1)$, (4, 3))	100	100	100	100	100	100	100	100	71	100	100	31

Table 2 Empirical rejection rates for testing uniformity for the test $T_\gamma^{\{2\}}$ and competitors ($n = 50$, $m = 500$, $\alpha = 0.05$, 10,000 replications)

Alternative	$T_5^{\{2\}}$	$T_2^{\{2\}}$	$T_1^{\{2\}}$	$T_{.5}^{\{2\}}$	$T_{.25}^{\{2\}}$	$T_{.17}^{\{2\}}$	R_n	$NN_{.5}^{15}$	B_n	G_n	CA
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 2, 2))	23	22	21	20	19	18	19	25	23	24	20
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 3, 2))	25	21	18	18	16	14	15	29	29	21	17
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 4, 2))	32	24	20	18	16	14	15	38	38	23	17
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 2, 1))	7	6	6	5	5	5	5	8	10	6	5
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 2, 3))	56	55	51	48	44	42	43	58	54	56	49
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 2, 4))	79	76	71	68	61	58	58	82	79	76	69
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (.5, 2, 1))	11	13	13	14	13	12	14	12	9	15	13
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (.25, 1, 2))	44	57	56	57	56	54	60	45	19	62	55
MMF((0.25, 0.25, 0.5), $(-1, \mu_2, \mu_1)$, (2, 0.25, 1))	14	17	18	18	18	17	20	16	9	20	17
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 2, 2))	91	97	97	97	97	97	98	93	44	98	97
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 3, 2))	91	96	97	97	97	96	97	92	46	98	97
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 4, 2))	90	96	96	96	95	95	97	91	47	98	95
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 2, 1))	29	41	42	43	43	42	48	29	8	48	41
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 2, 3))	100	100	100	100	100	100	100	100	91	100	100
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 2, 4))	100	100	100	100	100	100	100	100	100	100	100
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (.5, 2, 1))	37	52	53	54	55	53	59	39	8	59	52
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (.25, 1, 2))	96	99	99	99	99	99	100	97	43	100	99
MMF((0.1, 0.1, 0.8), $(-1, \mu_2, \mu_1)$, (2, 0.25, 1))	36	53	55	55	55	54	60	40	9	60	53

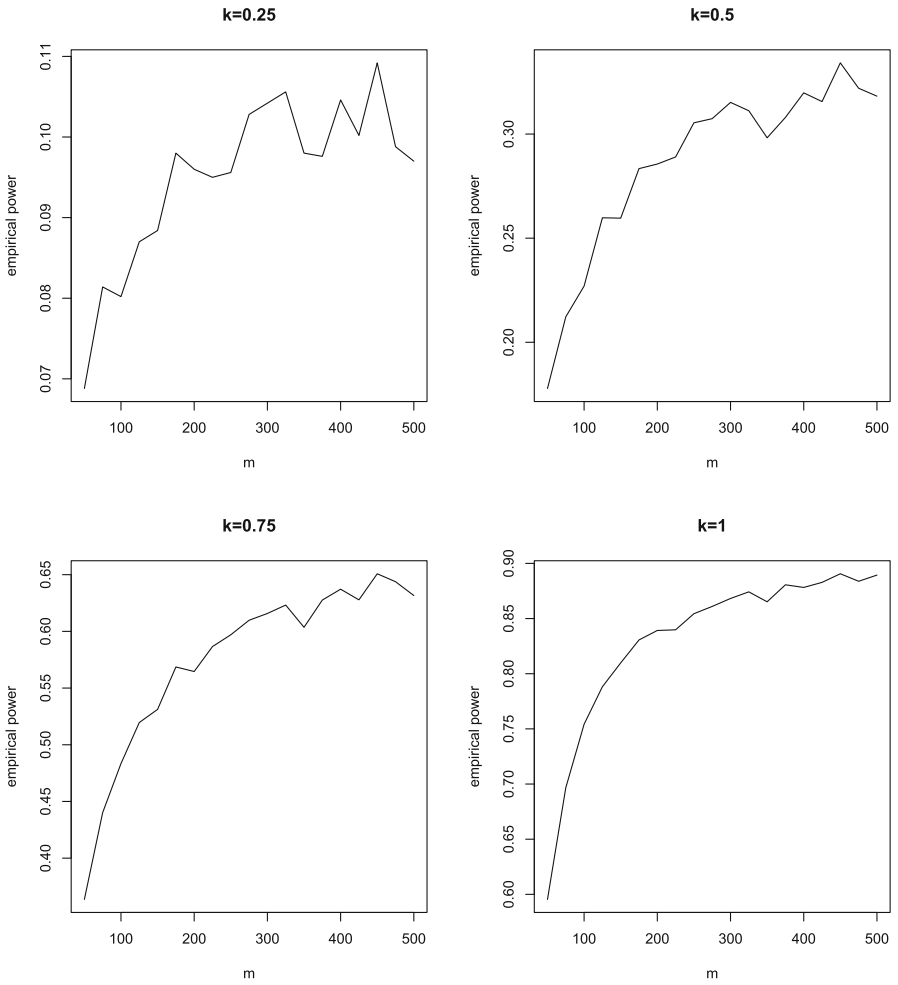


Fig. 1 Simulated empirical rejection rates for $T_1^{\{2\}}$ and different values of $10 \leq m \leq 500$ and concentration parameters $\kappa \in \{0.25, 0.5, 0.75, 1\}$ ($n = 50$, $\alpha = 0.05$, 10,000 replications)

where Σ is a symmetric positive definite $(d \times d)$-parameter matrix, which is identifiable up to multiplication by a positive scalar. For information regarding this model, see Mardia and Jupp (Mardia and Jupp 2000, Sect. 9.4.4), and for a numerical procedure to approximate the maximum likelihood estimator of the unknown parameter matrix Σ, see Tyler (1987). To the best of our knowledge, testing the fit to the angular central Gaussian family has not been considered in the literature.

The simulation parameters match the ones of Sect. 6.2. In complete analogy with previous simulations, we considered $T_\gamma^{\{2\}}$ in Eq. (22), for different values of the tuning parameter γ, and we implemented the parametric bootstrap procedure from Sect. 5. To simulate different models under the null hypothesis, we generated a realization of a random matrix

$$A = \begin{pmatrix} -0.846 & -0.531 & -0.779 \\ 0.609 & -0.096 & 0.761 \\ 0.851 & 0.133 & -0.666 \end{pmatrix}$$

and computed the covariance matrices $\Sigma_\ell = (A^\ell)^\top A^\ell$, $\ell \in \{1, 2, 3, 4\}$, where the power matrix $A^\ell = (a_{ij}^\ell)_{i,j=1,2,3}$ is defined component-by-component, and $\Sigma_0 = \mathrm{diag}(1, 2, 3)$. The corresponding alternatives are denoted by ACG_ℓ, $\ell = 0, \ldots, 4$. Results are presented in Table 5. In this case the bootstrap testing procedure controls the type I error, while performing well for most of the alternatives considered.

7 Real data

We revisit the paleomagnetic data in Scealy and Wood (2019), which is an example of spherical data. Paleomagnetic data consist of observations on the direction of magnetism in either rocks, sediment, or in archeological specimens. These data are measured at various geological points in time and spatial locations. The directions are usually measured as declination and inclination angles based on strike and dip coordinates, see Scealy and Wood (2019) and the references therein for more information. The data considered are taken from the GEOMAGIA50.v3 database, see Brown et al. (2015). For simplicity, we analyze the data provided in the supplementary material of Scealy and Wood (2019). The full data set consists of $n = 1137$ entries (variables are `age`, `dec`, `inc`, `lat`, and `lon`) collected at a single spatial location, which is the Eifel maars (EIF) lakes in Germany with relocated nearby data, for details see Scealy and Wood (2019). The analyzed directions are given by the variables declination D (`dec` defined on $[0°, 360°]$) and inclination I (`inc` defined on $[-90°, 90°]$). They are converted to Cartesian coordinates by $x_1 = \sin(I)$, $x_2 = \cos(I)\cos(D)$, and $x_3 = \cos(I)\sin(D)$, ensuring $x = (x_1, x_2, x_3) \in \mathcal{S}^2$. For a plot of the data, see Fig. 2 (left).

We test the composite hypothesis (23) of a von Mises–Fisher distribution by $T_{n,m,\gamma}^{\{2\}}$ of Eq. (22), for different values of the tuning parameter γ, and we fix $m = 500$ with a bootstrap sample size $b = 1000$. The bootstrap p-values are reported in Table 6. With the exception of the tuning parameter $\gamma = 0.5$, the p-values indicate that we are not able to reject the hypothesis of fit of an underlying von Mises–Fisher distribution at any level. For their analysis, the authors in Scealy and Wood (2019) consider rocks of age 1250 and hence determine a subset of the data of sample size $n = 50$, for a plot see Fig. 2 (right). They propose to use a new spherical model, namely a distribution of Kent type, by applying a transformation to the von Mises–Fisher density. The results of our test of fit to the von Mises–Fisher law for the subset are displayed in the second row of Table 6. For all significance levels and each choice of the tuning parameter, the tests reject the null hypothesis, indicating a poor fit of the von Mises–Fisher family for the subset of the data.

Table 3 Empirical rejection rates for testing uniformity for the test $T_\gamma^{\{\xi\}}$ for $\xi = 1, 1.5$ as well as $SR^{\{\xi\}}$ ($n = 50$, $m = 500$, $\alpha = 0.05$, 10,000 replications)

Alternative	$T_{.1}^{\{1\}}$	$T_{.5}^{\{1\}}$	$T_{.75}^{\{1\}}$	$T_1^{\{1\}}$	$T_2^{\{1\}}$	$T_3^{\{1\}}$	$T_{.1}^{\{1.5\}}$	$T_{.5}^{\{1.5\}}$	$T_{.75}^{\{1.5\}}$	$T_1^{\{1.5\}}$	$T_2^{\{1.5\}}$	$T_3^{\{1.5\}}$	$SR^{\{0.1\}}$	$SR^{\{0.5\}}$	$SR^{\{.75\}}$	$SR^{\{1\}}$	$SR^{\{2\}}$	$SR^{\{3\}}$
vMF($\mu_1, 0$) $= \mathcal{U}(\mathcal{S}^2)$	5	5	5	5	6	5	4	4	5	5	6	5	5	4	5	6	6	5
vMF($\mu_1, 0.25$)	11	11	12	11	11	10	12	12	13	11	11	10	11	11	12	12	12	12
vMF($\mu_1, 0.5$)	33	33	32	30	29	25	33	33	33	32	28	26	31	32	33	33	34	34
vMF($\mu_1, 0.75$)	65	64	64	61	59	52	66	66	65	63	58	53	63	64	66	65	67	66
vMF($\mu_1, 1$)	90	89	89	88	85	80	90	89	90	88	84	79	87	88	89	89	90	90
MMF((0.25, 0.75), (μ_1, μ_1), (0, 2))	99	99	99	99	98	97	99	99	99	99	98	97	99	99	99	99	99	99
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (2, 2))	79	80	82	82	83	81	77	80	82	82	84	82	82	82	81	80	79	76
MMF((0.5, 0.5), (μ_1, μ_1), (0, 2))	77	76	76	75	73	67	77	77	77	76	74	68	75	76	76	77	77	76
MMF((0.5, 0.5), ($-\mu_1, \mu_1$), (2, 2))	10	13	18	21	33	37	7	12	19	22	36	40	20	17	15	13	11	7
MMF((0.75, 0.25), (μ_1, μ_1), (0, 2))	25	25	25	24	23	21	24	25	25	24	23	21	24	25	26	25	25	25
MMF((0.75, 0.25), ($-\mu_1, \mu_1$), (2, 2))	79	80	82	83	84	81	76	80	83	83	85	83	81	82	83	81	81	78
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 0))	51	52	55	55	56	55	48	52	55	55	56	54	52	53	54	53	52	49
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 1))	12	15	18	20	32	35	8	13	17	20	33	36	20	17	15	13	12	8
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 2))	66	74	79	81	88	89	58	71	78	81	89	90	81	79	76	70	65	55
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 3))	96	98	99	99	100	100	91	97	99	99	100	100	99	99	98	97	95	88
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (5, 4))	100	100	100	100	100	100	98	100	100	100	100	100	100	100	100	100	100	97
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (0, 3))	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (1, 3))	99	99	99	99	99	99	99	99	100	100	99	99	100	99	100	99	100	9
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (2, 3))	98	98	99	99	99	99	97	98	99	99	99	99	99	99	99	98	98	97
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (3, 3))	96	97	98	98	99	99	94	98	99	99	99	99	99	99	98	97	96	93
MMF((0.25, 0.75), ($-\mu_1, \mu_1$), (4, 3))	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Table 4 Empirical rejection rates for testing the fit to a von Mises–Fisher distribution ($n = 50$, $m = 200$, $\alpha = 0.05$, 500 bootstrap sample size, 5000 replications)

Alternative	$T_{.1}^{\{2\}}$	$T_{.5}^{\{2\}}$	$T_{.75}^{\{2\}}$	$T_{1}^{\{2\}}$	$T_{2}^{\{2\}}$	$T_{3}^{\{2\}}$
vMF$(\mu_1, 0) = \mathcal{U}(\mathcal{S}^2)$	5	6	6	6	6	6
vMF$(\mu_1, 0.25)$	5	5	6	6	6	6
vMF$(\mu_1, 0.5)$	5	5	6	5	5	6
vMF$(\mu_1, 0.75)$	5	5	6	5	6	6
vMF$(\mu_1, 1)$	5	5	6	6	5	6
MMF$((0.25, 0.75), (\mu_1, \mu_1), (0, 2))$	6	7	7	7	7	6
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (2, 2))$	21	31	30	27	19	13
MMF$((0.5, 0.5), (\mu_1, \mu_1), (0, 2))$	6	7	8	8	7	6
MMF$((0.5, 0.5), (-\mu_1, \mu_1), (2, 2))$	36	49	47	45	30	20
MMF$((0.75, 0.25), (\mu_1, \mu_1), (0, 2))$	5	7	7	7	6	6
MMF$((0.75, 0.25), (-\mu_1, \mu_1), (2, 2))$	21	31	31	29	20	14
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (5, 0))$	16	18	17	16	12	9
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (5, 1))$	36	42	41	38	25	16
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (5, 2))$	59	73	72	69	54	37
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (5, 3))$	80	92	92	91	82	65
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (5, 4))$	92	98	98	98	94	87
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (0, 3))$	10	13	13	12	9	8
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (1, 3))$	23	32	31	29	20	15
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (2, 3))$	41	59	57	55	39	26
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (3, 3))$	60	78	78	75	59	44
MMF$((0.25, 0.75), (-\mu_1, \mu_1), (4, 3))$	25	34	32	30	19	13

As a second parametric family of distributions, we consider the Kent distribution, defined by the density

$$f(x) = \frac{1}{c(A, \kappa)} \exp(\kappa\, x^\top \theta + x^\top A x), \quad x \in \mathcal{S}^{d-1}. \tag{25}$$

Here, $\kappa > 0$ is a concentration parameter and $\theta \in \mathcal{S}^{d-1}$ is the mean direction. Moreover, A is a symmetric $(d \times d)$-matrix with $\mathrm{tr}(A) = 0$ and $A\theta = 0$ that depends on an 'ovality' parameter β, see Mardia and Jupp (2000). Hence we test the hypothesis that the data stems from a density of type (25), where the parameters κ, θ and β are unknown. These parameters have been estimated by the method of maximum-likelihood, and the same bootstrap parameters are applied as above. The bootstrap p-values are reported in Table 6. Interestingly, the full data set is rejected by each of the tests on every significance level. We thus conclude that the Kent distribution is not a suitable model. However, we obtain a different impression for the subset of the data with age fixed to 1250. For this data set, none of the tests can reject the hypothesis of an underlying Kent distribution

Table 5 Empirical rejection rates for testing the fit to an angular central Gaussian distribution ($n = 50$, $m = 200$, $\alpha = 0.05$, 500 bootstrap sample size, 5000 replications)

Alternative	$T_{.1}^{\{2\}}$	$T_{.5}^{\{2\}}$	$T_{1}^{\{2\}}$	$T_{3}^{\{2\}}$
ACG_0	4	4	4	3
ACG_1	5	4	4	4
ACG_2	4	6	4	5
ACG_3	4	3	5	5
ACG_4	6	6	5	4
$\mathrm{vMF}(\mu_1, 0) = \mathcal{U}(\mathcal{S}^2)$	3	3	1	1
$\mathrm{vMF}(\mu_1, 0.25)$	12	15	14	12
$\mathrm{vMF}(\mu_1, 0.5)$	23	34	26	18
$\mathrm{vMF}(\mu_1, 0.75)$	45	74	60	50
$\mathrm{vMF}(\mu_1, 1)$	74	91	91	79
$\mathrm{MMF}((0.25, 0.75), (\mu_1, \mu_1), (0, 2))$	95	100	100	95
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (2, 2))$	41	69	66	47
$\mathrm{MMF}((0.5, 0.5), (\mu_1, \mu_1), (0, 2))$	60	83	72	67
$\mathrm{MMF}((0.5, 0.5), (-\mu_1, \mu_1), (2, 2))$	4	3	3	1
$\mathrm{MMF}((0.75, 0.25), (\mu_1, \mu_1), (0, 2))$	6	17	11	9
$\mathrm{MMF}((0.75, 0.25), (-\mu_1, \mu_1), (2, 2))$	54	59	58	76
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (5, 0))$	25	23	41	46
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (5, 1))$	2	2	2	0
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (5, 2))$	28	44	39	28
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (5, 3))$	46	62	63	49
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (5, 4))$	76	91	84	80
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (0, 3))$	98	100	99	99
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (1, 3))$	95	99	98	98
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (2, 3))$	83	94	92	86
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (3, 3))$	69	86	81	74
$\mathrm{MMF}((0.25, 0.75), (-\mu_1, \mu_1), (4, 3))$	100	100	100	100

The results obtained in this section confirm the statements about the data sets made in Scealy and Wood (2019).

8 Discussion

We have studied goodness-of-fit tests for spherical data. Our tests apply to both simple hypotheses with all parameters assumed known and to composite hypotheses, with parameters estimated from the data at hand. Limit theory is developed under the null hypothesis as well as under alternatives, while the asymptotic validity of a resampling version of the tests is established. The new procedures perform well in finite samples, and they are competitive against other methods, whenever such methods are available.

Table 6 MLE estimates $(\widehat{\kappa}_n, \widehat{\theta}_n)$ of the parameters of the von Mises–Fisher distribution and bootstrap p-values of the test $T^{\{2\}}_{n,m,\gamma}$ applied to the paleomagnetic data set

Model	Sample	$\widehat{\kappa}_n$	$\widehat{\theta}_n$	$\widehat{\beta}_n$	$T^{\{2\}}_{n,m,0.1}$	$T^{\{2\}}_{n,m,0.5}$	$T^{\{2\}}_{n,m,0.75}$	$T^{\{2\}}_{n,m,1}$	$T^{\{2\}}_{n,m,2}$	$T^{\{2\}}_{n,m,3}$
vMF	Full sample	0.131	(0.947, −0.319, 0.039)		0.99	0.024	0.264	1	1	1
	Subsample	0.587	(0.282, −0.930, 0.237)		0	0	0	0	0	0
Kent	Full sample	0.131	(0.947, −0.319, 0.039)	0.070	0	0	0	0	0	0
	Subsample	0.597	(0.282, −0.930, 0.237)	1.043	0.206	0.376	0.287	0.19	0.291	0.419

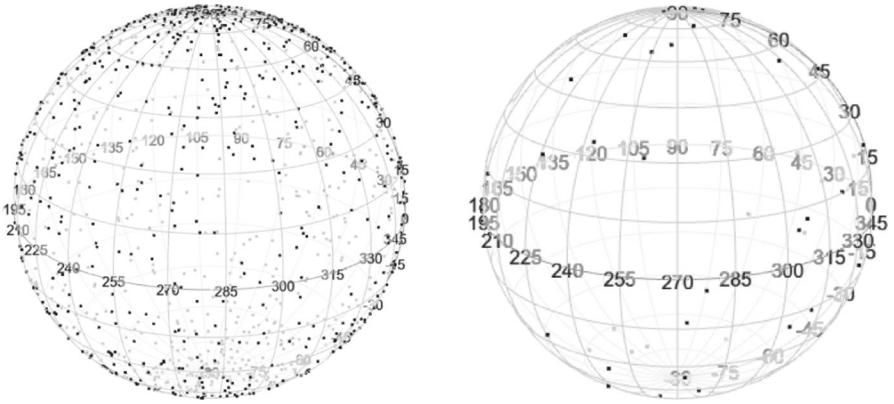


Fig. 2 Full data set ($n = 1137$) of archeomagnetic directions (left) and subsample ($n = 50$) consisting of directions the same age 1250 (right)

An application illustrates the usefulness of the new tests for data-modelling on the sphere. In closing we note that, despite the fact that our study focuses on spherical distributions, the method suggested herein also applies to higher dimensions, provided that one can draw samples from the distribution under test and, in the case of composite hypotheses, one is able to estimate the distributional parameters. This remark notwithstanding, we acknowledge that computational problems become progressively more involved with increasing dimension, and thus the feasibility of our method should be assessed on a case-by-case basis.

Acknowledgements The authors thank an associate editor and two anonymous referees for numerous comments that led to an improvement of the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests No funding was received to assist with the preparation of this manuscript. The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: Proofs

Proof of Theorem 3.1 Let $\varrho(t) := \mathbb{E}[\mathrm{CS}(t^\top X)]$, $t \in \mathbb{R}^d$. Notice that $\varrho(t) = \mathbb{E}[\mathrm{CS}(t^\top Y)]$ since $\varphi = \psi$. Let

$$a_{m,n} = \sqrt{\frac{m}{m+n}}, \quad b_{m,n} = \sqrt{\frac{n}{m+n}}, \tag{26}$$

and put

$$U_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \big\{\mathrm{CS}(t^\top X_j) - \varrho(t)\big\}, \quad V_m(t) = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \big\{\mathrm{CS}(t^\top Y_k) - \varrho(t)\big\}, \tag{27}$$

$t \in M$. Then $U_n = U_n(\cdot)$ and $V_m = V_m(\cdot)$ are centered random elements of $\mathbb{H}$, and by the central limit theorem in separable Hilbert spaces (see, e.g., Theorem 2.7 of Bosq 2000), we have $U_n \xrightarrow{\mathcal{D}} U$ as $n \to \infty$ and $V_m \xrightarrow{\mathcal{D}} V$ as $m \to \infty$, where U and V are centered random elements of $\mathbb{H}$ having covariance kernel K given in (14). From (13), we obtain

$$T_{n,m} = \|W_{n,m}\|_{\mathbb{H}}^2, \tag{28}$$

where $W_{n,m} = a_{m,n} U_n - b_{m,n} V_m$. Since U_n and V_m are independent for each pair (n, m), also U and V are independent, and we have $(U_n, V_m) \xrightarrow{\mathcal{D}} (U, V)$ as $n, m \to \infty$ (see, e.g., Theorem 2.8. of Billingsley 1999). Notice that $n, m \to \infty$ means that, as $s \to \infty$, we have $n = n(s) \to \infty$ and $m = m(s) \to \infty$ in an arbitrary manner. Now, if some subsequence of the bounded sequence $(a_{m,n})$ converges to some $\tau \in [0, 1]$ then, because of the continuous mapping theorem and $a_{m,n}^2 + b_{m,n}^2 = 1$, it follows that $W_{n,m} \xrightarrow{\mathcal{D}} \sqrt{\tau} U - \sqrt{1-\tau} V$ as $n, m \to \infty$. This limit random element has the same distribution as W, irrespective of τ. Consequently, we have $W_{n,m} \xrightarrow{\mathcal{D}} W$ as $n, m \to \infty$ (see, e.g., Theorem 2.6 of Billingsley 1999). In view of (28) and the continuous mapping theorem, the assertion follows. □

Proof of Theorem 3.2 Let

$$A_n(t) = \frac{1}{n} \sum_{j=1}^{n} \mathrm{CS}(t^\top X_j), \quad B_m(t) = \frac{1}{m} \sum_{k=1}^{m} \mathrm{CS}(t^\top Y_k) \tag{29}$$

and, regarded as random elements of $\mathbb{H}$, put $A_n = A_n(\cdot)$ and $B_m = B_m(\cdot)$. Likewise, write $a = \varrho(\cdot)$ and $b = b(\cdot)$ for the degenerate random elements of $\mathbb{H}$ that are the expectations of A_n and B_m, respectively. By the strong law of large numbers in Banach spaces (see, e.g., Hoffmann-Jørgensen and Pisier (1976)), we have $\|A_n - a\|_{\mathbb{H}} \to 0$ as $n \to \infty$ and $\|B_m - b\|_{\mathbb{H}} \to 0$ as $m \to \infty$ $\mathbb{P}$-almost surely. It follows that

$$\frac{m+n}{mn} T_{n,m} = \|A_n - B_m\|_{\mathbb{H}}^2 \to \|a - b\|_{\mathbb{H}}^2 \quad \mathbb{P}\text{-almost surely as } n, m \to \infty.$$

□

Proof of Theorem 3.3 The proof follows the lines of the proof of Theorem 3 of Baringhaus et al. (2017). In view of Theorem 3.2, condition (22) of Baringhaus et al. (2017) holds. Let A_n and B_m as in (29), and put $Z_{n,m} := A_n - B_m$. Furthermore, write $z = z(\cdot)$ for the degenerate random element of $\mathbb{H}$, where $z(t)$ is given in (19), and define

$$c_{m,n} = \sqrt{\frac{mn}{m+n}}.$$

Notice that

$$c_{m,n}\big(Z_{n,m} - z\big) = a_{m,n}\sqrt{n}(A_n - a) - b_{m,n}\sqrt{m}(B_m - b),$$

where $a_{m,n}$ and $b_{m,n}$ are given in (26). By the central limit theorem for $\mathbb{H}$-valued random elements, we have $\sqrt{n}(A_n - a) \xrightarrow{\mathcal{D}} A$ as $n \to \infty$ and $\sqrt{m}(B_m - b) \xrightarrow{\mathcal{D}} B$ as $m \to \infty$, where A and B are independent centered Gaussian random elements of $\mathbb{H}$ with covariance kernels K_1 and K_2, respectively, where K_1 and K_2 are given in (17). In view of (16), the continuous mapping theorem yields $c_{m,n}(Z_{n,m} - z) \xrightarrow{\mathcal{D}} Z$, where $Z := \sqrt{\tau}A - \sqrt{1-\tau}B$ is a centred Gaussian random element of $\mathbb{H}$ having covariance kernel K^* given in (18). Thus, also condition (23) of Baringhaus et al. (2017) holds, and the proof of Theorem 3.3 follows in view of

$$c_{m,n}\left(\frac{T_{n,m}}{c_{n,m}^2} - \Delta\right) = 2\langle c_{m,n}(Z_{n,m} - z), z\rangle + \frac{1}{c_{m,n}}\|c_{m,n}(Z_{n,m} - z)\|_{\mathbb{H}}^2.$$

Notice that the second summand on the right-hand side is $o_{\mathbb{P}}(1)$ in view of the tightness of $(c_{m,n}(Z_{n,m} - z))$, and the first summand converges in distribution to $2\langle Z, z\rangle$, which has the stated normal distribution $\mathrm{N}(0, \sigma^2)$. □

Proof of Theorem 4.2 The proof is similar to that of Theorem 2 of Baringhaus and Kolbe (2015) and will thus only be sketched. Notice that

$$X_{m,n,j} := \frac{1}{\sqrt{n}}\left(\mathrm{CS}(\cdot^\top X_j^{m,n}) - \rho_{m,n}(\cdot)\right), \quad j \in \{1, \ldots, n\}, \tag{30}$$

are i.i.d. centered random elements of the Hilbert space $\mathbb{H} = \mathrm{L}^2(M, \mathcal{B}(M), \mu)$ that, for a fixed complete orthonormal system of $\mathbb{H}$, satisfy $\mathbb{E}(\langle X_{m,n,j}, e_\ell\rangle) = 0$ and $\mathbb{E}\|X_{m,n,j}\|_{\mathbb{H}}^2 < \infty$ for each $j \in \{1, \ldots, n\}$ and each $\ell \geq 1$. The covariance function of the process $U_{m,n} = U_{m,n}(\cdot)$ is given by

$$\begin{aligned} c_{m,n}(s,t) &:= \mathrm{Cov}(U_{m,n}(s), U_{m,n}(t)) \\ &= \mathbb{E}[\mathrm{CS}(s^\top X_1^{m,n})\mathrm{CS}(t^\top X_1^{m,n})] - \varrho_{m,n}(s)\varrho_{m,n}(t), \quad s,t \in M, \end{aligned}$$

and the covariance operator $C_{m,n}$ (say) of $U_{m,n}$ satisfies

$$\langle C_{m,n}g, h\rangle = \int_M \int_M c_{m,n}(s,t)\, g(s)h(t)\, \mu(\mathrm{d}s)\mu(\mathrm{d}t), \quad g, h \in \mathbb{H}.$$

Since the function CS is bounded and continuous, $X_1^{m,n} \xrightarrow{\mathcal{D}} X_\infty$ yields $c_{m,n}(s,t) \to c_\infty(s,t)$, $s,t \in M$. By dominated convergence, we obtain $\langle C_{m,n}g, h\rangle \to \langle C_\infty g, h\rangle$ ($g,\ h \in \mathbb{H}$), which shows that condition (i) of Theorem 4.1 holds. The proof of condition (ii) of Theorem 4.1 follows the reasoning given on p. 603 of Baringhaus and Kolbe (2015) by replacing $J_0(2\sqrt{t}\cdot)$ with $\mathrm{CS}(t^\top \cdot)$, $\mathcal{H}_{m,n}$ with $\varrho_{m,n}$, $\mathcal{H}_\infty$ with ϱ_∞, ν with μ, and the region of integration with M. To prove condition (iii) of Theorem 4.1, notice that, with $X_{m,n,j}$ defined in (30), the fact that $|\mathrm{CS}(\cdot)| \le 2$ and Hölder's inequality give

$$\big|\langle X_{m,n,j}, h\rangle\big| \le \frac{2}{\sqrt{n}}(\mu(M))^{1/2}\|h\|_{\mathbb{H}}^2, \quad h \in \mathbb{H}.$$

Consequently,

$$\sum_{j=1}^{n} \mathbb{E}\big(\langle X_{m,n,j}, h\rangle^2 \mathbf{1}\big\{|\langle X_{m,n,j}, h\rangle| > \varepsilon\big\}\big) \le 4\mu(M)\|h\|_{\mathbb{H}}^2 \mathbb{P}(|\langle X_{m,n,1}, h\rangle| > \varepsilon),$$

and thus also condition (iii) of Theorem 4.1 holds. According to Theorem 4.1, we have $U_{m,n} \xrightarrow{\mathcal{D}} W_\infty$. In the same way, $V_{m,n} \xrightarrow{\mathcal{D}} \widetilde{W}_\infty$, where, due to the independence of $U_{m,n}$ and $V_{m,n}$, $\widetilde{W}_\infty$ is an independent copy of W_∞. In view of (20) and the continuous mapping theorem, it follows that $\widetilde{W}_{n,m} \xrightarrow{\mathcal{D}} \sqrt{\tau}\, W_\infty - \sqrt{1-\tau}\, \widetilde{W}_\infty$. The latter limit has the same distribution as W_∞. □

References

Alba Fernández V, Jiménez Gamero MD, Muñoz García J (2008) A test for the two-sample problem based on empirical characteristic functions. Comput Stat Data Anal 52(7):3730–3748. https://doi.org/10.1016/j.csda.2007.12.013

Alba-Fernández MV, Batsidis A, Jiménez-Gamero M, Jodrá P (2017) A class of tests for the two-sample problem for count data. J Comput Appl Math 318:220–229. https://doi.org/10.1016/j.cam.2016.09.050

Arboretti S, Ceccato R, Salmaso S (2021) Permutation testing for goodness-of-fit and stochastic ordering with multivariate mixed variables. J Stat Comput Simul 91:876–896

Baringhaus L, Kolbe D (2015) Two-sample tests based on empirical Hankel transforms. Stat Pap 56(3):597–617. https://doi.org/10.1007/s00362-014-0599-1

Baringhaus L, Ebner B, Henze N (2017) The limit distribution of weighted L^2-goodness-of-fit statistics under fixed alternatives, with applications. Ann Inst Stat Math 69(5):969–995

Billingsley P (1999) Convergence of probability measures, 2nd edn. Wiley, Chichester

Bingham C (1974) An antipodally symmetric distribution on the sphere. Ann Stat 2(6):1201–1225. https://doi.org/10.1214/aos/1176342874

Bosq D (2000) Linear processes in function spaces. Theory and Applications. Lecture notes in statistics, vol. 149. Springer, New York

Brown MC, Donadini F, Korte M, Nilsson A, Korhonen K, Lodge A, Lengyel SN, Constable CG (2015) Geomagia50.v3: 1. General structure and modifications to the archeological and volcanic database. Earth Planets Space 67(1):83. https://doi.org/10.1186/s40623-015-0232-0

Chen H, Xia Y (2023) A normality test for high-dimensional data based on the nearest neighbor approach. J Am Stat Assoc 118:719–731

Chen F, Jiménez-Gamero M, Meintanis S, Zhu L (2022) A general Monte Carlo method for multivariate goodness-of-fit testing applied to elliptical families. Comput Stat Data Anal. https://doi.org/10.1016/j.csda.2022.1075480167-9473

Cuesta-Albertos JA, Cuevas A, Fraiman R (2009) On projection-based tests for directional and compositional data. Stat Comput 19(4):367–380

Ebner B, Henze N (2020) Tests for multivariate normality–a critical review with emphasis on weighted L^2-statistics. TEST 29(4):845–892. https://doi.org/10.1007/s11749-020-00740-0

Ebner B, Henze N, Yukich JE (2018) Multivariate goodness-of-fit on flat and curved spaces via nearest neighbor distances. J Multivar Anal 165:231–242. https://doi.org/10.1016/j.jmva.2017.12.009

Friedman JH (2003) On multivariate goodness-of-fit and two-sample testing. In: PHYSTAT2003: statistical problems in particle physics, astrophysics and cosmology, pp 311–313

García-Portugués E, Verdebout T (2018) An overview of uniformity tests on the hypersphere. arXiv:1804.00286

García-Portugués E, Verdebout T (2020) Sphunif: uniformity tests on the circle, sphere, and hypersphere. R package version 0.1.0.90002. https://github.com/egarpor/sphunif

Giné E (1975) Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. Ann Stat 3:1243–1266

Henze N, Penrose MD (1999) On the multivariate runs test. Ann Stat 27(1):290–298. https://doi.org/10.1214/aos/1018031112

Hoffmann-Jørgensen J, Pisier G (1976) The law of large numbers and the central limit theorem in Banach spaces. Ann Probab 4(4):587–599. https://doi.org/10.1214/aop/1176996029

Jammalamadaka SR, Jiménez-Gamero MD, Meintanis SG (2019) A class of goodness-of-fit tests for circular distributions based on trigonometric moments. Stat Oper Res Trans (SORT) 43(2):317–336. https://doi.org/10.2436/20.8080.02.90

Jammalamadaka SR, Meintanis S, Verdebout T (2020) On Sobolev tests of uniformity on the circle with an extension to the sphere. Bernoulli 26(3):2226–2252. https://doi.org/10.3150/19-BEJ1191

Karling MJ, Genton MG, Meintanis SG (2023) Goodness-of-fit tests for multivariate skewed distributions based on the characteristic function. Stat Comput. https://doi.org/10.1007/s11222-023-10260-0

Kundu S, Majumdar S, Mukherjee K (2000) Central limit theorems revisited. Stat Probab Lett 47(3):265–275. https://doi.org/10.1016/S0167-7152(99)00164-9

Ley C, Verdebout T (2017) Modern directional statistics. Chapan & Hall/CRC, Boca Raton

Mardia KV, Jupp PE (2000) Directional statistics, 2nd edn. Wiley, Chichester

Meintanis SG (2005) Permutation tests for homogeneity based on the empirical characteristic function. J Nonparam Stat 17(5):583–592. https://doi.org/10.1080/10485250500039494

Meintanis SG, Ngatchou-Wandji J, Taufer E (2015) Goodness-of-fit tests for multivariate stable distributions based on the empirical characteristic function. J Multivar Anal 140:171–192. https://doi.org/10.1016/j.jmva.2015.05.006

R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna (2022). https://www.R-project.org/

Scealy JL, Wood ATA (2019) Scaled von Mises-Fisher distributions and regression models for paleomagnetic directional data. J Am Stat Assoc 114(528):1547–1560. https://doi.org/10.1080/01621459.2019.1585249

Székely GJ, Rizzo ML (2013) Energy statistics: a class of statistics based on distances. J Stat Plan Inference 143(8):1249–1272. https://doi.org/10.1016/j.jspi.2013.03.018

Tsagris M, Athineou G, Sajib A, Amson E, Waldstein MJ (2021) Directional: A Collection of R Functions for Directional Data Analysis. R package version 4.9. https://CRAN.R-project.org/package=Directional

Tyler DE (1987) Statistical analysis for the angular central Gaussian distribution on the sphere. Biometrika 74(3):579–589. https://doi.org/10.2307/2336697

Xu W, Matsuda T (2020) A Stein goodness-of-fit test for directional distributions. In: Proceedings of the 23rd international conference on artificial intelligence and statistics

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.