Manuscript Details

Manuscript number	JMVA_2019_11_R2
Title	Measures of goodness of fit obtained by almost-canonical transformations on Riemannian manifolds
Article type	Research Paper

Abstract

The standard method of transforming a continuous distribution on the line to the uniform distribution on $[0, 1]$ is the probability integral transform. Analogous transforms exist on compact Riemannian manifolds, $\mathcal{X}$, in that for each distribution with continuous positive density on $\mathcal{X}$, there is a continuous mapping of $\mathcal{X}$ to itself that transforms the distribution into the uniform distribution. In general, this mapping is far from unique. This paper introduces the construction of an almost canonical version of such a probability integral transform. The construction is extended to shape spaces, Cartan--Hadamard manifolds, and simplices. The probability integral transform is used to derive tests of goodness of fit from tests of uniformity. Illustrative examples of these tests of goodness of fit are given involving (i) Fisher distributions on S^2, (ii) isotropic Mardia--Dryden distributions on the shape space Σ^5_2. Their behaviour is investigated by simulation.

Keywords	Cartan--Hadamard manifold; compositional data; directional statistics ; exponential map; probability integral transform; shape space; simplex

Corresponding Author	Peter Jupp
Corresponding Author's Institution	University of St. Andrews
Order of Authors	Peter Jupp, Alfred Kume

Suggested reviewers	Christophe Ley, John Kent, Stephan Huckemann

Submission Files Included in this PDF

File Name [File Type]

JMVAcover221019.docx [Cover Letter]

PEJAK_JMVAresponse221019.pdf [Response to Reviewers]

PEJAK221019.pdf [Manuscript File]

Submission Files Not Included in this PDF

File Name [File Type]

JMVA221019.zip [LaTeX Source File]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Research Data Related to this Submission

There are no linked research data sets for this submission. The following reason is given:
No data was used for the research described in the article

University of St Andrews

Emeritus Professor P. E. Jupp
School of Mathematics and Statistics

22 October 2019

Dear Dietrich,

JMVA_2019_11_R1

Please find attached a new version of this manuscript, which Alfred Kume and I have revised in the light of the comments of the Associate Editor and the reviewers. We have also taken the opportunity to make a few minor typographical changes. Changes from the first revision have been highlighted in red.

Best wishes,

Peter

Mathematical Institute, North Haugh, St Andrews, Fife KY16 9SS, Scotland
Tel: 01334 476161 Ext. 3704 Direct Line: 01334 463704 Fax: 01334 463748 email: pej@st-andrews.ac.uk

JMVA-2019-11-R1

Measures of goodness of fit obtained by almost-canonical transformations on Riemannian manifolds

Response to Associate Editor and reviewers

We are grateful to the reviewers for their comments on the first revision and have made further changes in the light of these. Changes from the first revision have been highlighted in red in this revision. Our responses to the reviewers' comments are given below.

Reviewer 1

Regarding my comments on the previous version,

For (2), your comment should be reflected in the paper. Since your remark 2 says your construction can be used in the case of two arbitrary distributions, your comment should not be restricted to the case where one of the distributions is uniform.

We have re-written Remark 2 to clarify that it refers to all appropriate transformations, not just the almost-canonical transformations introduced in Section 2.

Partly in further response to your earlier comment, we have inserted new text into the first paragraph of Section 3 to emphasise that tangent spaces do not play an intrinsic role in the tests ϕ^*T; they arise merely in the construction of some suitable transformations ϕ.

For (3), since all your examples are '1-dimensional', is the relevant ψ_{d-1} just the identity map (when one of the distributions is uniform)? If so, there is no application of Proposition 2 involved. Am I missing anything here? A non-trivial example demonstrating how Proposition 2 is used should be included.

Yes, for distributions with symmetry of the type described just after Remark 2, ψ_{d-1} is the identity. Providing details of an example without such symmetry would involve more work than we can manage at present.

For (4), the (r, u) coordinates in Proposition 2 are polar coordinates on the tangent space, rather than the Riemannian normal coordinates on the manifold. You should perhaps reconsider my comment.

The $(r, \mathbf{u})$ coordinates are the polar form of the usual Riemannian normal coordinates. We have clarified this near equation (10) and have introduced the term 'polar Riemannian normal coordinates'.

For (5), since the aim of the paper is to 'propose and explore a new method as stated in your response to the AE, a comparison with existing methods should be included as it becomes more relevant.

Our aims are (i) to show how almost-canonical versions of probability transforms can be defined in a fairly general context, (ii) to show how these can be used to transform tests, e.g. of uniformity, into tests of goodness of fit, (iii) to explore some basic properties of these tests. There are many established tests of uniformity and several classes of tests of goodness of fit. Detailed comparison of the behaviour of the latter with tests obtained by the general machinery that we describe here would be a large undertaking, which we believe to be best left to future publications.

Reviewer 2

...

One confusion remains on my side that may have an easy fix.

The authors now (p. 4) define the mean as the extrinsic (not the intrinsic with respect to spherical distance) Fréchet median. Btw: Speaking for potential readers, I suggest to reverse engineer: First introduce the extrinsic Fréchet median, then talk about its properties.

Further, I believe the following uniformity statements (e.g. for μ_0) are only true for _intrinsic_ Fréchet medians. So, maybe, it's an intrinsic Fréchet median after all?

Yes, it should indeed be the Riemannian (intrinsic) Fréchet median. The use of the definition of the extrinsic (instead of Riemannian) Fréchet median was an error. We have inserted the correct definitions just before equation (4) and just before Proposition 2.

Measures of goodness of fit obtained by almost-canonical transformations on Riemannian manifolds

P.E. Jupp[a,*], A. Kume[b]

[a]School of Mathematics and Statistics, University of St Andrews, St Andrews KY16 9SS, UK
[b]School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7FS, UK

Abstract

The standard method of transforming a continuous distribution on the line to the uniform distribution on $[0, 1]$ is the probability integral transform. Analogous transforms exist on compact Riemannian manifolds, $\mathcal{X}$, in that for each distribution with continuous positive density on $\mathcal{X}$, there is a continuous mapping of $\mathcal{X}$ to itself that transforms the distribution into the uniform distribution. In general, this mapping is far from unique. This paper introduces the construction of an almost-canonical version of such a probability integral transform. The construction is extended to shape spaces, Cartan–Hadamard manifolds, and simplices.

The probability integral transform is used to derive tests of goodness of fit from tests of uniformity. Illustrative examples of these tests of goodness of fit are given involving (i) Fisher distributions on S^2, (ii) isotropic Mardia–Dryden distributions on the shape space Σ_2^5. Their behaviour is investigated by simulation.

Keywords: Cartan–Hadamard manifold, Compositional data, Directional statistics, Exponential map, Probability integral transform, Shape space, Simplex
2010 MSC: 62F03

1. Introduction

Directional statistics, shape analysis and compositional data analysis are concerned with probability distributions on Riemannian manifolds, shape spaces and simplices, respectively. The aim of this paper is to introduce and explore a canonical method of constructing transformations from such manifolds, $\mathcal{X}$, to certain associated manifolds, $\mathcal{Y}$, that send (almost) arbitrary continuous distributions on $\mathcal{X}$ into standard distributions on $\mathcal{Y}$. More precisely, $\mathcal{Y}$ is $\mathcal{X}$ itself, or a tangent space to $\mathcal{X}$, or a star-shaped open subset of a tangent space. Given a basepoint x in $\mathcal{X}$ and a standard continuous distribution, ν, on $\mathcal{Y}$, for any continuous distribution, μ, on $\mathcal{X}$ with positive density, we construct a function $\phi : \mathcal{X} \to \mathcal{Y}$ that is an almost-diffeomorphism (a diffeomorphism on the complements of some null sets in $\mathcal{X}$ and $\mathcal{Y}$) that sends μ to ν. Under mild conditions on uniqueness of medians of μ and of some distributions drived from it, ϕ as constructed here is canonical (in that any two versions differ only on a null set). These almost-diffeomorphisms, ϕ, are used to obtain tests of goodness of fit to μ from tests of goodness of fit to ν. If $\mathcal{X}$ is a compact Riemannian manifold then we can take $\mathcal{Y} = \mathcal{X}$, ν as the uniform distribution, and ϕ can be regarded as a form of probability integral transformation. On compact manifolds our tests of goodness of fit complement the general Wald-type tests of Beran [1], the score tests of Boulerice and Ducharme [3] and the Sobolev tests of Jupp [13], as well as the more specific tests in [17], [23], [5], [2] (see [22, Section 12.3]), [11, Section 4.2], [25] and [14, Section 4.4].

If $\mathcal{X}$ is a connected compact Riemannian manifold of dimension at least 2 and $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are sets of disjoint points in $\mathcal{X}$ then there is a diffeomorphism $\psi : \mathcal{X} \to \mathcal{X}$ that preserves the uniform distribution and such that $\psi(x_i) = y_i$ for $i = 1, \ldots, n$. (This follows from a straight-forward argument involving rotations on small embedded discs, as in Section 2.1.) Thus 'all data sets of size n are equivalent' up to diffeomorphism. The inference

*Corresponding author
Email addresses: pej@st-andrews.ac.uk (P.E. Jupp), A.Kume@kent.ac.uk (A. Kume)

obtained from applying almost any test of uniformity to $\phi(x_1), \ldots, \phi(x_n)$ is usually different from that obtained from applying it to $(\phi_\circ\psi)(x_1), \ldots, (\phi_\circ\psi)(x_n)$. For well-defined inference it is therefore necessary to use an agreed almost-diffeomorphism ϕ.

Although there is no unique canonical choice of the transformations ϕ, we introduce in Section 2 a sensible construction of 'almost-canonical' transformations, first for spheres and then for compact Riemannian manifolds, shape spaces, Cartan–Hadamard manifolds and simplices. Section 3 shows how these transformations send general tests of uniformity (or of goodness of fit to some standard distribution) into general tests of goodness of fit. The behaviour of these goodness-of-fit tests is illustrated in Section 4 by some simulation studies on the sphere, S^2, and on the shape space, Σ_2^5.

2. Almost-canonical transformations

2.1. Spheres

Let X be a random variable on the unit circle and suppose that an orientation and an initial direction on the circle have been chosen. Then the *probability integral transformation* of the distribution is the transformation of the circle which sends θ to U, where $U = 2\pi \Pr(0 < X \leq \theta)$. If the distribution of X is continuous then U is distributed uniformly on the circle. Thus the probability integral transformation can be used to transform any test of uniformity into a corresponding test of goodness of fit (see [22, Section 6.4]). For continuous distributions (with positive density) μ, on S^{p-1}, the unit sphere in $\mathbb{R}^p$, with $p > 2$, there are analogues $\phi : S^{p-1} \to S^{p-1}$ of the probability integral transformation that transform μ into the uniform distribution, ν. Such ϕ are far from unique, since if $\psi : S^{p-1} \to S^{p-1}$ preserves ν then the composite function $\psi \circ \phi : S^{p-1} \to S^{p-1}$ also transforms μ into ν. Homeomorphisms ψ that preserve ν can be constructed from any embeddings $\gamma : D^{p-1} \to S^{p-1}$ that map the uniform distribution on the $(p-1)$-dimensional disc, D^{p-1}, to the uniform distribution on $\gamma(D^{p-1})$, together with functions $t \mapsto \mathbf{U}_t$ from $[0, 1]$ to the rotation group $SO(p-1)$ with $\mathbf{U}_t = \mathbf{I}_3$ for t near 0 or 1. Then ψ is the identity outside $\gamma(D^{p-1})$ and is given by $\psi\{\gamma(r, \theta)\} = \gamma\{r, \mathbf{U}_r(\theta)\}$ on $\gamma(D^{p-1})$, where (r, θ) are polar coordinates on D^{p-1}.

Our construction of almost-canonical versions of the probability integral transformation ϕ on S^{p-1} is based on a set $S^{p-1} \supset S^{p-2} \cdots \supset S^s$ of nested spheres for which

$$S^{k-1} \text{ is the great sphere in } S^k \text{ normal to } \mathbf{m}_k \text{ in } S^k, \text{ for } k = p-1, \ldots, s+1, \tag{1}$$

where $\mathbf{m}_k$ is some point in S^k. The tangent-normal decomposition [22, (9.1.20)] expresses each $\mathbf{x}$ in S^k as

$$\begin{aligned} \mathbf{x} &= t\,\mathbf{m}_k + (1-t^2)^{1/2}\,\mathbf{u} && (2)\\ &= \cos(r)\,\mathbf{m}_k + \sin(r)\,\mathbf{u}, && (3) \end{aligned}$$

where $t = \mathbf{x}^\top \mathbf{m}_k$, $\mathbf{u} \in S^{k-1}$, the sphere normal to $\mathbf{m}_k$, and $r = \arccos t$ is the colatitude of $\mathbf{x}$. The function $\mathbf{x} \mapsto \mathbf{u} = (1-t^2)^{-1/2}(\mathbf{x} - t\mathbf{m}_k)$ sends $S^k \setminus \{\pm\mathbf{m}_k\}$ into S^{k-1}, so that, given a distribution μ on S^{p-1}, we can define distributions $\mu_{p-1}, \ldots, \mu_s$ on $S^{p-1}, S^{p-2}, \ldots, S^s$ recursively by $\mu_{p-1} = \mu$ and μ_{k-1} as the marginal distribution of $\mathbf{u}$ on S^{k-1} for $k = p-1, \ldots, s+1$. Although the points $\mathbf{m}_{p-1}, \ldots, \mathbf{m}_{s+1}$ could be chosen as any orthonormal points in S^{p-1} (see Remark 1), we need to ensure that if $s = 0$ then μ_0 is the uniform distribution on S^0. If $\mathbf{m}_1$ is a (circular) median [22, (3.4.18)] of μ_1 then μ_0 is uniform; if the median is unique then uniformity on S^0 is equivalent to $\mathbf{m}_1$ being a median of μ_1. It is therefore convenient to take $\mathbf{m}_k$ $(k = p-1, l \ldots, s+1)$ to be the (Riemannian, alias intrinsic) Fréchet median of μ_k, i.e., the point $\mathbf{m}$ in S^k minimising the expected spherical distance $\mathrm{E}_{\mu_k}\{\arccos(\mathbf{x}^\top\mathbf{m})\}$. We shall assume that

$$\begin{aligned} &\mu \text{ is either uniform or has a unique Fréchet median } \mathbf{m}_{p-1}, && (4)\\ &\text{for } k = p-2, \ldots, s+1,\ \mu_k \text{ has a unique Fréchet median } \mathbf{m}_k, && (5)\\ &\mu_s \text{ is the uniform distribution on } S^s. && (6) \end{aligned}$$

The nested spheres in (1) are reminiscent of the principal nested spheres of [12] but, whereas principal nested spheres may be small spheres and are chosen to give closest fit to the data, the spheres in (1) are great spheres and are chosen to be orthogonal to $\mathbf{m}_{p-1}, \ldots, \mathbf{m}_{s+1}$. In cases in which (4)–(6) hold, Proposition 1 provides an almost-canonical version of the probability integral transformation on S^{p-1}.

Proposition 1. *Let μ be a probability distribution on S^{p-1} such that the density of μ with respect to the uniform distribution, ν, is continuous and positive. Suppose that μ satisfies conditions (4)–(6). Then homeomorphic almost-diffeomorphisms $\phi_k : S^k \to S^k$ for $k = s, \ldots, p-1$ can be defined inductively by (a) ϕ_s is the identity, (b) for $k = s+1, \ldots, p-1$,*

$$\phi_k(r, \mathbf{u}) = \psi_{k|\phi_{k-1}(\mathbf{u})}(r)\ \phi_{k-1}(\mathbf{u}), \tag{7}$$

where

$$\psi_{k|\mathbf{u}} = \tilde{F}_{\mathbf{u}}^{-1} \circ F_{\mathbf{u}}$$

with

$$F_{\mathbf{u}}(v) = \Pr(0 < R \le v | \mathbf{U} = \mathbf{u}) \quad \text{under } \mu_k \tag{8}$$

$$\tilde{F}_{\mathbf{u}}(v) = \Pr(0 < R \le v | \mathbf{U} = \mathbf{u}) \quad \text{under } \nu_k \tag{9}$$

for $0 \le v \le \pi$, points $\mathbf{x}$ in S^{k+1} are identified with their coordinates $(r, \mathbf{u})$ as in (3), $(R, \mathbf{U})$ denotes a random element of S^{k+1}, and ν_k is the uniform distribution on S^k. Then ϕ_{p-1} is a homeomorphic almost-diffeomorphism that transforms μ into ν.

Proof. From (7) and continuity of the density, ϕ_k is a homeomorphism of S^k and its restriction to $S^k \setminus \{\pm\mathbf{m}_k\}$ is a diffeomorphism. It is straightforward to show that ϕ_{p-1} transforms μ into ν.

2.2. *Compact Riemannian manifolds*

We now show how the probability integral transformation can be extended to arbitrary compact Riemannian manifolds in an almost-canonical way.

Let $\mathcal{X}$ be a compact Riemannian manifold. The Riemannian metric determines the volumes of infinitesimal cubes, and so equips $\mathcal{X}$ with a unique uniform probability measure, $\nu_{\mathcal{X}}$. Let μ be a probability distribution on $\mathcal{X}$ having continuous positive density with respect to $\nu_{\mathcal{X}}$. If $\mathcal{X}$ is connected then there are homeomorphisms of $\mathcal{X}$ that transform μ into $\nu_{\mathcal{X}}$; see [14, Proposition 1]. One way of constructing such homeomorphisms, ϕ, is by using the multivariate probability integral transformation (alias *Rosenblatt transformation*, [26]) in coordinate neighbourhoods, as in the first proof in [24]. In the case in which the density is smooth, there is also a slick differential-geometric proof [24, Theorem 2]. This proof can be used to provide a canonical choice of ϕ but this involves solving a differential equation and does not give ϕ explicitly. If $\mathcal{X} = S^1$ or $\dim \mathcal{X} > 1$ then, as in the spherical case, the homeomorphism ϕ is far from unique and it is not obvious how to make a canonical choice of ϕ. To obtain a canonical choice of ϕ by extending the construction in Proposition 1 to compact Riemannian manifolds, we exploit the fact that, if $\mathcal{X}$ is a Riemannian manifold and m is any point in $\mathcal{X}$ then the exponential map (see, e.g., [8, Section 1.6]) from the tangent space, $T\mathcal{X}_m$, at m into $\mathcal{X}$ can be used to identify suitable open discs round the origin of 0 in $T\mathcal{X}_m$ with suitable open sets in $\mathcal{X}$. Define the open set $\mathcal{B}$ of $\mathcal{X}$ by

$$\mathcal{B} = \{\exists r \ge 0, \exists \mathbf{u} \in T_1\mathcal{X}_m : \exists \text{ unique minimising geodesic from } m \text{ to } \exp(r\mathbf{u})\},$$

where $T_1\mathcal{X}_m$ denotes the set of unit tangent vectors at m. Then the restriction of $\exp^{-1}$ to $\mathcal{B}$ identifies $\mathcal{B}$ with $\{(r, \mathbf{u}) : 0 \le r < r_{\mathbf{u}}, \mathbf{u} \in T_1\mathcal{X}_m\}$, where

$$r_{\mathbf{u}} = \sup\{r : \exists \text{ unique minimising geodesic from } m \text{ to } \exp(r\mathbf{u})\}.$$

The mapping

$$\exp(r\mathbf{u}) \mapsto (r, \mathbf{u}) \tag{10}$$

can be regarded as giving 'polar Riemannian normal coordinates on $\mathcal{B}$'.

For $\mathcal{X} = S^{p-1}$, the tangent-normal decomposition (2) is related to these coordinates by $t = \cos r$. If $\mathcal{X}$ is compact then $\mathcal{X} \setminus \mathcal{B}$ has measure zero. See, e.g., [6, Proposition 2.113, Corollary 3.77, Lemma 3.96]. Thus absolutely continuous probability distributions on $\mathcal{X}$ can be identified with absolutely continuous probability distributions on $\{(r, \mathbf{u}) : 0 \le r < r_{\mathbf{u}}, \mathbf{u} \in T_1\mathcal{X}_m\}$. In particular, such a distribution induces a marginal distribution on $T_1\mathcal{X}_m$.

Recall that on a Riemannian manifold, $\mathcal{X}$, a Riemannian (alias intrinsic) Fréchet median of a probability distribution, μ, on $\mathcal{X}$ is a point m in $\mathcal{X}$ that minimises the expected distance $\mathrm{E}_\mu\{d(x, m)\}$, where d denotes Riemannian distance.

Proposition 2. *Let μ be a probability distribution on a compact Riemannian manifold $\mathcal{X}$ of dimension d such that the density of μ with respect to the uniform distribution, ν, is continuous and positive. Suppose that μ has a unique Fréchet median, m. Let $\{(r, \mathbf{u}) : 0 \leq r < r_{\mathbf{u}}, \mathbf{u} \in T_1\mathcal{X}_m\}$ be (maximal) polar Riemannian normal coordinates on $\mathcal{B}$ with m corresponding to the origin. Assume that the marginal distributions on $T_1\mathcal{X}_m$ obtained from μ and ν by using (10) satisfy conditions (4)–(6).*

Let $\phi_{d-1}, \tilde{\phi}_{d-1} : T_1\mathcal{X}_m \to T_1\mathcal{X}_m$ be the almost-canonical uniformising almost-diffeomorphisms corresponding to μ and ν, respectively, given by Proposition 1 and identification of $T_1\mathcal{X}_m$ with S^{d-1}. Put $\psi_{d-1} = \tilde{\phi}_{d-1}^{-1} \circ \phi_{d-1}$ and define the function $\phi : \mathcal{X} \to \mathcal{X}$ by

$$\phi\{\exp(r\mathbf{u})\} = \exp\left[\tilde{F}^{-1}_{\psi_{d-1}(\mathbf{u})}\{F_{\mathbf{u}}(r)\}\, \psi_{d-1}(\mathbf{u})\right] \qquad r\mathbf{u} \in \exp^{-1}(\mathcal{B}) \tag{11}$$

and arbitrarily on $\mathcal{X} \setminus \mathcal{B}$, where $F_{\mathbf{u}}$ and $\tilde{F}_{\mathbf{u}}$ are defined by (8) and (9) with $k = d - 1$. Then ϕ is a diffeomorphism almost everywhere and transforms μ into ν. If μ is the uniform distribution then (11) is the identity.

Proof. This is a straightforward calculation.

We call the almost-diffeomorphism ϕ of Propositions 1 or 2 the *probability integral transformation*. It is almost canonical, since it is determined (except on null sets) by unique medians at each stage.

Remark 1. *The appropriate general mathematical setting for the constructions in Propositions 1 and 2 is that of orthonormal frames in a tangent space. An orthonormal frame at a point m in a d-dimensional Riemannian manifold $\mathcal{X}$ is an ordered set of orthonormal vectors in the tangent space $T\mathcal{X}_m$. Let μ be a probability distribution on $\mathcal{X}$ such that the density of μ with respect to the uniform distribution, ν, is continuous and positive. Let $(m_{d-1}, \ldots, m_{s+1})$ be an orthonormal frame at m and suppose that the distribution on the u-dimensional sphere normal to $m_{d-1}, \ldots, m_{s+1}$ is uniform. Then replacing the successive medians in Propositions 1 and 2 by $m, m_{d-1}, \ldots, m_{s+1}$ defines an almost-diffeomorphism ϕ of $\mathcal{X}$ that takes μ to ν.*

Remark 2. *Almost-homeomorphisms can be used in the simulation of arbitrary continuous distributions on $\mathcal{X}$. Let μ and ν be probability distributions on $\mathcal{X}$ (with ν not necessarily being the uniform distribution) and ϕ any transformation (not necessarily an almost-canonical homeomorphism as introduced in this Section) that takes μ into ν. If $x_1, \ldots, x_n$ in $\mathcal{X}$ are a random sample from ν then $\phi^{-1}(x_1), \ldots, \phi^{-1}(x_n)$ are a random sample from μ.*

A class of distributions for which the probability integral transformation takes a particularly simple form consists of those with unique median m on $\mathcal{X}$ and for which the corresponding marginal distribution on $T_1\mathcal{X}_m$ (obtained using (10)) is uniform. If $\mathcal{X}$ is the sphere S^{p-1}, the projective space $\mathbb{R}P^{p-1}$, the rotation group $SO(3)$ or the complex projective space $\mathbb{C}P^{k-2}$ then this class includes the distributions that have rotational symmetry about the unique median. Some examples are:

(a) For a distribution μ on S^{p-1} that is rotationally symmetric about a unit vector $\boldsymbol{\mu}$, the transformation ϕ given by (11) that sends μ into the uniform distribution has the form

$$\phi(\mathbf{x}) = u\boldsymbol{\mu} + \{(1-u^2)/(1-t^2)\}^{1/2}\left(\mathbf{I}_p - \boldsymbol{\mu}\boldsymbol{\mu}^\top\right)\mathbf{x},$$

where $t = \mathbf{x}^\top\boldsymbol{\mu}$, $\mathbf{I}_p$ denotes the $p \times p$ identity matrix and $u = G_0^{-1}(G_\mu(t))$, G_μ and G_0 denoting the cumulative distribution functions of $\mathbf{x}^\top\boldsymbol{\mu}$ when $\mathbf{x}$ has distribution μ and the uniform distribution, respectively. In particular, for the Fisher distribution, $\mathcal{F}(\boldsymbol{\mu}, \kappa)$, on S^2 with mean direction $\boldsymbol{\mu}$ and concentration κ,

$$u = \left(2e^{\kappa t} - e^{\kappa} - e^{-\kappa}\right) / \left(e^{\kappa} - e^{-\kappa}\right), \qquad \kappa > 0 \tag{12}$$

and $u = t$ for $\kappa = 0$ (see [14, Example 1]).

(b) The angular central Gaussian distributions on the real projective space $\mathbb{R}P^{p-1}$ have probability density functions

$$f(\pm\mathbf{x}; \mathbf{A}) = |\mathbf{A}|^{-1/2}\left(\mathbf{x}^\top\mathbf{A}^{-1}\mathbf{x}\right)^{-p/2}, \mathbf{x} \in \mathbb{R}^p, \tag{13}$$

where $\mathbf{A}$ is positive definite (see [22, Section 9.4.4]). Those distributions with probability density functions (13) that are symmetrical about the modal axis $\pm\boldsymbol{\mu}$ have $\mathbf{A} = a^2\,\boldsymbol{\mu}\boldsymbol{\mu}^\top + b^2\,(\mathbf{I}_p - \boldsymbol{\mu}\boldsymbol{\mu}^\top)$ with $a > b > 0$. Then ϕ is given by

$$\phi(\pm\mathbf{x}) = \pm\left[u\boldsymbol{\mu} + \{(1-u^2)/(1-t^2)\}^{1/2}\left(\mathbf{I}_p - \boldsymbol{\mu}\boldsymbol{\mu}^\top\right)\mathbf{x}\right],$$

where $t = \mathbf{x}^\top\boldsymbol{\mu}$ and

$$u = t\{a/b + (1-a/b)t^2\}^{-1/2}.$$

The transformation ϕ coincides with the standard transformation $\pm\mathbf{x} \mapsto \pm\|\mathbf{A}^{-1/2}\mathbf{x}\|^{-1}\mathbf{A}^{-1/2}\mathbf{x}$ to uniformity on $\mathbb{R}P^{p-1}$ [22, Section 9.4.4], where $\mathbf{A}^{1/2}$ denotes the positive definite square root of $\mathbf{A}$.

(c) For the matrix Fisher distribution on $SO(3)$ with density proportional to $\exp\{\mathrm{trace}\,(\kappa\mathbf{X}^\top\mathbf{M})\}$ for $\kappa \geq 0$ and $\mathbf{M}$ in $SO(3)$, [14, Example 2] shows that $\mathbf{M}^\top\mathbf{X}$ and $\mathbf{M}^\top\phi(\mathbf{X})$ have the same rotation axis, and that the rotation angle, u, of $\mathbf{M}^\top\phi(\mathbf{X})$ is related to the rotation angle, t, of $\mathbf{M}^\top\mathbf{X}$ by

$$\tilde{F}_0(u)/\tilde{F}_0(\pi) = \tilde{F}_\kappa(t)/\tilde{F}_\kappa(\pi),$$

where $\tilde{F}_\kappa(\theta) = \int_0^\theta e^{4\kappa\cos^2(\omega/2)}\sin^2(\omega/2)d\omega$.

(d) On the shape space Σ_2^k of k non-identical labelled landmarks in $\mathbb{R}^2$, the isotropic Mardia–Dryden distributions, alias isotropic offset normal distributions, $\mathcal{MD}([\boldsymbol{\mu}],\kappa)$ [4, Section 11.1.2] of shapes $[\mathbf{X}]$ obtained by isotropic Gaussian perturbation of the landmarks of shapes $[\boldsymbol{\mu}]$ have densities

$$f([\mathbf{X}];[\boldsymbol{\mu}],\kappa) = e^{\kappa\{1-\cos^2\rho([\mathbf{X}],[\boldsymbol{\mu}])\}}\mathcal{L}_{k-2}\{-\kappa\cos^2\rho([\mathbf{X}],[\boldsymbol{\mu}])\}, \tag{14}$$

where $\mathcal{L}_{k-2}$ is the Laguerre polynomial of order $k-2$, ρ is the Riemannian shape distance and κ is a concentration parameter [4, equations (11.11), (11.15)]. Identification of $2\times(k-1)$ real matrices $\mathbf{Z}$ satisfying $\mathrm{trace}(\mathbf{Z}\mathbf{Z}^\top) = 1$ with unit vectors $\mathbf{z}$ in $\mathbb{C}^{k-1}$ leads to identification of the space Σ_2^k with the complex projective space $\mathbb{C}P^{k-2}$. Calculation shows that for the distribution with density (14), the homeomorphism ϕ is

$$\phi([\mathbf{z}]) = \left[u\boldsymbol{\mu} + \{(1-u^2)/(1-t^2)\}^{1/2}\left\{\mathbf{z} - (\mathbf{z}^\top\boldsymbol{\mu})\boldsymbol{\mu}\right\}\right],$$

where $t = \cos\rho([X],[\mu])$, $u^2 = F^{-1}_{[X],0}\left\{F_{[X],\kappa}(t^2)\right\}$ with $F_{[X],\kappa}$ defined by

$$F_{[X],\kappa}(x) = (k-2)e^\kappa\sum_{i=0}^{k-2}\sum_{r=0}^{k-3}\binom{k-2}{i}\binom{k-3}{r}\frac{(-1)^r\kappa^i}{i!}\int_0^x e^{-\kappa s}s^{r+i}ds.$$

For $\kappa = 0$ (corresponding to the uniform distribution) $F_{[X],\kappa}$ takes the simple form

$$F_{[X],0}(x) = 1 - (1-x)^{k-2}.$$

2.3. Shape spaces

The probability integral transformation can be defined also for the shape spaces, Σ_m^k, of shapes of k non-identical labelled landmarks in $\mathbb{R}^m$. As indicated after (14), the space Σ_2^k can be identified with the complex projective space $\mathbb{C}P^{k-2}$, and so is a compact Riemannian manifold. For $m > 2$, Σ_m^k is not a manifold but for our purposes, it is enough to work on the *non-singular part* of Σ_m^k, which is the open set consisting of the shapes of k non-identical labelled landmarks in $\mathbb{R}^m$ that do not lie in any $(m-2)$-dimensional affine subspace.

It follows from [16, Section 6.3 and Theorem 6.5] that, for x in the non-singular part of Σ_m^k there is a system of Riemannian normal coordinates with inverse that maps an open set $\{(r,\mathbf{u}) : 0 \leq r < r_\mathbf{u}, \mathbf{u} \in T_1\mathcal{X}_x\}$ diffeomorphically onto an open set $\mathcal{B}$ of Σ_m^k by (10), where $T_1\mathcal{X}_x$ denotes the set of unit tangent vectors at x, and $\Sigma_m^k \setminus \mathcal{B}$ has measure zero. If the distribution on $T_1\mathcal{X}$ satisfies conditions (4)–(6) then the probability integral transform can be defined as in Proposition 2.

A referee has pointed out that this construction can be extended to more general shape spaces. For a quotient of a Riemannian manifold by a proper isometric Lie group action, the singularity set has dimension less than that of the manifold [9], and so (10) is an almost-diffeomorphism.

2.4. Cartan–Hadamard manifolds

The Cartan–Hadamard manifolds are the complete simply-connected manifolds with non-positive curvature. It follows from the Cartan–Hadamard theorem [8, Theorem I 13.3], [18] that on a Cartan–Hadamard manifold, $\mathcal{X}$, the inverse of the exponential map at any basepoint x identifies $\mathcal{X}$ with $T\mathcal{X}_x$. Then the choice of a distribution ν (with positive density) on $\mathcal{X}$ enables an extension of the approach used in Section 3. Important instances of such manifolds are the simplicial shape spaces of shapes of m-simplices in $\mathbb{R}^m$ with positive volume, equipped with a Riemannian metric derived from a natural metric on $SL(m)$ [27, Section 3.6.2], [20, Section 2]. The case $m = 2$ gives the space of shapes of non-degenerate triangles in the plane, which can be identified with the Poincaré half-plane, $\mathbb{H}^2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$, with Riemannian metric $g_{ij} = \delta_{ij} x_2^{-2}$. This space was used in [10] as a sample space for electrical impedances.

Proposition 3. *Let μ and ν be probability distributions with positive densities on a Cartan–Hadamard manifold, $\mathcal{X}$, of dimension d. Let m be a point of $\mathcal{X}$ and $\{(r, \mathbf{u}) : 0 \le r, \mathbf{u} \in T_1\mathcal{X}_m\}$ be polar Riemannian normal coordinates on $\mathcal{X}$ with m corresponding to the origin. Let $\phi_{d-1}, \tilde{\phi}_{d-1} : T_1\mathcal{X}_m \to T_1\mathcal{X}_m$ be the almost-canonical uniformising almost-diffeomorphisms corresponding to μ and ν, respectively, given by Proposition 1 and identification of $T_1\mathcal{X}_m$ with S^{d-1}. Put $\psi_{d-1} = \tilde{\phi}_{d-1}^{-1} \circ \phi_{d-1}$ and define the function $\phi : \mathcal{X} \to \mathcal{X}$ by*

$$\phi\{\exp(r\mathbf{u})\} = \exp\left[\tilde{F}^{-1}_{\psi_{d-1}(\mathbf{u})}\{F_{\mathbf{u}}(r)\}\ \psi_{d-1}(\mathbf{u})\right],$$

where $F_{\mathbf{u}}$ and $\tilde{F}_{\mathbf{u}}$ are defined by (8) and (9) with $k = d - 1$. Then ϕ is an almost-diffeomorphism that maps geodesics through m into geodesics through m and transforms μ into ν.

2.5. Simplices

The open $(p-1)$-simplex is

$$\Delta_{p-1} = \left\{(y_1, \dots, y_p) : y_j > 0,\ \sum_{j=1}^{p} y_j = 1\right\}.$$

There is a canonical base point, the centroid, $\mathbf{c} = (p^{-1}, \dots p^{-1})$ and a canonical Riemannian metric obtained by regarding Δ_{p-1} as an affine subspace of $\mathbb{R}^p$. The unit tangent sphere at $\mathbf{c}$ is

$$T_1\Delta_{p-1,\mathbf{c}} = \left\{\mathbf{u} = (v_1, \dots, v_p) : \sum_{j=1}^{p} v_j = 0,\ \sum_{j=1}^{p} v_j^2 = 1\right\}$$

and the exponential map is

$$\exp(r\mathbf{u}) = \mathbf{c} + r\mathbf{u} \tag{15}$$

for $r \in [0, 1/(p\max_{1\le j\le p}|v_j|)]$. The uniform distribution is a scaled version of Lebesgue measure on Δ_{p-1} and the corresponding marginal distribution on the unit tangent sphere is the uniform distribution on $T_1\Delta_{p-1,\mathbf{c}}$.

2.5.1. Using the exponential map

The manifold Δ_{p-1} is simply connected and has curvature 0 but it is not complete. The exponential map (15) is a diffeomorphism between a star-shaped portion of $T\Delta_{p-1,\mathbf{c}}$ and Δ_{p-1}. Let μ be a distribution on Δ_{p-1} with continuous positive density with respect to the uniform distribution, ν. Then a minor variant of Proposition 3 produces a canonical almost-diffeomorphism $\phi : \Delta_{p-1} \to \Delta_{p-1}$ that transforms μ into ν.

Proposition 4. *Let μ be a probability distribution on Δ_{p-1} having continuous positive density with respect to Lebesgue measure. Let $\mathbf{c}$ be the barycentre of Δ_{p-1} and $\left\{(r, \mathbf{u}) : 0 \le r, \mathbf{u} \in T_1\Delta_{p-1,\mathbf{c}}\right\}$ be polar Riemannian normal coordinates on Δ_{p-1} with $\mathbf{c}$ corresponding to the origin. Let $(R, \mathbf{U})$ be the normal coordinates of a random element of Δ_{p-1}. Let $\psi : T_1\Delta_{p-1,\mathbf{c}} \to T_1\Delta_{p-1,\mathbf{c}}$ be the almost-canonical homeomorphism such that $\psi(\mathbf{U})$ is uniformly distributed. Identification of $T_1\Delta_{p-1,\mathbf{c}}$ with S^{p-2} leads to definition of $F_{\mathbf{u}}$ and $\tilde{F}_{\mathbf{u}}$ by (8) and (9) with $k = p-2$. Define the function $\phi : \Delta_{p-1} \to \Delta_{p-1}$ by*

$$\phi\{\exp(r\mathbf{u})\} = \exp\left[\tilde{F}^{-1}_{\psi(\mathbf{u})}\{F_{\mathbf{u}}(r)\}\ \psi(\mathbf{u})\right].$$

Then ϕ is a diffeomorphism almost everywhere, maps geodesics through $\mathbf{c}$ into geodesics through $\mathbf{c}$, and transforms μ into ν.

2.5.2. *Using radial projection*

An alternative to using the exponential map (15) is to use 'radial projection' of $\Delta_{p-1} \setminus \{\mathbf{c}\}$ onto its boundary $\partial\Delta_{p-1}$. The coordinates $(r, z_1, \ldots, z_p)$ given by radial projection are defined by

$$r = \begin{cases} 0 & \text{if} \quad \mathbf{x} = \mathbf{c}, \\ 1 - py_{(1)} & \text{if} \quad \mathbf{x} \neq \mathbf{c} \end{cases} \tag{16}$$

$$z_j = r^{-1}(y_j - y_{(1)}) \quad j = 1, \ldots, p, \tag{17}$$

$y_{(1)}$ denoting the smallest of $y_1, \ldots, y_p$. Then $r \in [0, 1)$. A simple calculation shows that the density of the uniform distribution with respect to $dr\, dz_1 \ldots dz_{i-1}\, dz_{i+1} \ldots dz_p$ is proportional to $r^{-(p-1)}$. It follows that, for $i = 1, \ldots, p$, radial projection of $\Delta_{p-1,i} = \left\{(y_1, \ldots, y_p) \in \Delta_{p-1} \setminus \{\mathbf{c}\} : y_{(1)} = y_i\right\}$ onto the face $\partial_i\Delta_{p-1} = \left\{(z_1, \ldots, z_p) : z_i = 0\right\}$ sends the uniform distribution on $\Delta_{p-1,i}$ to the uniform distribution on the $(p-2)$-simplex $\partial_i\Delta_{p-1}$. The boundary, $\partial\Delta_{p-1}$, of Δ_{p-1} is the union of $\partial_1\Delta_{p-1}, \ldots, \partial_p\Delta_{p-1}$.

The next proposition shows that radial projection provides canonical uniformising homeomorphic almost-diffeomorphisms of simplices that are analoguous to those for spheres that are described in Proposition 1. Unlike the construction in Proposition 1, the construction in Proposition 5 does not assume uniqueness of medians, as in (4)–(5).

Proposition 5. *Let μ be a probability distribution on Δ_{p-1} having continuous positive density with respect to Lebesgue measure. For $k = 0, \ldots, p-2$, denote by $\partial^{p-1-k}\Delta_{p-1}$, the union of the k-dimensional faces of Δ_{p-1}. Then repeated radial projection sends μ to a probability distribution μ_k on $\partial^{p-1-k}\Delta_{p-1}$. Let s be the largest value of k for which μ_k is uniform. For $k = s+1, \ldots, p-1$, let $r, z_1, \ldots, z_{i-1}, z_{i+1}, \ldots z_k$ be coordinates (defined analogously to those in (16)–(17)) on the part of the $(p-1-k)$-simplex in $\partial^{p-1-k}\Delta_{p-1}$ on which $z_i = 0$. Define functions $\phi_k : \partial^{p-k-1}\Delta_{p-1} \to \partial^{p-k-1}\Delta_{p-1}$ for $k = s, \ldots, p-1$ recursively by (a) ϕ_s is the identity, (b) for $k = s+1, \ldots, p-1$,*

$$\phi_k(r, \mathbf{z}) = F_{\mathbf{z}}(r)^{1/(k+2-p)}\ \phi_{k-1}(\mathbf{z}),$$

where

$$F_{\mathbf{z}}(t) = \Pr(0 < R \leq t | \mathbf{Z} = \mathbf{z}) \quad \text{under } \mu_k$$

for $0 \leq t \leq 1$ and $\mathbf{z} = (z_1, \ldots, z_{i-1}, z_{i+1}, \ldots z_k)$, Then ϕ_{p-1} is a homeomorphic almost-diffeomorphism that transforms μ into ν.

Proof. This is a straightforward calculation using the fact that $\Pr(0 < R \leq t | \mathbf{Z} = \mathbf{z}) = t^{k+2-p}$ under the uniform distribution on this $(p-1-k)$-simplex.

3. Goodness-of-fit tests via transformation

For many of the sample spaces that we consider there are well-established tests of, e.g., uniformity. Transformations can be used to adapt these to give tests of goodness of fit. Let μ and ν be probability distributions on $\mathcal{X}$. Then any transformation, ϕ, that takes μ into ν can be used to transform any test, T, of goodness of fit to ν into a test, ϕ^*T, of goodness of fit to μ. Given points $x_1, \ldots, x_n$ in $\mathcal{X}$, ϕ^*T is obtained by applying T to the transformed data, $\phi(x_1), \ldots, \phi(x_n)$. The null distribution of ϕ^*T is the same as that of T. (Although our almost-canonical construction in Section 2 of suitable transformations ϕ proceeds via a tangent space, $T\mathcal{X}_m$, this tangent space plays no key role in the test ϕ^*T.)

Often the null hypothesis about the distribution generating the data is not that it is some specified distribution but that it is a distribution in a given parametric model, $\{\mu_\theta : \theta \in \Theta\}$. For each θ in Θ, let ϕ_θ be a transformation that takes μ_θ into ν. Let $\hat{\theta}$ be an estimate of θ. Then goodness of fit to $\{\mu_\theta : \theta \in \Theta\}$ is tested by applying T to the transformed data, $\phi_{\hat{\theta}}(x_1), \ldots, \phi_{\hat{\theta}}(x_n)$. Significance can be assessed by simulation from the fitted distribution. If a good approximation to the null distribution of T is available then simulation can be avoided by using this approximation.

Provided that the estimator giving $\hat{\theta}$ is consistent, the consistency properties of ϕ^*T are inherited from those of T. In particular, if $\hat{\theta}$ is the maximum likelihood estimate then ϕ^*T is consistent against all alternatives if and only if T is consistent against all alternatives.

3.1. Spheres

On a sphere the uniform distribution provides a canonical choice for ν. Then the transformation, ϕ, of Proposition 1 that takes μ into ν can be used to transform tests of uniformity into tests of goodness of fit to μ. If the test of uniformity is (like the Sobolev tests of [7]) invariant under isometries of the sphere then ϕ need not be specified fully but only up to composition with a rotation. In this case, ϕ_s defined in (a) of Proposition 1 need not be the identity of S^s but can be any rotation.

One nice characterisation of the uniform distributions on S^2 is that, for a uniformly distributed random vector with longitude ψ and colatitude θ, (a) ψ is uniformly distributed on $[0, 2\pi]$, (b) $\cos\theta$ is uniformly distributed on $[-1, 1]$, (c) ψ and θ are independent. Thus combining any tests of (a), (b) and (c) gives a test of uniformity on S^2. Using the general construction given in the previous paragraph with $\phi : S^2 \to S^2$ given by (11) but with (12) replaced by the approximation $2e^{\kappa(t-1)} - 1$ to (12) for κ not close to 0, taking the tests in (a), (b) and (c) to be Kuiper's V_n, the Kolmogorov–Smirnov test, and a rather special '2-variable' test yields the standard method [22, Section 12.3.1] of investigating goodness of fit of Fisher distributions on S^2.

3.2. Compact Riemannian manifolds and shape spaces

On a compact Riemannian manifold or a shape space the uniform distribution provides a canonical choice for ν. Then the transformation, ϕ, of Proposition 2 that takes μ into ν can be used to transform tests of uniformity into tests of goodness of fit to μ.

3.3. Cartan–Hadamard manifolds

Let m be a point in a Cartan–Hadamard manifold, $\mathcal{X}$, and let μ and ν be probability distributions on $\mathcal{X}$ and $T\mathcal{X}_m$, respectively, such that the density of μ with respect to ν is positive. By Proposition 3, there is an almost-canonical almost-diffeomorphism $\phi : \mathcal{X} \to T\mathcal{X}_m$ that transforms μ into ν. Since $T\mathcal{X}_m$ can be identified with $\mathbb{R}^d$ (where d is the dimension of $\mathcal{X}$), standard goodness-of-fit tests on $\mathbb{R}^d$ can be adapted to give goodness-of-fit tests on $\mathcal{X}$.

3.4. Simplices

On the simplex Δ_{p-1} the uniform distribution provides a canonical choice for ν. Then the transformation, ϕ, of Proposition 4 or Proposition 5 that takes μ into ν can be used to transform tests of uniformity into tests of goodness of fit to μ.

An appealing test of uniformity on Δ_{p-1} is the score test of uniformity ($\alpha_1 = \ldots = \alpha_p = 1$) within the Dirichlet family with densities (with respect to the uniform distribution)

$$f(y_1, \ldots, y_p; \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{j=1}^{p} \alpha_j)}{\prod_{j=1}^{p} \Gamma(\alpha_j)} \prod_{j=1}^{p} y_j^{\alpha_j - 1},$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$ with $\alpha_i > 0$ for $i \in \{1, \ldots, p\}$. For independent observations $\mathbf{y}_1, \ldots, \mathbf{y}_n$ on Δ_{p-1} with $\mathbf{y}_i = (y_{i1}, \ldots, y_{ip})$ (for $i = 1, \ldots, n$), this score test rejects uniformity for large values of

$$S_n = \frac{n}{\psi'(1)} \left\{ \frac{\psi'(p)}{p\psi'(p) - \psi'(1)} \sum_{j=1}^{p} \sum_{k=1}^{p} w_j w_k - \sum_{j=1}^{p} w_j^2 \right\},$$

where $w_j = n^{-1} \sum_{i=1}^{n} \ln y_{ij}$ and ψ denotes the digamma function. Under uniformity the large-sample asymptotic distribution of S_n is χ^2_p.

4. Simulation studies

In order to assess the performance of our tests, we consider three simulation studies. The first involves the goodness-of-fit test on S^2 based on the Rayleigh test of uniformity. First 10,000 random samples of size 50 were simulated from the Fisher distribution $\mathcal{F}(\boldsymbol{\mu}, \kappa)$ with given mode $\boldsymbol{\mu}$ and concentration $\kappa = 10$. For each sample, goodness of fit to (a) the true $\mathcal{F}(\boldsymbol{\mu}, 10)$ distribution, (b) the fitted $\mathcal{F}(\hat{\boldsymbol{\mu}}, \hat{\kappa})$ distribution, where $\hat{\boldsymbol{\mu}}$ and $\hat{\kappa}$ are the maximum likelihood estimates of $\boldsymbol{\mu}$ and κ, was assessed. Then 10,000 random samples of size 50 were simulated from the projected normal $\mathcal{PN}_3(\boldsymbol{\mu}, \mathbf{I}_3)$ distribution (obtained by projecting the trivariate normal $\mathcal{N}_3(\boldsymbol{\mu}, \mathbf{I}_3)$ distribution radially onto S^2) and goodness of fit to the $\mathcal{F}(\boldsymbol{\mu}, 10)$ distribution was assessed. The resulting p-values (based on the large-sample asymptotic χ^2_3 distribution) are shown in the histograms on the left of Fig. 1. Corresponding histograms for 1,000 samples of size 500 are given on the right of Fig. 1. The fairly uniform distribution of p-values for fit to the true distribution indicates that the test tends not to reject the null hypothesis when it is true, whereas the clustering of p-values near 1 when assessing goodness of fit to the *fitted* distribution shows the anticipated excellent fit to the fitted distribution. For samples generated from $\mathcal{PN}_3(\boldsymbol{\mu}, \mathbf{I}_3)$, the p-values for fit to the $\mathcal{F}(\boldsymbol{\mu}, 10)$ distribution also cluster near 1, meaning that this test does not detect that the data come from the wrong model.

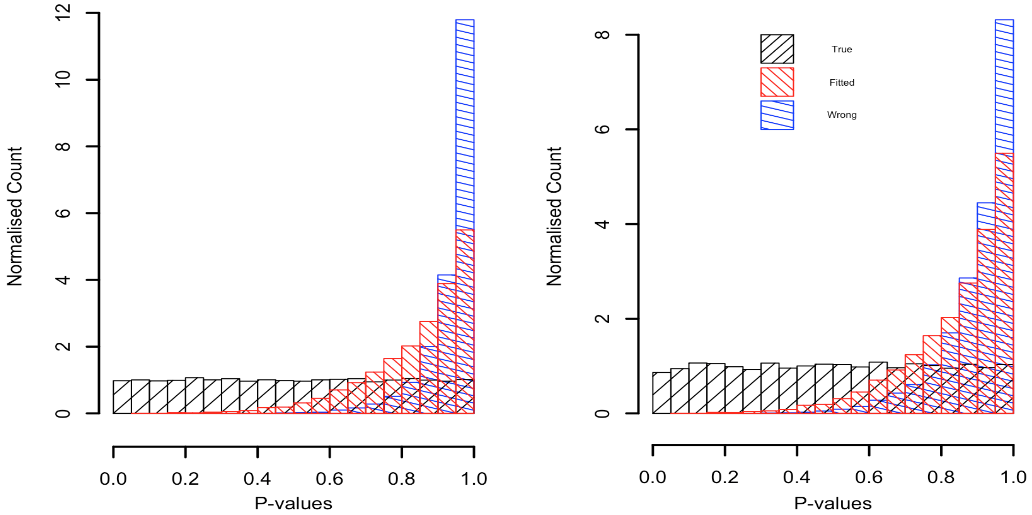

Fig. 1: Behaviour of test of goodness of fit (a) to true $\mathcal{F}(\mu, 10)$ distribution on S^2 (black), (b) to fitted $\mathcal{F}(\hat{\mu}, \hat{\kappa})$ distribution (red), (c) to projected normal $\mathcal{PN}_3(\mu, \mathbf{I}_3)$ distribution (blue), using test based on Rayleigh's test of uniformity. The histograms are of p-values (based on the large-sample asymptotic χ^2_3 distribution). Each histogram summarizes 10,000 simulations, each of size 50 (left) or 500 (right).

One possible explanation for the inability of the above test to detect that the data come from the wrong model is that the Rayleigh test of uniformity is not consistent against all alternatives. Therefore a second simulation study was carried out, which was like the first but with the Rayleigh test replaced by Giné's [7] F_n test [22, Section 10.4.1], which is consistent against all alternatives to uniformity on S^2. Histograms of the resulting values of F_n are shown in Fig. 2 for sample sizes, n, of 50 (left) and 500 (right). Significance was assessed using the asymptotic quantiles given in [15] and [22, Section 10.4.1]. For assessing goodness of fit to the true distribution, the proportions of the values of the statistic that exceeded the asymptotic 10%, 5% and 1% upper quantiles were 0.10, 0.05 and 0.01 ($n = 50$) and 0.10, 0.04 and 0.01 ($n = 500$), respectively, indicating that the test detects good fit when it is present. For fit to the *fitted* distribution, none of the values of F_n exceeded the asymptotic 10% quantile, indicating the anticipated excellent fit to the fitted distribution. For samples generated from $\mathcal{PN}_3(\boldsymbol{\mu}, \mathbf{I}_3)$, the proportions of the values of F_n that exceeded the asymptotic 10%, 5% and 1% upper quantiles were 0.58, 0.34 and 0.05 for $n = 50$, while for $n = 500$, all the values of F_n far exceeded the asymptotic 1% upper quantile. This indicates clearly that the test can detect bad fit.

The third simulation study involves the goodness-of-fit test on Σ^5_2 based on Mardia's [21] test of uniformity. First, 10,000 random samples of size 50 were simulated from the isotropic Mardia–Dryden $\mathcal{MD}([\boldsymbol{\mu}], 0.125)$ distribution with given mode $[\boldsymbol{\mu}]$. For each sample, goodness of fit to (a) the true $\mathcal{MD}([\boldsymbol{\mu}], 0.125)$ distribution, (b) the fitted $\mathcal{MD}([\hat{\boldsymbol{\mu}}], \hat{\kappa})$ distribution, where $[\hat{\boldsymbol{\mu}}]$ and $\hat{\kappa}$ are the maximum likelihood estimates of $[\boldsymbol{\mu}]$ and κ (calculated by the EM method of [19]), was assessed using Mardia's uniformity test on Σ^5_2. Then 10,000 random samples of size 50 were simulated from the non-isotropic Mardia–Dryden distribution obtained by Gaussian $\mathcal{N}_2(\mathbf{0}, \Sigma)$ perturbations of $\boldsymbol{\mu}$, where $\Sigma = \mathrm{diag}(1, 25)$, and goodness of fit to the $\mathcal{MD}([\boldsymbol{\mu}], 0.125)$ distribution was assessed. The resulting p-values based on the large-sample asymptotic χ^2_{15} distribution are shown in the histograms on the left of Fig. 3. Corresponding histograms for 10,000 samples of size 500 are given on the right. The fairly uniform distribution of p-values for fit

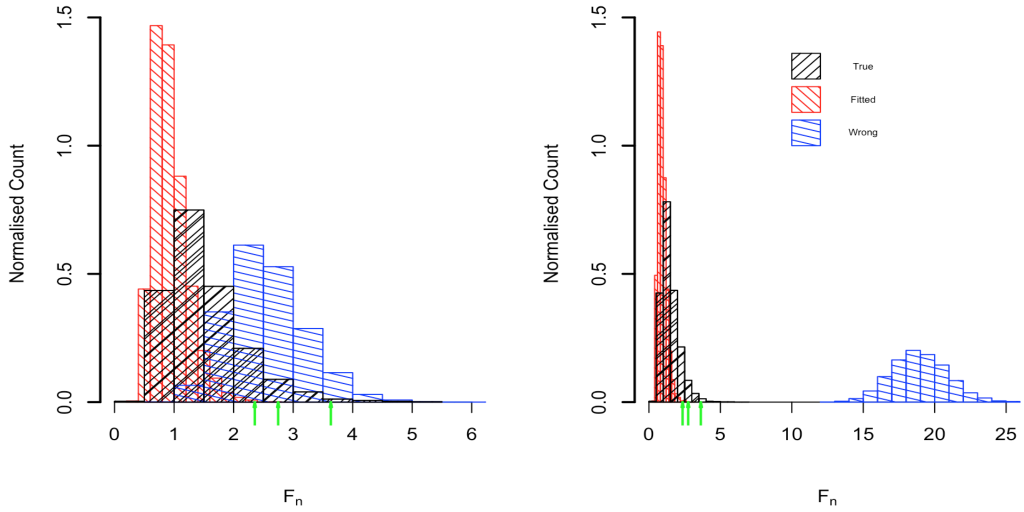

Fig. 2: Behaviour of test of goodness of fit (a) to true $\mathcal{F}(\boldsymbol{\mu}, 10)$ distribution on S^2 (black), (b) to fitted $\mathcal{F}(\hat{\boldsymbol{\mu}}, \hat{\kappa})$ distribution (red), (c) to projected normal $\mathcal{PN}_3(\boldsymbol{\mu}, \mathbf{I}_3)$ distribution (blue), using test based on Giné's F_n test of uniformity. The histograms are of values of F_n. Each histogram summarizes 10,000 simulations, each of size 50 (left) or 500 (right). Green arrows on horizontal axes are 10%, 5% and 1% upper quantiles of asymptotic distribution.

to the true distribution indicates that the test tends not to reject the null hypothesis when it is true. The clustering of p-values near 1 for fit to the *fitted* distribution shows the anticipated excellent fit to the fitted distribution. For samples generated from the non-isotropic distribution, the p-values cluster near 0, indicating that the test can detect bad fit.

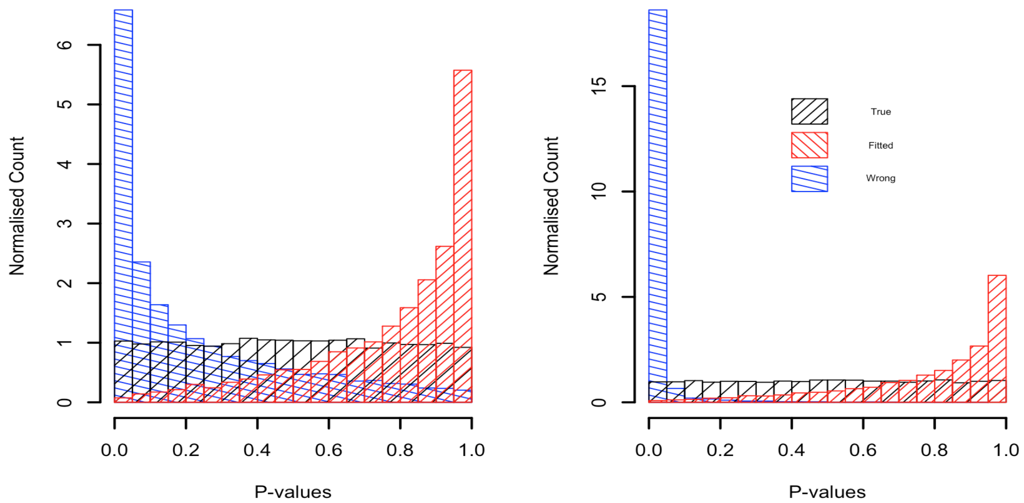

Fig. 3: Behaviour of test of goodness of fit (a) to true isotropic Mardia–Dryden $\mathcal{MD}([\boldsymbol{\mu}], 0.125)$ distribution on Σ_2^5 (black), (b) to fitted isotropic $\mathcal{MD}([\hat{\boldsymbol{\mu}}], \hat{\kappa})$ distribution (red), (c) to non-isotropic Mardia–Dryden distribution obtained by Gaussian $\mathcal{N}_2(\mathbf{0}, \mathrm{diag}(1, 25))$ perturbations of $\boldsymbol{\mu}$ (blue), using test based on Mardia's test of uniformity. The histograms are of p-values (based on the large-sample asymptotic χ^2_{15} distribution). Each histogram summarizes 10,000 simulations, each of size 50 (left) or 500 (right).

References

[1] R. Beran, Exponential models for directional data, Ann. Statist. 7 (1979) 1162–1178.

[2] D.J. Best, N.I. Fisher, Goodness-of-fit and discordancy tests for samples from the Watson distribution on the sphere, Austral. J. Statist. 28 (1986) 13–31.
[3] B. Boulerice, G. Ducharme, Smooth tests of goodness-of-fit for directional and axial data, J. Multivariate Anal. 60 (1997) 154–175.
[4] I.L. Dryden, K.V. Mardia, Statistical Shape Analysis with applications in R, second ed., Wiley, Chichester, 2016.
[5] N.I. Fisher, D. Best, Goodness-of-fit tests for Fisher's distribution on the sphere. Austral. J. Statist. 26 (1984) 142–150.
[6] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry, second ed., Springer-Verlag, Berlin, 1993.
[7] E. Giné M., Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms, Ann. Statist. 3 (1975) 1243–1266.
[8] S. Helgason, Differential Geometry, Lie Groups and Symmetric Spaces. Academic Press, New York, 1978.
[9] S. Huckemann, T. Hotz, A. Munk, Intrinsic shape analysis: Geodesic principal component analysis for Riemannian manifolds modulo Lie group actions (with discussion), Statist. Sinica 20 (2010) 1–100.
[10] S.F. Huckemann, P.T. Kim, J.Y. Koo, A. Munk, Möbius deconvolution on the hyperbolic plane with application to impedance density estimation, Ann. Statist. 38 (2010) 2465–2498.
[11] M.C. Jones, A. Pewsey, S. Kato, On a class of circulas for circular distributions, Ann. Inst. Statist. Math. 67 (2015) 843–862.
[12] S. Jung, I.L. Dryden, J.S. Marron, Analysis of principal nested spheres, Biometrika 99 (2012) 551–568.
[13] P.E. Jupp, Sobolev tests of goodness of fit of distributions on compact Riemannian manifolds, Ann. Statist. 33 (2005) 2957–2966.
[14] P.E. Jupp, Copulae on products of compact Riemannnian manifolds, J. Multivariate Anal. 14 (2015) 92–98.
[15] J. Keilson, J., D. Petrondas, U. Sumita, J. Wellner, Significance points for some tests of uniformity on the sphere, J. Statist. Comput. Simulation 17 (1983), 195–218.
[16] D.G. Kendall, D. Barden, T.K. Carne, H. Le, Shape and Shape Theory, Wiley, Chichester, 1999.
[17] J.T. Kent, The Fisher–Bingham distribution on the sphere, J. R. Stat. Soc. B 44 (1982) 71–180.
[18] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry, Volume II, Interscience, New York, 1969.
[19] A.Kume, M. Welling, Maximum likelihood estimation for the offset-normal shape distributions using EM, J. Comp. Graphical Stat. 19 (2010) 702–723.
[20] H. Le, C.G. Small, Multidimensional scaling of simplex shapes, Pattern Recognition 32 (1999) 1601–1613.
[21] K.V. Mardia, Directional statistics and shape analysis, J. Appl. Statist. 26 (1999) 949–957.
[22] K.V. Mardia, P.E. Jupp, Directional Statistics, Wiley, Chichester, 2000.
[23] K.V. Mardia, D. Holmes, J.T. Kent, A goodness-of-fit test for the von Mises–Fisher distribution, J. R. Stat. Soc. B 46 (1984) 72–78.
[24] J. Moser, On the volume elements on a manifold, Trans. Amer. Math. Soc. 120 (1965) 286–294.
[25] A. Pewsey, S. Kato, Parametric bootstrap goodness-of-fit testing for Wehrly–Johnson bivariate circular distributions, Statistics and Computing 26 (2016) 1307–1317.
[26] M. Rosenblatt, Remarks on a multivariate transformation, Ann. Math. Statist. 23 (1952) 470–1472.
[27] C.G. Small, The Statistical Theory of Shape, Springer, New York, 1996.