

Finite Sample Smeariness of Fréchet Means and Application to Climate

Shayan Hundrieser*, Benjamin Eltzner* and Stephan F. Huckemann*

Abstract

Fréchet means on manifolds are minimizers of expected squared distance, just as their Euclidean kin. On manifolds, Fréchet means for most distributions feature a classical asymptotic Gaussian central limit theorem with \sqrt{n} -rate, except for some distributions, some of which are known on spheres. These slower and non-Gaussian rates are called *smeariness* rates. Here we first describe exhaustively all smeary distributions on circles, comprising not only *power* smeariness, but also *logarithmic* smeariness. Then we make the concept of *finite sample smeariness* (FSS) precise and show for the circle that FSS affects all distributions that are spread on both sides beyond an open half circle. An analog result holds for tori. We statistically quantify the *scale of FSS* and find two different qualities: If asymptotically, the sample Fréchet variance remains strictly above the population Fréchet variance, we speak of Type I FSS, this is the case, e.g., for all (!) von Mises distributions. Else, we have Type II FSS, e.g., for distributions with support excluding the antipode of its mean. For both types of FSS it turns out that the nominal level of asymptotic quantile based tests for the circular mean is higher than the true level. Simulations indicate, however, that suitably designed bootstrap tests preserve the level. For illustration of the relevance of FSS in real data, we apply our method to directional wind data from two European cities. It turns out that quantile based tests, not correcting for FSS, find a multitude of significant wind changes. This multitude condenses to a few years featuring significant wind changes, when our bootstrap tests are applied, correcting for FSS.

Key words and phrases: Fréchet means, smeariness, one- and two-sample tests, nonparametric asymptotic quantile based tests, bootstrap tests, directional data on circles and tori, wind directions.

Figure *AMS 2000 Subject Classification:* Primary 62 H 11
Secondary 60 F 05, 62 G 10

1 Introduction

Generalizing the *expected value* for random variables X on arbitrary metric spaces (M, d) , Fréchet (1948) proposed to consider minimizers of the expected squared distance:

$$\operatorname{argmin}_{x \in M} F(x) \quad \text{where} \quad F(x) = \mathbb{E} [d(x, X)^2] . \quad (1)$$

Under mild assumptions, Ziezold (1977) derived strong asymptotic consistency for

$$\operatorname{argmin}_{x \in M} F_n(x) \quad \text{where} \quad F_n(x) = \frac{1}{n} d(x, X_j)^2 \quad (2)$$

from an i.i.d. sample $X_1, \dots, X_n \sim X$. Under stronger conditions, among others, under uniqueness of the minimizer μ of (1), called a *Fréchet population mean*, and that M is a manifold, Bhattacharya and Patrangenaru (2005) derived, in a local chart near μ , a central limit theorem for a measurable selection

*Institute for Mathematical Stochastics, Georg-August-Universität Göttingen

$\widehat{\mu}_n$ of (2), called a *Fréchet sample mean*, with a Gaussian limiting distribution and the usual rate of $n^{-1/2}$.

For the subsequent decade, it was unclear, whether any other limiting distributions were possible, possibly with different rates. Then Hotz and Huckemann (2015) discovered certain distributions on the circle, featuring rates of $n^{-\frac{1}{2(r+1)}}$, for all integer $r \in \{0, 1, \dots\}$, with powers of Gaussians as limiting distributions. Such rates have been called *smeariness* of order r , where *zero-order smeariness* corresponds to the classical CLT. Recently, Eltzner and Huckemann (2019); Eltzner (2019); Tran (2019) derived distributions on spheres of arbitrary dimension, featuring smeariness of order $r = 2$.

This present article shows that, in contrast to naïve belief, initially also held by the authors, the phenomenon of smeariness extends to a wide class of distributions manifesting as *finite sample smeariness* (FSS), namely that, depending on the distribution, there are integers $n_- < n_+$ and $0 < \alpha_- < \alpha_+ < 1$ such that

$$\mathbb{E}[d(\widehat{\mu}_n, \mu)^2] \quad \text{is of order between} \quad \frac{1}{n^{\alpha_+}} \quad \text{and} \quad \frac{1}{n^{\alpha_-}} \quad \text{for} \quad n \in \{n_-, n_- + 1, \dots, n_+\} \quad (3)$$

cf. Figure 1.

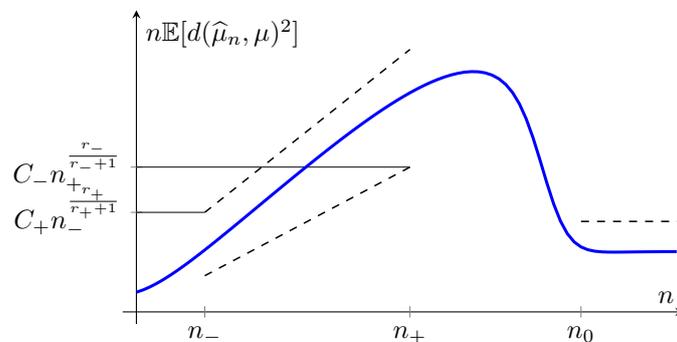


Figure 1: *Scaled variance curve $n \mapsto n\mathbb{E}[d(\widehat{\mu}_n, \mu)^2]$ (blue) of an exemplary distribution featuring FSS. Along $[n_-, n_+]$ the curve is between the lower ($C_- n_+^{\alpha_-}$) and upper ($C_+ n_-^{\alpha_+}$) bounds (dashed), satisfying the condition $C_+ n_-^{\alpha_+} \leq C_- n_+^{\alpha_-}$, and for $n \geq n_0$ it is below the horizontal upper bound (dashed). The exponents relate to power smeariness $0 < r_- < r_+$ via $\alpha_- = \frac{r_-}{r_-+1}$ and $\alpha_+ = \frac{r_+}{r_++1}$ as detailed in Section 2 and Definition 3.1.*

We will make the above informal definition (3) of FSS precise and, on the circle, we will show that FSS affects all distributions which are spread on both sides beyond an open half circle. In particular, FSS affects all distributions with nowhere vanishing density such as all von Mises families, Bingham families, etc., an overview of many more such distributions is given by Mardia and Jupp (2000). Upon closer inspection, we find two types of FSS. All of the distributions mentioned before feature FSS of *Type I*. Distributions with FSS and zero density at the antipode of their Fréchet mean μ , feature FSS of *Type II*. This is the case, for instance, when resampling from samples that are spread around beyond a half circle.

Further, we show how FSS affects quantile based one- and two-sample tests such that their nominal size may be drastically different from their effective size, cf. Table 1.

Suitably designed bootstrap tests, however, seem to preserve the size rather well. In application to wind direction data we see that asymptotic quantile based two-sample test, on a circle and a torus, due to FSS present in the data, give a wrong impression of the extent of extreme wind change events for two continental European cities. Using bootstrap tests which preserve the nominal level, extreme wind events reduce to a few concise events, which in particular make the years 2003, 2005 and 2017, 2018 exceptional (the latter strongly), hinting towards a trace of recent climate change.

We begin our exposition by a new and rather exhaustive framework describing all types of smeariness possible on the circle and the torus, choosing a tradeoff between generality and technical simplicity. In

λ	0	1/4	1/2	3/4
$n = 100$	0.333	0.446	0.564	0.676
$n = 1000$	0.349	0.496	0.604	0.768

Table 1: *Effective size under the null hypothesis of equal von-Mises mixture distributions as defined in (9) (with $\kappa = 3, \alpha = 1/2$ and varying λ) of the quantile-based two-sample Test 4.1(ii) for equality of Fréchet means with nominal size 0.05.*

this framework, *order smeariness* as discovered before by Hotz and Huckemann (2015) turns out to be a special case of arbitrary *power smeariness* and there is a new and different phenomenon of *logarithmic smeariness* that simultaneously features limiting behaviors of two different regimes. Then, we rigorously define FSS in terms of power smeariness, explore its two qualitatively different types and provide for an estimator of the *scale of FSS* including a statistical significance test. We continue with exploring the quantile based test by Bhattacharya and Patrangenaru (2005) and an implementation of the bootstrap test, which was also suggested by Bhattacharya and Patrangenaru (2005). We corroborate via simulations that the bootstrap test preserves the nominal size, also in the presence of FSS. Then we apply both tests to climate data. We conclude with an outlook to higher dimensions and list open problems that arise from our findings.

In the following we consider only distributions with unique population means μ . In case of nonunique sample means, we agree to pick a unique element from a realisation of an independent uniform variable on the set of sample means. Obviously, this is a measurable selection.

2 Smeariness on the Circle and the Torus

The *circle* is the space $\mathbb{S}^1 = [-\pi, \pi)$ with $-\pi$ and π identified and the usual arc length distance

$$d(x, y) = \min \{|x - y|, 2\pi - |x - y|\}.$$

Theorem 2.1 (Central Limit Theorem on the Circle). *Let $n \in \mathbb{N}$ and consider i.i.d. random elements $X_1, \dots, X_n \sim X$ on \mathbb{S}^1 with unique Fréchet population mean $\mu = 0$, Fréchet sample mean $\hat{\mu}_n$ and Euclidean variance $\mathbb{E}[X^2]$, that feature a continuous density f on $(\pi - \delta, \pi) \cup [-\pi, -\pi + \delta)$ w.r.t. the arc length measure for some $\delta > 0$. Then, for $n \rightarrow \infty$, we have*

(i) *in case of $f(-\pi) = \lim_{x \searrow -\pi} f(x) = \lim_{x \nearrow \pi} f(x) < \frac{1}{2\pi}$, that*

$$\sqrt{n} \hat{\mu}_n \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{\mathbb{E}[X^2]}{(1 - 2\pi f(-\pi))^2} \right)$$

(ii) *in case of $f(-\pi + \epsilon) = \frac{1}{2\pi} - G'(\epsilon) + o(G'(\epsilon))$ and $f(\pi - \epsilon) = \frac{1}{2\pi} - G'(\epsilon) + o(G'(\epsilon))$, with a non-negative, locally at 0 strictly convex, continuously differentiable function $G: [0, \delta) \rightarrow \mathbb{R}_{\geq 0}$ fulfilling $G(0) = G'(0) = 0$, that*

$$\sqrt{n} \operatorname{sign}(\hat{\mu}_n) 2\pi G(|\hat{\mu}_n|) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbb{E}[X^2]).$$

Proof. Assertion (i) has been shown by McKilliam et al. (2012); Hotz and Huckemann (2015). For Assertion (ii), let $0 < x < \delta$ and consider the empirical Fréchet function from (2)

$$\begin{aligned} F_n(x) &= \frac{1}{n} \sum_{X_j \in [x-\pi, \pi)} (X_j - x)^2 + \frac{1}{n} \sum_{X_j < x-\pi} (X_j + 2\pi - x)^2 \\ &= \frac{1}{n} \sum_{j=1}^n (X_j - x)^2 + \frac{4\pi}{n} \sum_{X_j < x-\pi} (X_j - x + \pi) \end{aligned}$$

In consequence, with the Euclidean mean $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$, almost surely,

$$\frac{1}{2} \text{grad } F_n(x) = x - \bar{X} - \frac{2\pi}{n} \sum_{X_j < x - \pi} 1. \quad (4)$$

Since

$$\begin{aligned} \mathbb{E} \left[\frac{1}{n} \sum_{X_j < x - \pi} 1 \right] &= \int_{-\pi}^{x - \pi} f(t) dt = \int_{-\pi}^{x - \pi} \left(\frac{1}{2\pi} - G'(\pi + t) + o(G'(\pi + t)) \right) dt \\ &= \frac{x}{2\pi} - G(x) + o(G(x)), \end{aligned}$$

where the last equality is due to the properties of G , we have for the sum of Bernoulli variables

$$\frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{X_j < x - \pi\}} = \frac{x}{2\pi} - G(x) + o(G(x)) + O_p \left(\frac{\sqrt{|x - G(x)|}}{\sqrt{n}} \right).$$

In conjunction with the analog argument for $-\delta < x < 0$ we thus obtain from (4) that

$$\frac{\sqrt{n}}{2} \text{grad } F_n(x) = \sqrt{n}(\text{sign}(x) 2\pi G(|x|) - \bar{X}) + o(G(x)) + O_p \left(\frac{\sqrt{|x + G(|x|)|}}{\sqrt{n}} \right). \quad (5)$$

This yields Assertion (ii) since the above l.h.s. vanishes by definition at $x = \hat{\mu}_n$, $\hat{\mu}_n = o_p(1)$ due to Ziezold (1977), and since $\sqrt{n}\bar{X} \xrightarrow{D} \mathcal{N}(0, \mathbb{E}[X^2])$ by the classical Euclidean central limit theorem. \square

Remark 2.2. *The above theorem gives a rather general version for the CLT on \mathbb{S}^1 , balancing between simplicity and technical generality. In particular, $f(-\pi) > \frac{1}{2\pi}$ is not possible due to Hotz and Huckemann (2015, Theorem 1). As detailed below, special cases of Theorem 2.1 are power smeariness and logarithmic smeariness.*

Our assumption on local strict convexity of G at 0 is equivalent to G' being a strictly monotone increasing function in a neighborhood around 0. In addition, by $G(0) = 0$ we make sure that $\int_0^\epsilon G'(x) dx = G(\epsilon)$ and that $G'(\epsilon) > 0$ for all $0 < \epsilon < \delta$.

Example 2.3 (Power Smeariness). *For arbitrary $r > 0$ consider*

$$\begin{aligned} d\mathbb{P}^r(x) &= \frac{\pi r + \pi - r}{\pi r + \pi} \mathbb{1}_{[-1/2, 1/2]}(x) dx + \frac{1}{2\pi} \cdot \left(1 - (\pi - x)^r \right) \mathbb{1}_{[\pi-1, \pi]}(x) dx \\ &\quad + \frac{1}{2\pi} \cdot \left(1 - (\pi + x)^r \right) \mathbb{1}_{[-\pi, -\pi+1]}(x) dx. \end{aligned}$$

Symmetry and Hotz and Huckemann (2015, Proposition 1) then yield a unique Fréchet population mean at $\mu = 0$ with $G(\epsilon) = \frac{\epsilon^{r+1}}{2\pi(r+1)}$ in $[0, 1)$. In consequence

$$n^{\frac{1}{2(r+1)}} \hat{\mu}_n$$

has a nontrivial limiting distribution, so that $\hat{\mu}_n$ is power smeary of order r . A special case is power smeariness of integer order $r \in \{0, 1, 2, \dots\}$ as derived in Hotz and Huckemann (2015).

Corollary 2.4 (Logarithmic Smeariness). *For arbitrary $r > 0$ and suitable $c_r > 0$ consider $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} X$ on \mathbb{S}^1 following the probability distribution*

$$\begin{aligned} d\mathbb{P}_r(x) &= c_r \cdot \mathbb{1}_{[-1/2, 1/2]}(x) dx + \mathbb{1}_{[\pi-1/2, \pi]}(x) \frac{\max\{1 - G'_r(\pi - x), 0\}}{2\pi} dx \\ &\quad + \mathbb{1}_{[-\pi, -\pi+1/2]}(x) \frac{\max\{1 - G'_r(\pi + x), 0\}}{2\pi} dx, \end{aligned}$$

where $G_r(x) = \exp(-1/|x|^r) \mathbb{1}_{(0, \infty)}(x)$ is a smooth function. Then

(i) $\mu = 0$ is the unique Fréchet population mean, and

(ii) for every Fréchet sample mean $\hat{\mu}_n$ we have

(a) $\hat{\mu}_n \xrightarrow{\text{a.s.}} 0$,

(b) $(\log \sqrt{n})^{1/r} \hat{\mu}_n \xrightarrow{\mathcal{D}} \frac{1}{2} d\delta_{-1} + \frac{1}{2} d\delta_1$,

(c) $r (\log \sqrt{n})^{(1+r)/r} \left(\hat{\mu}_n - \frac{\text{sign } \hat{\mu}_n}{(\log \sqrt{n})^{1/r}} \right) \xrightarrow{\mathcal{D}} \text{sign } Z \cdot \log |Z|$ with $Z \sim \mathcal{N}\left(0, \frac{\mathbb{E}[X^2]}{4\pi^2}\right)$.

The proof is deferred to the appendix.

Remark 2.5. These results extend at once to the m -torus $\mathbb{T}^m = (\mathbb{S}^1)^m = \prod_{i=1}^m [-\pi, \pi)$, $m \in \mathbb{N}$, equipped with the canonical product metric

$$d_{\mathbb{T}^m}(x, y) = \sqrt{\sum_{i=1}^m d(x^{(i)}, y^{(i)})^2}, \quad x = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(m)} \end{pmatrix}, \quad y = \begin{pmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{pmatrix} \in \mathbb{T}^m,$$

because $\mu \in \mathbb{T}^m$ is a minimizer of the population Fréchet function (1) of a random variable X on \mathbb{T}^m if and only if all of its coordinates $\mu^{(i)}$ are minimizers of the population Fréchet functions of the marginals $X^{(i)}$ on the i -th circle, $i \in \{1, \dots, m\}$. Notably, every marginal may feature different smeariness.

In consequence, due to Hotz and Huckemann (2015, Corollary 3), for a sample $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} X$ there are at most n^m minimizers of the sample Fréchet function (2). In particular every one of the n^m grid points

$$\left(\left(-\pi + \frac{2\pi(2j_1 - 1)}{2n} \right), \dots, \left(-\pi + \frac{2\pi(2j_m - 1)}{2n} \right) \right), \quad j_1, \dots, j_m \in \{1, \dots, n\}$$

is a minimizer of the sample Fréchet function of $X_j = \left(-\pi + \frac{2\pi(j-1)}{n} \right) (1, \dots, 1)^T \in \mathbb{T}^m$, $j \in \{1, \dots, n\}$, cf. Figure 2. Indeed, their i -th marginal ($1 \leq i \leq m$) is given by the sample $X_j^{(i)} = \left(-\pi + \frac{2\pi(j-1)}{n} \right)$, $j \in \{1, \dots, n\}$ which has exactly n minimizers of its sample Fréchet function which are equally spaced between the sample points.

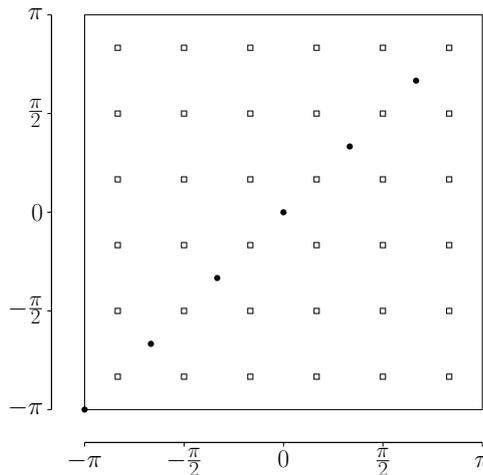


Figure 2: Six sample points (black circles) on the diagonal of the two-torus \mathbb{T}^2 and their equally spaced $6^2 = 36$ Fréchet sample means (empty squares), cf. Remark 2.5.

3 Finite Sample Smeariness (FSS)

For data sampled from nonsmeary (order $r = 0$ smeariness) distributions that are in some sense near smeary distributions, for rather large sample sizes n , a rate of $d(\hat{\mu}_n, \mu)^2$ considerably lower than the usual asymptotic $1/n$ rate has been observed by Eltzner and Huckemann (2019) on circles and spheres of arbitrary dimension.

Definition 3.1. *The Fréchet population mean μ of a random element X on a metric space (M, d) is finite sample smeary (FSS) if there are constants $C_+, C_- > 0$, $K > 0$, $0 < r_- < r_+$ and integers $1 \leq n_- < n_+ < n_0$ satisfying the condition $C_+ n_-^{\frac{r_+}{r_++1}} \leq C_- n_+^{\frac{r_-}{r_-+1}}$ and satisfying for Fréchet sample means $\hat{\mu}_n$ that*

$$(i) \forall n \in (n_-, n_+] \cap \mathbb{N} : \quad \mathbb{E} \left[d(X, \mu)^2 \right] \leq C_- n^{\frac{r_-}{r_-+1}} \leq n \mathbb{E} \left[d(\hat{\mu}_n, \mu)^2 \right] \leq C_+ n^{\frac{r_+}{r_++1}}.$$

$$(ii) \forall n \in (n_0, \infty) \cap \mathbb{N} : \quad \mathbb{E} \left[d(\hat{\mu}_n, \mu)^2 \right] \leq K n^{-1}.$$

Further, if there are $C > 0, n_0 \leq n_C \in \mathbb{N}$ such that

$$(iii) \forall n \in (n_C, \infty) \cap \mathbb{N} : \quad \mathbb{E} \left[d(\hat{\mu}_n, \mu)^2 \right] > \left(\mathbb{E} \left[d(X, \mu)^2 \right] + C \right) n^{-1} \text{ then it is Type I FSS,}$$

(iv) else it is Type II FSS.

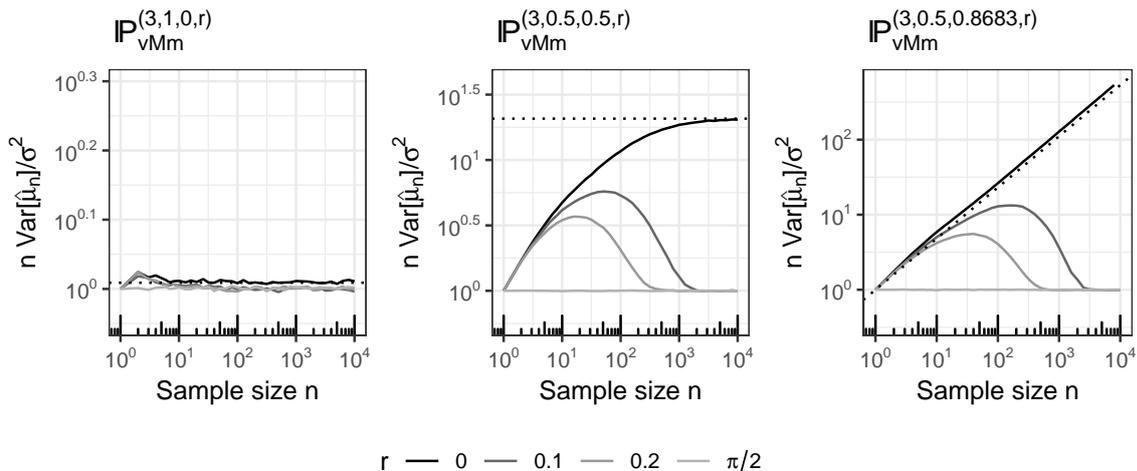


Figure 3: Log-log plots of curves of scaled variances (also divided by Euclidean variance σ^2), cf. Figure 1, for varying von Mises mixture distributions (vMmds) defined in (9) in black ($r = 0$) and vMmds with a disk cut out of radius $r = 0.1, 0.2, \pi/2$ (in fading gray) about the antipode of the mean, as defined in (10). In the left and middle display, the vMmds (black) feature Type I FSS, with the dotted line giving the asymptotic scaled variance, the two vMmds with a disk cut out (gray curves) of radius $0 < r < \pi/2$ feature Type II FSS. In the right display, the vMmd (black) features smeariness, with theoretical scaled variance (dotted), the vMmds with a disk cut out (gray curves) of radius $0 < r < \pi/2$ feature Type II FSS. Only the vMmds with a disk cut out of size $r = \pi/2$ (light gray) features no FSS at all.

The behavior of the scaled variance $n \mathbb{E} [d(\hat{\mu}_n, \mu)^2]$ in case of FSS is sketched in Figure 1. For $n_- \leq n \leq n_+$, Definition (i) requires $n \mathbb{E} [d(\hat{\mu}_n, \mu)^2]$ to start above the Euclidean variance $\sigma^2 = \mathbb{E} [(d(X, \mu))^2]$ of the classical Euclidean CLT and to increase at least with some polynomial power $n^{r_-/(r_-+1)} > 0$ that

does not exceed (ensured by the condition) $n^{r+/(r+1)} > 0$. Asymptotically, guaranteed by (ii), the rate of $\mathbb{E}[d(\widehat{\mu}_n, \mu)^2]$ is then the classical $1/n$, ruling out smeariness.

Simple cases of FSS on the circle are illustrated for some nonaltered (Type I FSS), introduced in (9), and altered (Type II FSS) introduced in (10), von Mises mixtures in Figure 3. Indeed, in conjunction with Theorem 2.1 on the circle, Type I FSS occurs if X features a density f nonvanishing at the antipode $\bar{\mu}$ of its unique intrinsic mean μ , cf. Definition 3.1 (iii). If the density vanishes at the antipode, Type II FSS occurs, Definition 3.1 (iv). So in case of Type I FSS, the curve $\mathbb{E}[(d(\widehat{\mu}_n, \mu))^2]$ rises from σ^2 to $\sigma^2(1 - 2\pi f(\bar{\mu}))^{-2}$. In case of Type II FSS, the curve first rises from σ^2 and eventually drops down to σ^2 .

Depending on the probability distribution near $\bar{\mu}$, also more complicated versions of increase and decrease may occur, as Example 3.2 and Figure 4 teach: every pair of bumps of the density near the antipode may result in a bump of the scaled variance. As before, however, it starts at σ^2 and eventually settles at $\sigma^2(1 - 2\pi f(\bar{\mu}))^{-2}$, producing Type I FSS if $f(\bar{\mu}) > 0$ and Type II FSS else.

Example 3.2. Investigating the relationship between the scaled variance of intrinsic sample means $n \mapsto n\mathbb{E}[d(\widehat{\mu}_n, \mu)^2]$ and the density $f^{(t,w)}$ of X near the antipode of the intrinsic population mean $\mu = 0$, let $l \in \mathbb{N}$, $w = (w_1, \dots, w_l) \in [0, 1]^l$, $t = (t_1, \dots, t_l) \in [0, \pi]^l$ with $t_0 := 0 < t_1 < \dots < t_l < \pi$ and define the distribution $\mathbb{P}_U^{(t,w)}$ by

$$d\mathbb{P}_U^{(t,w)}(x) := k \cdot d\delta_0(x) + \frac{1}{2\pi} \cdot f^{(t,w)}(x)dx \quad \text{with}$$

$$f^{(t,w)}(x) := \begin{cases} w_i & \text{if } x \in [-\pi + t_{i-1}, -\pi + t_i] \cup (\pi - t_i, \pi - t_{i-1}] \text{ for some } i \in \{1, \dots, l\} \\ 0 & \text{if } x \in [-\pi + t_l, \pi - t_l], \end{cases}$$

where $k = k(t, w) > 0$ is a normalization constant to ensure that $\mathbb{P}_U^{(t,w)}$ is a probability measure. In Table 2 we list cases (a) – (e) of parameter choices considered.

	(a)	(b)	(c)	(d)	(e)
l	1	2	4	2	4
t	1.5	(0.8, 2)	(0.1, 0.2, 0.5, 2)	(0.8, 2)	(0.1, 0.2, 0.5, 2)
w	0.5	(0.5, 1)	(0.5, 0.8, 0.0, 1)	(0.0, 1)	(0.0, 0.85, 0.0, 1)
Type of FSS	I	I	I	II	II

Table 2: Selected values for the parameters l, t, w and their resulting type of FSS.

In all cases the population mean of $\mathbb{P}_U^{(t,w)}$ is unique and located at $\mu = 0$. Whenever the density at the antipode is strictly between zero and $1/2\pi$ we have FSS of Type I. Regardless of the type of FSS, every pair of bumps in the density near the antipode corresponds to a single bump in the scaled variance curve. In case of Type I FSS the curve approaches asymptotically a value strictly above the Euclidean variance, in case of Type II FSS it approaches asymptotically the Euclidean variance. Numerical experiments indicate that the convergence occurs from below in case of Type I and from above in case of Type II.

Remark 3.3. As Example 3.2 and Figure 4 teach, a distribution may feature FSS of different r_-, r_+ on different intervals n_-, n_+ .

On the circle we have a complete characterization of FSS.

Theorem 3.4. Let X be a random variable on the circle \mathbb{S}^1 with unique Fréchet population mean $\mu = 0$. Then the following hold.

- (i) If the support of X is contained in $(-\pi/2, \pi/2)$ then μ is not finite sample smeary.
- (ii) If the support of X is contained in $[-\pi/2, \pi/2]$ and if X features a density near $\pm\pi/2$, then μ is not finite sample smeary.

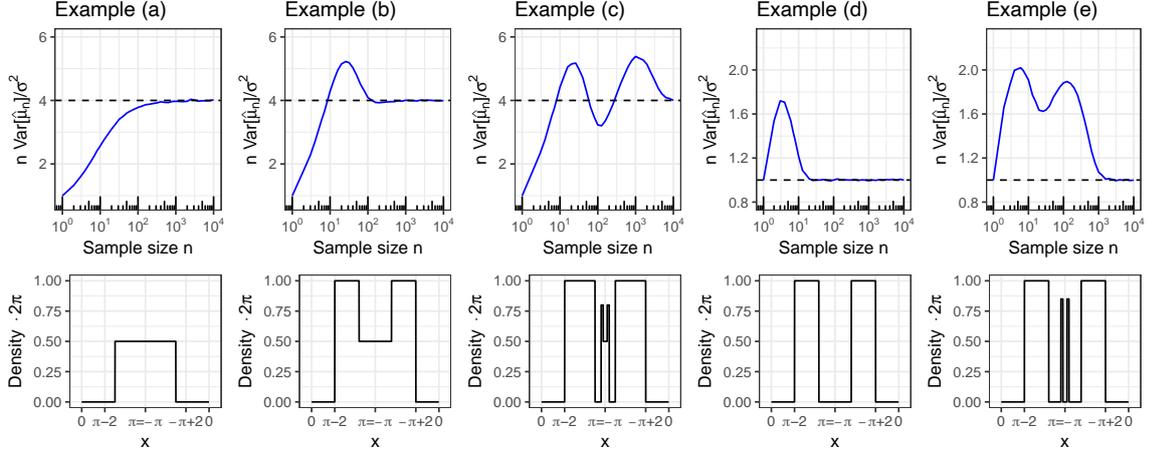


Figure 4: Top row: Curves of scaled variances (also divided by Euclidean variance σ^2), as in Figure 3 (here only the horizontal is in log-scale), for each distribution in Table 2 of Example 3.2. Bottom row: Density part of the corresponding distributions (that comprise a δ -measure at the origin).

(iii) If $\mathbb{P}^X[-\pi, -\pi/2] \neq 0 \neq \mathbb{P}^X[\pi/2, \pi]$ then μ is finite sample smeary if it is not smeary.

Proof. Let $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} X$. In cases (i) and (ii) we have a.s. that $d(X_i, X_j) = |X_i - X_j|$ for all $1 \leq i, j \leq n$, so that $-\pi/2 < \hat{\mu}_n = \frac{1}{n} \sum_{j=1}^n X_j < \pi/2$ a.s., giving $\mathbb{E}[d(\hat{\mu}_n, \mu)^2] = \frac{1}{n} \mathbb{E}[X^2]$ leading to the absence of FSS.

In case (iii), set $n_- = 1$ and $n_+ = 2$, $\bar{X} = \frac{1}{2}(X_1 + X_2)$ and $\bar{x} = \frac{1}{2}(x_1 + x_2)$. Also denote

$$\begin{aligned} A_0 &:= \{x_1, x_2 \in \mathbb{S}^1 : |x_1 - x_2| < \pi\} \\ A_+ &:= \left\{x_1, x_2 \in \mathbb{S}^1 : |x_1 - x_2| \geq \pi, -\frac{\pi}{2} < \bar{x} < 0\right\} \\ A_- &:= \left\{x_1, x_2 \in \mathbb{S}^1 : |x_1 - x_2| \geq \pi, 0 \leq \bar{x} < \frac{\pi}{2}\right\} \end{aligned}$$

and $d\mathbb{P}^X(x_1, x_2) := d\mathbb{P}^X(x_1) d\mathbb{P}^X(x_2)$. Then (since there can be no mass of X at $x = -\pi$, due to Hotz and Huckemann (2015, Theorem 1))

$$\begin{aligned} \mathbb{E}[d(\hat{\mu}_2, \mu)^2] &= \int_{A_0} \bar{x}^2 d\mathbb{P}^X(x_1, x_2) + \int_{A_+} (\bar{x} + \pi)^2 d\mathbb{P}^X(x_1, x_2) + \int_{A_-} (\bar{x} - \pi)^2 d\mathbb{P}^X(x_1, x_2) \\ &= \frac{1}{2} \mathbb{E}[d(X, \mu)^2] + \pi \left(\int_{A_+} (\pi + 2\bar{x}) d\mathbb{P}^X(x_1, x_2) + \int_{A_-} (\pi - 2\bar{x}) d\mathbb{P}^X(x_1, x_2) \right) \end{aligned}$$

By construction, both integrands above are positive. Hence, by hypothesis the two integrals are positive, yielding $n\mathbb{E}[d(\hat{\mu}_n, \mu)^2] > \mathbb{E}[d(X, \mu)^2]$ for (all) integer $2 = n \in (n_-, n_+]$, giving FSS in case of nonsmearyness. \square

The following example illustrates Theorem 3.4 with two point masses at or beyond the equator (case (iii)), or before (case (i)), possibly (in case (iii)), a nonunique sample mean set, of which uniformly at random, a sample mean is selected.

Example 3.5. Letting $\epsilon \in [-\pi/2, \pi/2)$ and $w \in (0, 1/4]$, define the circular distribution $\mathbb{P}_E^{(\epsilon, w)}$ by

$$d\mathbb{P}_E^{(\epsilon, w)}(x) := (1 - 2w) \cdot \mathbb{1}_{[-1/2, 1/2]}(x) dx + w d\delta_{\pi/2 + \epsilon}(x) + w d\delta_{-\pi/2 - \epsilon}(x),$$

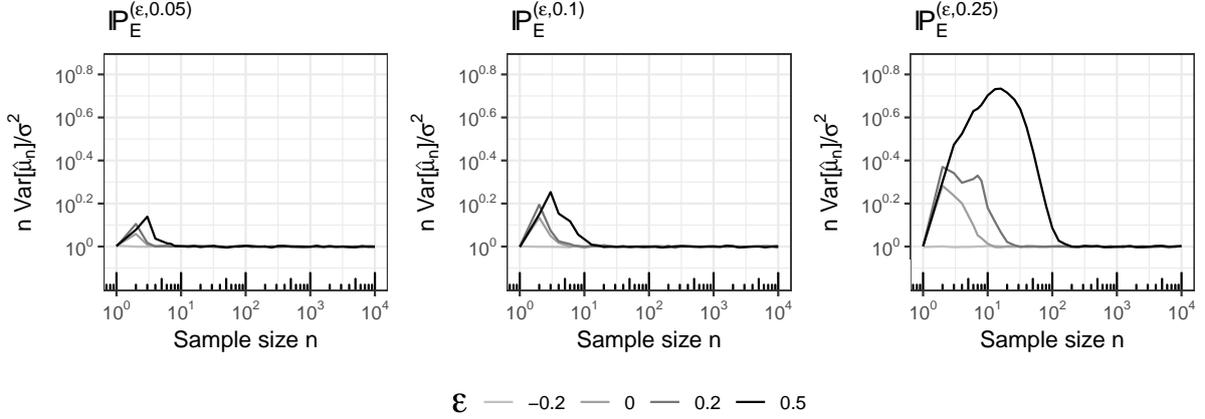


Figure 5: Curves of scaled variances (also divided by Euclidean variance σ^2), as in Figure 3, for the three types of distributions $\mathbb{P}_E^{(\epsilon, w)}$ from Example 3.5 with weight w of point mass at $\pm(\pi/2 + \epsilon)$. Black: $\epsilon = 0.5$, dark gray: $\epsilon = 0.2$, gray: $\epsilon = 0$, light gray: $\epsilon = -0.2$. We have Type II FSS for all values of $\epsilon \geq 0$ and no FSS (constant scaled variance) for $\epsilon < 0$.

which assigns at least half the mass to $[-1/2, 1/2]$ and the rest is evenly distributed close to the equator at $\pi/2 + \epsilon$ and $-\pi/2 - \epsilon$.

For $\epsilon < 0$ these distribution are supported in $(-\pi/2, \pi/2)$ and thus by Theorem 3.4 (i) feature no FSS.

For $\epsilon > 0$, in contrast by Theorem 3.4 (iii), they always feature FSS, which is, by Theorem 2.1 (i) of Type II.

For $\epsilon = 0$, by the same argument, they also feature FSS of Type II. Indeed, samples featuring only points at $\pm\pi/2$ and no others, occurring with positive probability, have nonunique sample means, namely one closer to 0 and one closer to $-\pi$. As agreed, of these with probability 1/2 the one closer to $-\pi$ is chosen which accounts for a higher variance than the Euclidean.

Each panel in Figure 5 illustrates the three cases, with larger effect on FSS the larger the weight w of each of the point masses.

Thus far, we have defined FSS for arbitrary metric spaces. However, we have investigated FSS only for the circle. These results extend at once to the torus, as in the corollary below. For other spaces, a generalization is beyond the scope of this paper. Already for spheres, finding examples of smeary distributions is rather involved, and to date, only two-smeariness could be confirmed by Eltzner and Huckemann (2019), and they left no doubt that Type I FSS also occurs on spheres. While in the light of *geometrical smeariness* introduced by Eltzner (2019), it seems dubious that Type II FSS also exists on spheres, it seems that Type I FSS is present on all positively curved spaces. For these Afsari (2009) has shown that the Hessian of the Fréchet function is smaller than its Euclidean equivalent, which is twice the identity matrix. The CLT by Bhattacharya and Patrangenaru (2005) therefore suggests that FSS is always present on such spaces except for single point masses. Indeed, corresponding results, even for small sample size and highly concentrated random variables have recently been derived by Pennec (2019). Since the magnitude of FSS seems very small in these highly concentrated cases, below we derive a heuristic for the effect size of FSS.

Corollary 3.6. *Let $m \in \mathbb{N}$ and let $X = (X^{(1)}, \dots, X^{(m)})$ be a random variable on the torus \mathbb{T}^m with unique Fréchet population mean $\mu = 0$. Then the following hold.*

- (i) *If the support of X is contained in $(-\pi/2, \pi/2)^m$ then μ is not finite sample smeary.*

- (ii) If the support of X is contained in $[-\pi/2, \pi/2]^m$ and if every marginal $X^{(i)}$ ($1 \leq i \leq m$) features a density near $\pm\pi/2$, then μ is not finite sample smeary.
- (iii) If for a marginal $\mathbb{P}^{X^{(i)}}[-\pi, -\pi/2] \neq 0 \neq \mathbb{P}^{X^{(i)}}[\pi/2, \pi]$ ($1 \leq i \leq m$) then μ is finite sample smeary if it is not smeary.

In order to heuristically assess the effect of FSS consider random variables $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} X$ on a metric space (Q, d) with unique population mean μ and unique sample mean $\hat{\mu}_n$ (in case of nonuniqueness, take a uniformly random selection as agreed at the end of Section 1) and set

$$V := F(\mu), \quad \hat{V}_n := F_n(\hat{\mu}_n), \quad W := \mathbb{E}[d(\mu, X)^4], \quad \hat{W}_n := \frac{1}{n} \sum_{j=1}^n d(\hat{\mu}_n, X_j)^4, \quad (6)$$

where we assume that the 4-th moment exists. Then, under mild additional assumptions the CLT for the Fréchet variance by Dubey and Müller (2019) teaches that

$$\sqrt{n}(\hat{V}_n - V) \xrightarrow{D} \mathcal{N}(0, W - V^2),$$

where $\hat{W}_n - \hat{V}_n^2$ is an asymptotically unbiased estimator for $W - V^2$, that also satisfies \sqrt{n} asymptotic normality.

Further, let $B > 0$ be a large integer and μ_b^* be a mean of an n -out-of- n bootstrap sample of X_1, \dots, X_n for $b = 1, \dots, B$. Further, let $\hat{\mu}^*$ be the mean of the μ_1^*, \dots, μ_B^* and set

$$\hat{V}^* := \frac{1}{B} \sum_{b=1}^B d(\hat{\mu}^*, \mu_b^*)^2, \quad \hat{W}^* := \frac{1}{B} \sum_{b=1}^B d(\hat{\mu}^*, \mu_b^*)^4. \quad (7)$$

Then, under absence of smeariness and absence of FSS, $\sqrt{B}(\hat{V}^* - V_n/n)$ is approximately normal with zero expectation and its variance can be estimated by $\hat{W}^* - (\hat{V}^*)^2$. Hence, in this case, with the standard-normal α -quantile ϕ_α we expect that

$$\mathbb{P} \left\{ \hat{V}^* > \frac{V_n}{n} + \phi_{1-\alpha} \sqrt{\frac{\hat{W}^* - (\hat{V}^*)^2}{B}} \right\} \approx \alpha.$$

Test 3.7. For random variables X_1, \dots, X_n on a metric space (Q, d) with sample mean $\hat{\mu}_n$, bootstrap means μ_1^*, \dots, μ_B^* , and their mean $\hat{\mu}^*$, as above and the notation from (6) and (7),

$$S_{\text{FSS}, n} := \frac{n\hat{V}^*}{\hat{V}_n} \quad (8)$$

denotes the scale of FSS for sample size n . With the standard normal α -quantile ϕ_α , $\alpha \in [0, 1]$, reject the absence of FSS for sample size n at nominal level α if

$$S_{\text{FSS}, n} - 1 > h_{n, \alpha} \text{ where } h_{n, \alpha} = \frac{n\phi_\alpha}{\sqrt{B}} \frac{\sqrt{\hat{W}^* - (\hat{V}^*)^2}}{\hat{V}_n}.$$

Figure 6 shows simulations for distributions $\mathbb{P}_{vMm}^{(3, \beta, \lambda, r)}$ featuring Type I FSS for $r = 0$ as defined in (9) and Type II FSS for $r > 0$ as defined in (10). Indeed for $r = \pi/2$, the null hypothesis of absence of FSS is true and Test 3.7 keeps the level $\alpha = 0.05$. For Type II FSS, rejection probabilities are almost one within the regime of FSS (compare with Figure 3). Beyond that regime the null hypothesis is true again and detected with the correct size. For the von Mises mixture distribution (vMmd) consisting of only one von Mises distribution ($\beta = 1$, left panel, black) with almost no FSS visible in Figure 3, rejection probabilities increase visibly with sample size. As expected, the power of Test 3.7 increases with proximity to a smeary distribution ($\beta = 0.5, \lambda = 0.8683, r = 0$, right panel, black), and specifically in case of Type II FSS also with smaller hole size and proximity to a distribution featuring Type I FSS.

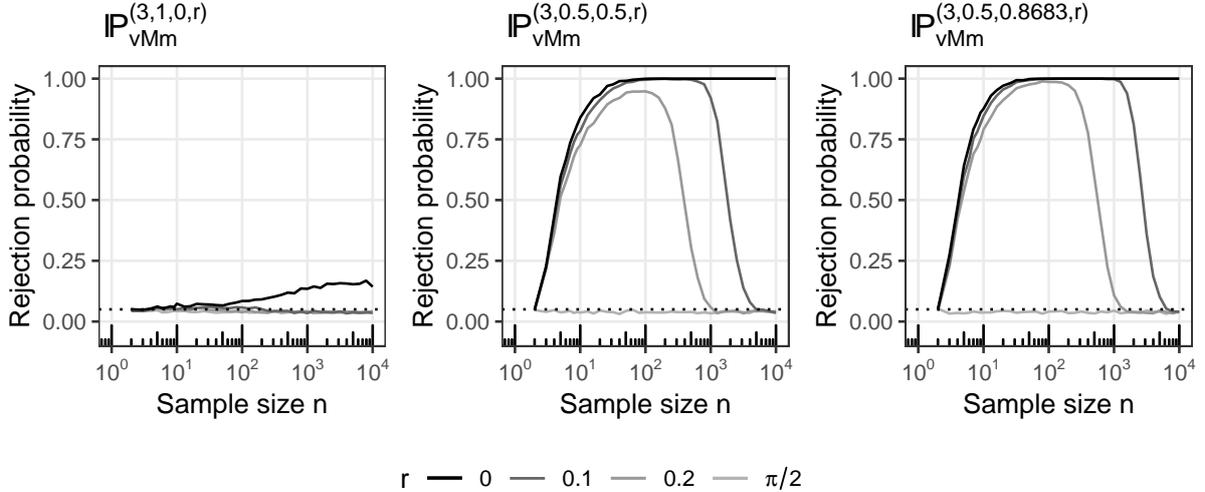


Figure 6: Empirical rejection probability curves for Test 3.7 with nominal size $\alpha = 0.05$ (dotted horizontal) for the vMmds from Figure 3. For the vMmds with a disk cut out of size $r = \pi/2$ about the antipode of its mean, which feature no FSS, all of the curves (light gray) keep the level. For Type II FSS (gray: $r = 0.2$, dark gray: $r = 0.1$) rejection probabilities are almost one within the regime of FSS; beyond, where there is no FSS any more, they drop back to the nominal size (see Figure 3). Also for the vMmd consisting of only one vMm (left panel, black) with almost no FSS visible, rejection probabilities increase visibly with sample size. For Type I FSS (middle, black) and smeariness (right, black), the rejection probabilities remain one also for higher sample sizes.

4 One- and Two-Sample Tests for Fréchet Means under FSS

We begin with a brief review of the celebrated central limit theorem (CLT) by Bhattacharya and Patrangenaru (2005) for a k -dimensional manifold M and corresponding tests proposed therein. The CLT states that, under suitable conditions (further clarified in Bhattacharya and Lin (2017); Eltzner and Huckemann (2019); Eltzner et al. (2019)), in a local chart $\phi : U \rightarrow \mathbb{R}^k$, $U \subset M$, the fluctuation $\phi(\hat{\mu}_n) - \phi(\mu)$ of the Fréchet sample mean $\hat{\mu}_n$ about the Fréchet population mean μ , rescaled with the square root of sample size \sqrt{n} , is asymptotically (for $n \rightarrow \infty$) Gaussian with zero mean and covariance given by

$$\Sigma = H^{-1}CH^{-1}.$$

Here, C is the population covariance of the gradient of the Fréchet function F from (1) at μ in the local chart and H is twice the expected value of the Hessian of the Fréchet function at μ in the local chart. One of the above mentioned conditions is that H be positive definite. In our language, this means that $\hat{\mu}_n$ is nonsmeary (i.e., on the circle we have the situation of Theorem 2.1 (i) where $H = 1 - 2\pi f(-\pi)$, as can be derived at once from the argument following (4)).

For mutually independent samples $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} X$ and $Y_1, \dots, Y_n \stackrel{\text{i.i.d.}}{\sim} Y$ with population Fréchet means $\mu^{(X)}$ and $\mu^{(Y)}$, respectively, and an additional $\mu_0 \in M$, consider the hypotheses

$$\begin{aligned} H_0^1 &: \mu^{(X)} = \mu_0 && \text{for the one-sample test,} \\ H_0^2 &: \mu^{(X)} = \mu^{(Y)} && \text{for the two-sample test.} \end{aligned}$$

Quantile based tests. With a local chart $\phi : U \rightarrow \mathbb{R}^k$ where $U \subset M$ contains μ (for H_0^1), or $\mu^{(X)}$ and $\mu^{(Y)}$ (for H_0^2), respectively, Bhattacharya and Patrangenaru (2005) thus suggest to consider the following

test statistics,

$$\begin{aligned} T^1 &= n (\phi(\widehat{\mu}_n^{(X)}) - \phi(\mu)) ^T \widehat{H}_X \widehat{C}_X^{-1} \widehat{H}_X (\phi(\widehat{\mu}_n^{(X)}) - \phi(\mu)) && \text{for the one-sample test,} \\ T^2 &= (n+m) (\phi(\widehat{\mu}_n^{(X)}) - \phi(\widehat{\mu}_m^{(Y)})) ^T \widehat{H}_{X,Y} \widehat{C}_{X,Y}^{-1} \widehat{H}_{X,Y} (\phi(\widehat{\mu}_n^{(X)}) - \phi(\widehat{\mu}_m^{(Y)})) && \text{for the two-sample test,} \end{aligned}$$

and use χ_k^2 as their asymptotic approximation under the respective null hypothesis. Here $\widehat{H}_X, \widehat{C}_X$ or $\widehat{H}_{X,Y}, \widehat{C}_{X,Y}$ are the usual plugin estimators of H and C based on the first sample, or the pooled sample, respectively, cf. also Bhattacharya and Bhattacharya (2012, Section 5.4.1).

Test 4.1 (Bhattacharya and Patrangenaru (2005); Bhattacharya and Bhattacharya (2012)). *For $0 \leq \alpha \leq 1$ and the $1 - \alpha$ quantile $\chi_{k,1-\alpha}^2$ of the χ^2 distribution, reject at nominal level α*

- (i) H_0^1 , if $T^1 \geq \chi_{k,1-\alpha}^2$,
- (ii) H_0^2 , if $T^2 \geq \chi_{k,1-\alpha}^2$.

Asymptotically, while both tests are independent of the chart chosen, in general, they do not keep the level promised.

Proposition 4.2. *On the circle, in case of Type I FSS, the Tests 4.1 have asymptotically a size strictly higher than their nominal size.*

Proof. Indeed with the notation of case (i) of Theorem 2.1, if $M = \mathbb{S}^1$ is the circle, $\mu_0 = 0$, or $\widehat{\mu}^{(X)} = 0 = \widehat{\mu}^{(Y)}$, respectively, and ϕ the identity on $(-\pi, \pi) \subset \mathbb{S}^1$, plugin estimators for $H^{-1}CH^{-1}$ are $\frac{1}{n} \sum_{j=1}^n X_j^2$, or $\frac{1}{n+m} \sum_{j=1}^n X_j^2 + \frac{1}{n+m} \sum_{j=1}^m Y_j^2$, respectively, cf. Eltzner and Huckemann (2019), whereas $H^{-1}CH^{-1} = \mathbb{E}[X^2](1-2\pi f(-\pi))^{-2}$, or $(\gamma \mathbb{E}[X^2] + (1-\gamma) \mathbb{E}[Y^2])(1-2\pi f(-\pi))^{-2}$, with $n/(n+m) \rightarrow \gamma$, respectively, in case (i) of Theorem 2.1. Hence, with these estimators, in case of Type I FSS, i.e., $0 < f(-\pi) < \frac{1}{2\pi}$, the asymptotic true level of Tests 4.1 is strictly higher than their nominal level. \square

Remark 4.3. *On the circle, in case of Type II FSS, Tests 4.1 may also have a true size higher than their nominal size in a range up to rather high sample sizes, as the following simulations in Section 5 show.*

Bootstrap based tests have also been proposed by Bhattacharya and Patrangenaru (2005); Bhattacharya and Bhattacharya (2012) where $\Sigma = H^{-1}CH^{-1}$ is estimated by bootstrapping from the samples. As they have not provided detail, in the following, we employ the bootstrap one- and two-sample tests by Eltzner and Huckemann (2017), for a given level $0 \leq \alpha \leq 1$.

Resample $B \in \mathbb{N}$ times n -out-of- n from the sample X_1, \dots, X_n and let $\mu_b^{(X),*}$ be the corresponding Fréchet sample means for $b = 1, \dots, B$. Mapping these sample means under ϕ to a Euclidean space yields the covariance estimate $\Sigma^{(X),*}$. From another round of resampling obtain new $\mu_b^{(X),*}$, set $T_b^* = (\phi(\mu_b^{(X),*}) - \phi(\widehat{\mu}_n^{(X)})) ^T (\Sigma^{(X),*})^{-1} (\phi(\mu_b^{(X),*}) - \phi(\widehat{\mu}_n^{(X)}))$, $b \in \{1, \dots, B\}$ and determine $e_{1-\alpha}^*$ such that

$$\frac{\#\{b \in \{1, \dots, B\} : T_b^* \leq e_{1-\alpha}^*\} - 1}{B} \leq 1 - \alpha \leq \frac{\#\{b \in \{1, \dots, B\} : T_b^* \leq e_{1-\alpha}^*\}}{B}.$$

Further, set

$$T^1 = (\phi(\widehat{\mu}_n^{(X)}) - \phi(\mu_0)) ^T (\Sigma^{(X),*})^{-1} (\phi(\widehat{\mu}_n^{(X)}) - \phi(\mu_0)).$$

Similarly, using m -out-of- m sampling with replacement from Y_1, \dots, Y_m , obtain a bootstrap covariance estimate $\Sigma^{(Y),*}$. Now, set $A = \Sigma^{(X),*} + \Sigma^{(Y),*}$. This choice of A , and not using the pooled variance, proves to be more robust, cf. Huckemann and Eltzner (2020). Then, from another round of sampling obtain $\mu_b^{(X),*}$ and $\mu_b^{(Y),*}$, set $d_b^{(X),*} = \phi(\widehat{\mu}_b^{(X),*}) - \phi(\widehat{\mu}_n^{(X)})$, $d_b^{(Y),*} = \phi(\widehat{\mu}_b^{(Y),*}) - \phi(\widehat{\mu}_m^{(Y)})$, define $T_b^* = (d_b^{(X),*} - d_b^{(Y),*}) ^T A^{-1} (d_b^{(X),*} - d_b^{(Y),*})$, $b \in \{1, \dots, B\}$ and determine $f_{1-\alpha}^*$ such that

$$\frac{\#\{b \in \{1, \dots, B\} : T_b^* \leq f_{1-\alpha}^*\} - 1}{B} \leq 1 - \alpha \leq \frac{\#\{b \in \{1, \dots, B\} : T_b^* \leq f_{1-\alpha}^*\}}{B}.$$

Further, set

$$T^2 = (\phi(\widehat{\mu}_n^{(X)}) - \phi(\widehat{\mu}_m^{(Y)})^T A^{-1} (\phi(\widehat{\mu}_n^{(X)}) - \phi(\widehat{\mu}_m^{(Y)})).$$

Tests 4.4 (Bootstrap Based). *With the notation above, for $0 \leq \alpha \leq 1$, reject at level α*

(i) H_0^1 , if $T^1 \geq e_{1-\alpha}^*$,

(ii) H_0^2 , if $T^2 \geq f_{1-\alpha}^*$.

Again, asymptotically, both tests are independent of the chart chosen.

Remark 4.5. *Simulations in Section 5 depicted in Figures 7 and 8 show that the Tests 4.4 keep the nominal level $\alpha = 0.05$ fairly well, in particular for Type I FSS. Upon (very) close inspection, for Type II FSS, the Tests 4.4 may be slightly too conservative, for Type I FSS too liberal, indicating that a nonsymmetric bootstrap may be more accurate. Investigating this effect is left for future research beyond the scope of this paper.*

5 Simulations

On \mathbb{S}^1 , as usual identified with $[-\pi, \pi)$, we consider von Mises mixtures with antipodal modes with respect to arc length measure:

$$d\mathbb{P}_{vMm}^{\kappa, \beta, \lambda}(x) := \beta \frac{\exp(\kappa \cos(x))}{I_0(\kappa)} dx + (1 - \beta) \frac{\exp(\lambda \cos(x + \pi))}{I_0(\lambda)} dx \quad \text{for } \kappa, \lambda \geq 0, \beta \in [0, 1], \quad (9)$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind of order 0, e.g., Mardia and Jupp (2000, p. 36). A given von Mises mixture $d\mathbb{P}_{vMm}^{\kappa, \beta, \lambda}$ may attain mean(s) at 0 or π or even at $\{-t, t\}$ for $t \in (0, \pi)$. Furthermore, we define for $r \geq 0$ the function cutting out and mirroring a disk of radius r about $-\pi$:

$$\zeta^r : \mathbb{S}^1 \rightarrow \mathbb{S}^1, \quad p \mapsto \begin{cases} p & \text{if } p \in [-\pi + r, \pi - r) \\ p + \pi & \text{if } p \in [-\pi, -\pi + r) \\ p - \pi & \text{if } p \in [\pi - r, \pi) \end{cases}$$

For the von Mises mixture $\mathbb{P}_{vMm}^{\kappa, \beta, \lambda}$ we then denote the push-forward measure under ζ^r by

$$\mathbb{P}_{vMm}^{\kappa, \beta, \lambda, r} := \zeta^r_* \mathbb{P}_{vMm}^{\kappa, \beta, \lambda}, \quad (10)$$

which preserves all mass except for that in a disk of radius r about $-\pi$, which is mirrored. Recall that by Theorem 3.4, all of the $\mathbb{P}_{vMm}^{\kappa, \beta, \lambda, r}$ with unique mean $\mu = 0$ and $r < \pi/2 - \epsilon$ for some $\epsilon > 0$ feature FSS (if they are not smeary themselves), which is of Type II if and only if $r > 0$. For the following simulations we selected parameters as described in Table 3. All of them give unique means at $\mu = 0$, none of which is smeary. Only the nearby $\mathbb{P}_{vMm}^{3, 0.5, \lambda, 0}$ is smeary (power smeary of order 2) for $\lambda \approx 0.8683$.

In Figures 7 and 8 we compare the power functions at nominal level $\alpha = 0.05$ of the quantile tests (gray lines, Test 4.1) to the power functions of the bootstrap tests (black lines, Test 4.4). For the one-sample tests (Figure 7) we consider simulations of $\mathbb{P}_{vMm}^{\kappa, \beta, \lambda, r}$ and rotated null hypotheses $\mu_0 \in [-\pi, \pi)$, for the two sample-tests (Figure 8) we consider two simulations of $\mathbb{P}_{vMm}^{\kappa, \beta, \lambda, r}$ that are rotated with respect to one another by the angle $p \in [-\pi, \pi)$. Random variables from von Mises distributions were generated using the R-package *circular* (Lund et al., 2017).

κ	3		3		3	
β	1		1/2		1/2	
λ	0		0		1/2	
r	0	0.1	0	0.1	0	0.1
$S_{FSS,30}$	1.0	1.0	3.7	3.0	7.2	5.4
$S_{FSS,100}$	1.0	1.0	4.2	2.7	11.6	6.6
$S_{FSS,300}$	1.0	1.0	4.4	2.0	15.9	5.6

Table 3: Parameters of $\mathbb{P}_{vM}^{\kappa,\beta,\lambda,r}$ with scales of FSS for simulations for the performance of Tests 4.1 and 4.4.

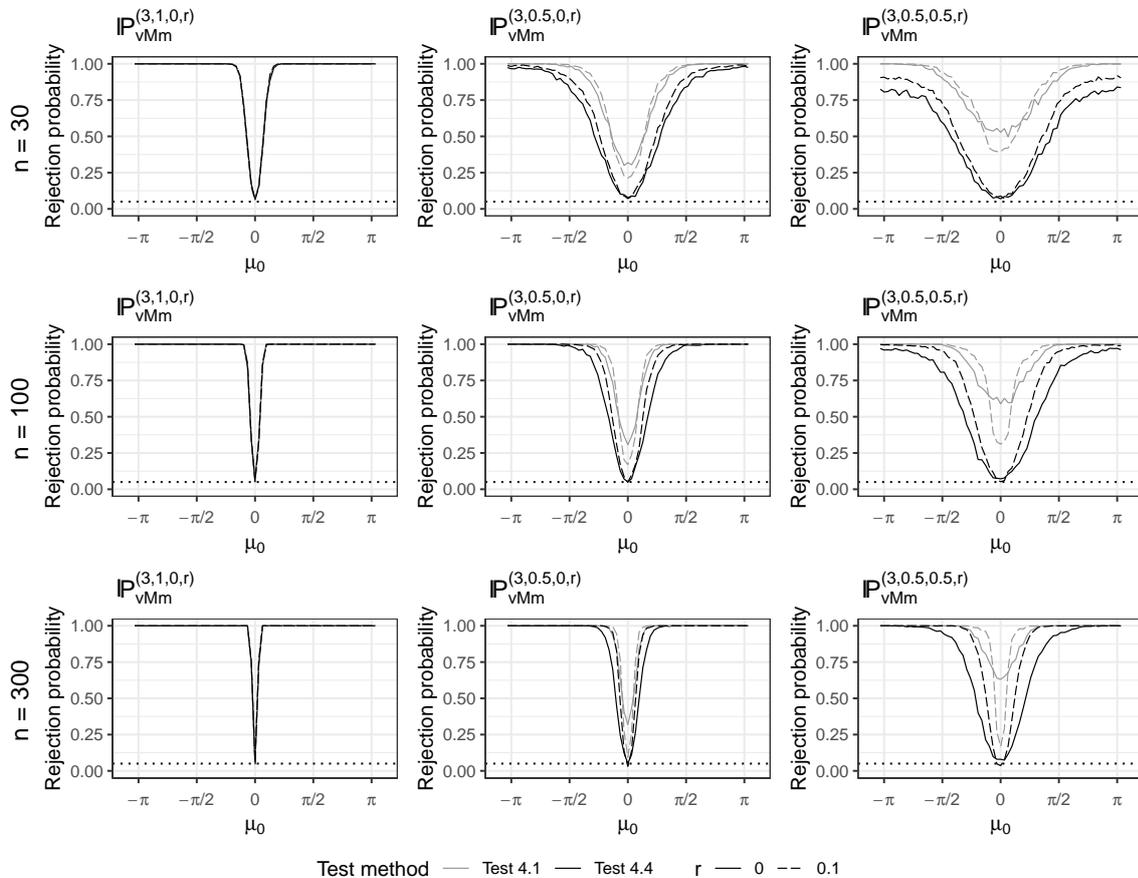


Figure 7: Rejection probabilities for H_0^1 for varying $\mu_0 \in [-\pi, \pi]$ under quantile based tests (Test 4.1, gray) and bootstrap based tests (Test 4.4, black) at nominal level 5% (dotted horizontal) with one sample of size $n = 30$ (top row), $n = 100$ (middle row), and $n = 300$ (bottom row). The solid lines represent samples which were generated from mixed von Mises distributions $\mathbb{P}_{vM}^{\kappa,\beta,\lambda,r}$, i.e., $r = 0$. The dashed lines correspond to samples from $\mathbb{P}_{vM}^{\kappa,\beta,\lambda,r}$ where all elements closer to $-\pi$ than $r = 0.1$ were mirrored. Table 3 gives an overview of parameters.

With increasing scale of FSS (see Table 3) we see that the quantile tests (gray) become more and more liberal while the bootstrap tests (black) maintain the correct level. In particular, the quantile based tests perform poorly in the presence of considerable Type I FSS, in consequence of the scaled variance of intrinsic sample means $n\mathbb{E}[d(\hat{\mu}_n, \mu)^2]$ being larger than the Euclidean variance σ^2 . Upon very close

inspection, in case of Type II FSS (dashed lines, $r > 0$) we see that the bootstrap tests may be slightly too conservative. Conversely, in case of Type I FSS (solid lines, $r = 0$) the bootstrap tests may be slightly to liberal. Both effects may be due to a systematic bias, cf. Remark 4.5.

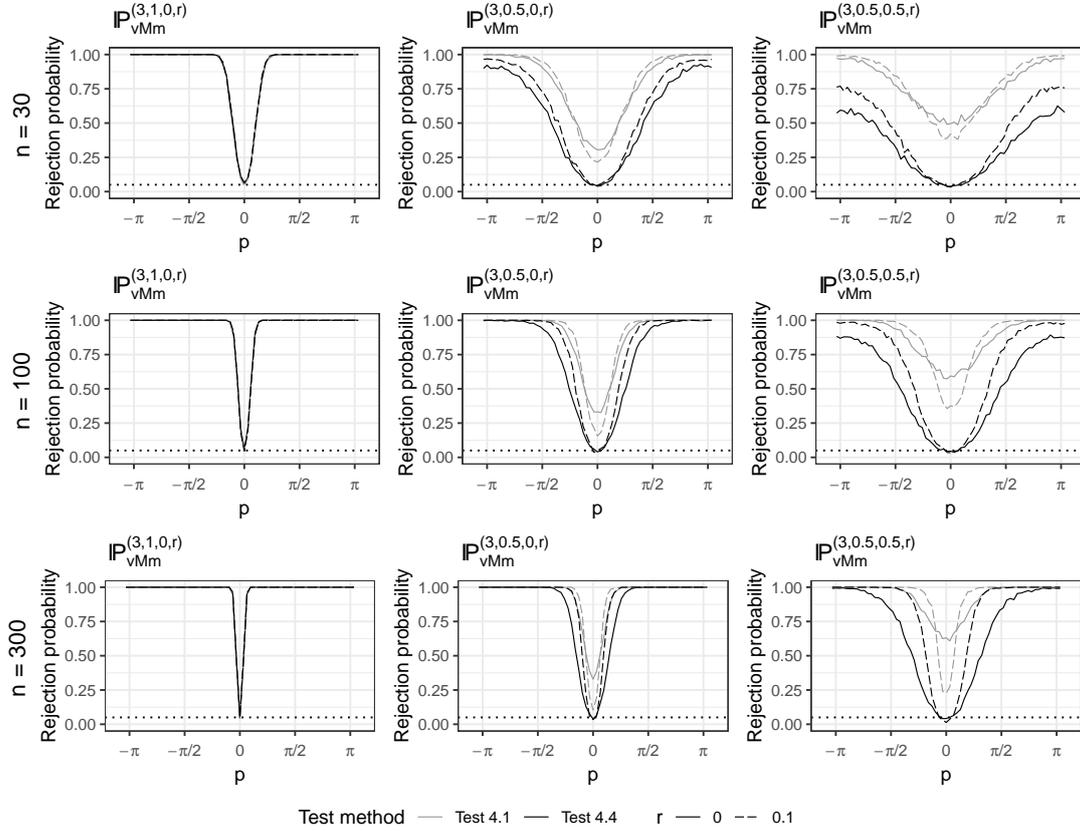


Figure 8: Rejection probabilities for H_0^2 over angle $p \in [-\pi, \pi]$ between $\mu^{(X)}$ and $\mu^{(Y)}$ under quantile based test (Test 4.1, gray) and bootstrap based test (Test 4.4, black) at nominal level 5% (dotted horizontal) with two samples from the same distribution of size $n = 30$ (top row), $n = 100$ (middle row), and $n = 300$ (bottom row) but where the latter sample is rotated by p . The solid lines represent samples which were generated from mixed von Mises distributions $\mathbb{P}_{vMm}^{\kappa, \beta, \lambda, r}$, i.e., $r = 0$. The dashed lines correspond to samples from $\mathbb{P}_{vMm}^{\kappa, \beta, \lambda, r}$ where all elements closer to $-\pi$ than $r = 0.1$ were mirrored.

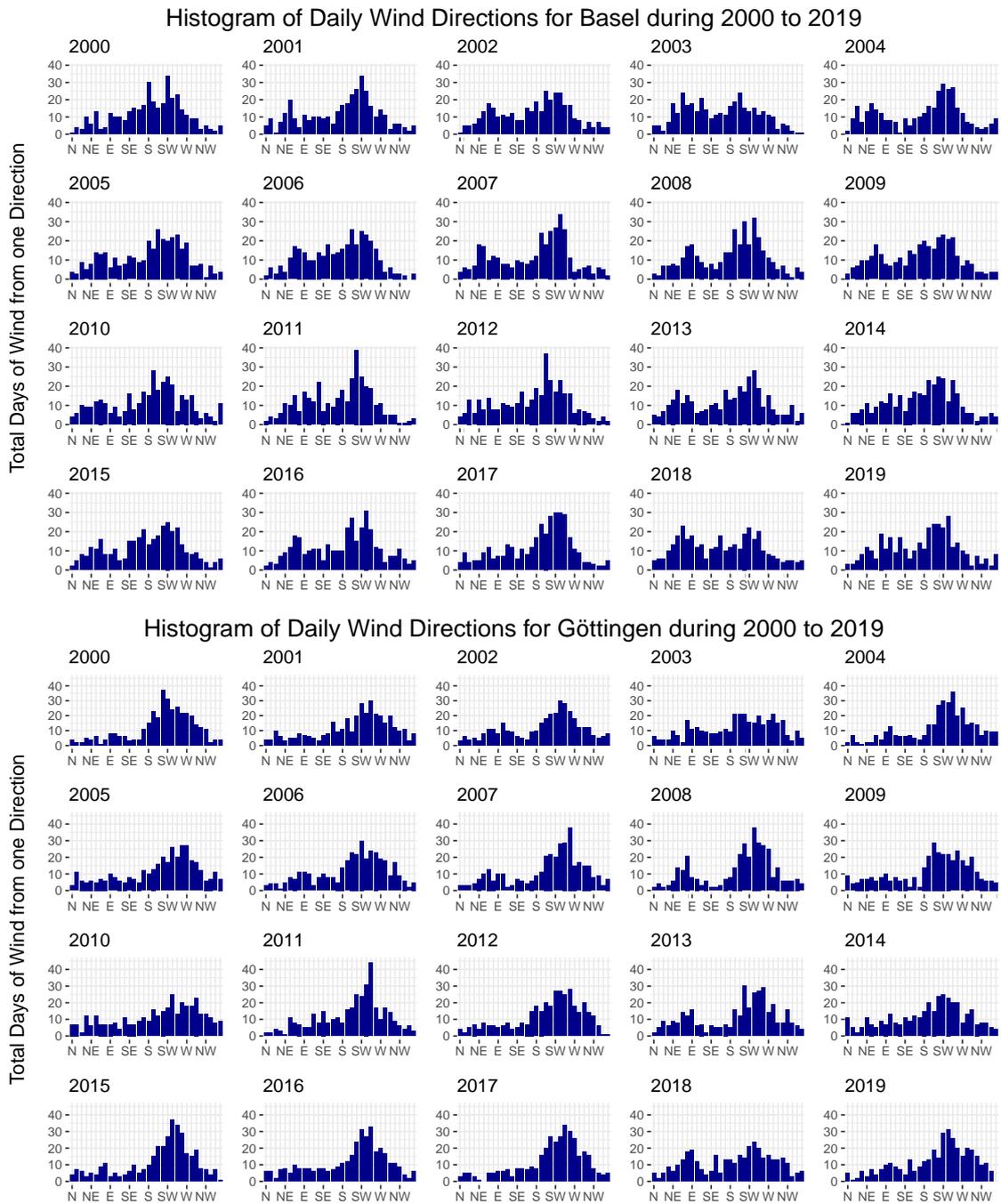


Figure 9: Histograms of Fréchet means of daily wind direction data by meteoblue AG (2020) for Basel (top) and Göttingen (bottom) for the years 2000 to 2019. The x-axis is divided into 32 segments with labels N : “North”, E : “East”, S : “South”, and W : “West”, indicating origin of average wind direction.

6 Assessing Significant Change of Wind Direction

In application of our methods dealing with FSS, we analyze wind data from Basel and Göttingen (the city of the authors' institution) provided by meteoblue AG (2020). For our purpose we consider daily Fréchet mean wind directions for the years 2000 to 2019 giving for each city 20 samples of two-dimensional circular data of size $n = 365$. The respective daily wind directions are illustrated in Figure 9. To assess a possible effect of climate change, we test for a significant change in wind direction.

For each of these 40 samples we estimated the *scale of FSS* from (8) with $B = 1000$, cf. Table 4. Remarkably, for both cities all of the scales of FSS statistically indicate presence of FSS ($p = B^{-1}$) except for Göttingen in the year 2017.

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Basel	1.601	4.067	1.542	1.197	32.695	2.683	1.845	2.096	1.676	2.344
Göttingen	3.317	2.734	2.760	3.065	1.492	4.171	1.605	15.493	3.861	8.796
Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Basel	5.010	1.775	1.466	3.211	1.491	1.741	1.705	5.229	2.694	2.481
Göttingen	8.003	1.685	4.457	36.365	3.446	3.305	5.636	1.037	1.804	2.592

Table 4: *Scale of FSS from (8) for daily Fréchet mean wind directions in Basel and Göttingen for the years 2000 to 2019. All of these values indicate presence of FSS (with the lowest p -value possible: $p = 10^{-3}$) except for Göttingen in the year 2017.*

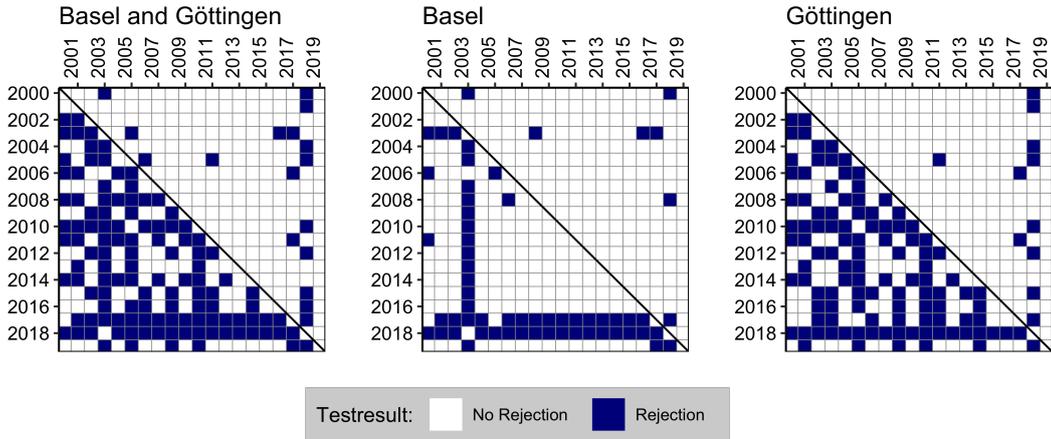


Figure 10: *Benjamini-Hochberg corrected test results (above diagonals: Test 4.4, below diagonals: Test 4.1) of comparisons of daily Fréchet means of wind direction data for the years 2000 to 2019 for Basel and Göttingen (left), Basel (middle), and Göttingen (right) for a significance level of 5%.*

In consequence, we expect that the quantile based two-sample Test 4.1 will feature a considerably high error of the first kind as compared to the bootstrap based Test 4.4. For each series of $20 \cdot 19/2 = 190$ tests at nominal level $\alpha = 0.05$ we performed a Benjamini-Hochberg correction, cf. Figure 10.

The first series of tests is performed on the two-torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ for Basel and Göttingen jointly (left panel of Figure 10). According to the quantile Test 4.1, a majority of years seem to be significantly different from a great number of others. Applying the bootstrap Test 4.4 we see this “noise” disappearing, leaving only the years 2003, 2005, 2017 and 2018 as moderately exceptional, where 2018 appears to be more exceptional than the other three.

Looking only at Basel (now testing only on \mathbb{S}^1 , middle panel of Figure 10), Test 4.1 seems to suggest very clearly that the years 2003, 2017 and 2018 are exceptional. Test 4.4, however, clarifies the picture: 2017 is not exceptional and 2003 and 2018 are only moderately exceptional.

Comparison with Göttingen only (again testing on \mathbb{S}^1 , right panel of Figure 10) shows the “noise” from Test 4.1 is rooted in Göttingen. Again, Test 4.4 clarifies the picture: only the years 2005, 2017 are mildly exceptional. The year 2018 remains more prominently exceptional. The year 2010, which appeared almost as strongly exceptional as year 2018 for Test 4.1, is no longer exceptional for Test 4.4. Notably, 2017 is also exceptional through absence of FSS, cf. Table 4.

In conclusion, using the bootstrap test which preserves the nominal level, the years 2003 and 2018 appear exceptional along with 2005 and 2017. These findings fall well into the climatological context of central European heat waves linked to exceptional wind constellations (Kornhuber et al. (2019) identify a *recurrent wave-7* wind pattern for the years 2003, 2006, 2015 and 2018). While the 2003 heat wave occurred in most of Europe, the 2018 heat wave manifested in a climatic dipole: hot and dry north of the Alps, comparably cool and moist across large parts of Mediterranean, cf. Buras et al. (2019). This is in agreement with our finding that the wind anomaly of 2018 was more prominent for Göttingen (in the northern part of Germany), rather than for Basel (edging the southern border of Germany) which is barely north of the Alps.

In a debate quantifying climate change, its anthropogenic component and future costs, linking to changes of wind patterns (e.g. McInnes et al. (2011)), our new inferential tools for cyclic data presented here, warrant a more detailed application in future work.

7 Discussion and Outlook

In this contribution, we have investigated two manifolds with codimension one cut loci, namely circles and tori and found FSS manifesting in two different types. We expect similar findings for other manifolds with codimension one cut loci, such as real projective spaces, modeling projective shapes, say, as in Mardia and Patrangenaru (2005); Hotz et al. (2016). We conjecture that this is, using the language of Eltzner (2019), a consequence of *topological smeariness*. On manifolds with higher codimension cut loci, in the language of Eltzner (2019), for instance on arbitrary spheres, there is the different phenomenon of *geometrical smeariness*. We conjecture that this leads to Type I FSS only.

We have demonstrated by simulation that on the circle and the torus, the proposed bootstrap tests preserve the level fairly well. Motivated by close inspection, we conjecture that a bias correction is necessary to obtain a corresponding theoretical result. This will be particularly challenging for sample sizes in the regimes of ascending or descending slopes of FSS, cf. Figures 1 and 4.

Acknowledgement

The authors gratefully acknowledge support by the DFG HU 1575/7, the DFG CRC 2088, the Niedersachsen Vorab of the Volkswagen Foundation and the Felix-Bernstein Institute for Mathematical Stochastics in the Biosciences. We also thank meteoblue AG for providing the wind data sets.

References

- Afsari, B. (2009). *Means and averaging on Riemannian manifolds*. University of Maryland.
- Bhattacharya, A. and R. Bhattacharya (2012). *Nonparametric Inference on Manifolds*. Cambridge University Press.
- Bhattacharya, R. and L. Lin (2017). Omnibus CLT for Fréchet means and nonparametric inference on non-Euclidean spaces.

- Bhattacharya, R. N. and V. Patrangenaru (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds II. *The Annals of Statistics* 33(3), 1225–1259.
- Buras, A., A. Rammig, and C. S. Zang (2019). Quantifying impacts of the drought 2018 on European ecosystems in comparison to 2003. *arXiv preprint arXiv:1906.08605*.
- Dubey, P. and H.-G. Müller (2019). Fréchet analysis of variance for random objects. *Biometrika* 106(4), 803–821.
- Eltzner, B. (2019). Measure Dependent Asymptotic Rate of the Mean: Geometrical Smeariness. *arXiv e-prints 1908.04233*.
- Eltzner, B., F. Galaz-Garcia, S. F. Huckemann, and W. Tuschmann (2019). Stability of the cut locus and a central limit theorem for Fréchet means of Riemannian manifolds. *arXiv e-prints 1909.00410*.
- Eltzner, B. and S. Huckemann (2017). Bootstrapping descriptors for non-Euclidean data. In *International Conference on Geometric Science of Information*, pp. 12–19. Springer.
- Eltzner, B. and S. F. Huckemann (2019). A smeary central limit theorem for manifolds with application to high dimensional spheres. *Annals of Statistics* 47(6), 3360–3381.
- Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'Institut de Henri Poincaré* 10(4), 215–310.
- Hotz, T. and S. Huckemann (2015). Intrinsic means on the circle: Uniqueness, locus and asymptotics. *Annals of the Institute of Statistical Mathematics* 67(1), 177–193.
- Hotz, T., F. Kelma, and J. T. Kent (2016). On the topology of projective shapes. *arXiv e-prints 1602.04330*.
- Huckemann, S. F. and B. Eltzner (2020). Statistical methods generalizing principal component analysis to non-Euclidean spaces. In *Handbook of Variational Methods for Nonlinear Geometric Data*, pp. 317–388. Springer.
- Kornhuber, K., S. Osprey, D. Coumou, S. Petri, V. Petoukhov, S. Rahmstorf, and L. Gray (2019). Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters* 14(5), 054002.
- Lund, U., C. Agostinelli, et al. (2017). Circular. <https://cran.r-project.org/web/packages/circular/index.html>. R package version 0.4-93.
- Mardia, K. and V. Patrangenaru (2005). Directions and projective shapes. *The Annals of Statistics* 33, 1666–1699.
- Mardia, K. V. and P. E. Jupp (2000). *Directional Statistics*. New York: Wiley.
- McInnes, K. L., T. A. Erwin, and J. M. Bathols (2011). Global climate model projected changes in 10 m wind speed and direction due to anthropogenic climate change. *Atmospheric Science Letters* 12(4), 325–333.
- McKilliam, R. G., B. G. Quinn, and I. V. L. Clarkson (2012). Direction estimation by minimum squared arc length. *IEEE Transactions on Signal Processing* 60(5), 2115–2124.
- meteoblue AG (2020). history+ platform. www.meteoblue.com/en/weather/archive/export/basel_switzerland_2661604, www.meteoblue.com/en/weather/archive/export/goettingen_germany_2918632. Last checked on 06/04/2020.
- Pennec, X. (2019). Curvature effects on the empirical mean in Riemannian and affine manifolds: a non-asymptotic high concentration expansion in the small-sample regime. *arXiv e-prints 1906.07418*.

Tran, D. (2019). Behavior of Fréchet mean and central limit theorems on spheres. *arXiv e-prints 1911.01985*.

Ziezold, H. (1977). Expected figures and a strong law of large numbers for random elements in quasi-metric spaces. *Transaction of the 7th Prague Conference on Information Theory, Statistical Decision Function and Random Processes A*, 591–602.

Appendix: Proofs

Proof of Corollary 2.4. Since the density of \mathbb{P}_r on $(-\pi, -\pi + 1/2) \cup (\pi - 1/2, \pi)$ is strictly smaller than $\frac{1}{2\pi}$ we see that $c_r > \frac{1}{2\pi}$. Hence, by Hotz and Huckemann (2015, Propostion 1) there is a unique Fréchet population mean μ in $[-1/2, 1/2]$, which, by symmetry, is $\mu = 0$. This gives (i) and the strong law by Ziezold (1977) yields (ii), (a).

Next, plug in $x = \hat{\mu}_n$ in (5) as done in the proof of Theorem 2.1, to obtain

$$\sqrt{n} \cdot \text{sign}(\hat{\mu}_n) \exp(-1/|\hat{\mu}_n|^r) = \frac{\sqrt{n}}{2\pi} \bar{X} + o_P(1).$$

Subjecting both sides above to $y \mapsto \text{sign } y \cdot \log(|y|) \mathbf{1}_{\mathbb{R} \setminus \{0\}}(y)$ gives

$$\begin{aligned} \text{sign}(\hat{\mu}_n) \left(\log \sqrt{n} - \frac{1}{|\hat{\mu}_n|^r} \right) &= \text{sign}(\bar{X}) \log \left| \frac{\sqrt{n}}{2\pi} \bar{X} + o_P(1) \right| \\ &= \text{sign}(\bar{X}) \log \left| \frac{\sqrt{n}}{2\pi} \bar{X} \right| + o_P(1). \end{aligned}$$

For the last equality, we have used $\frac{2\pi}{\sqrt{n}\bar{X}} = O_P(1)$ and $\log(|1 + o_P(1)|) = o_P(1)$. Setting $L_n = \text{sign}(\bar{X}) \log \left| \frac{\sqrt{n}}{2\pi} \bar{X} \right|$, proceed to

$$\text{sign}(\hat{\mu}_n) \log \sqrt{n} - L_n = \frac{\text{sign}(\hat{\mu}_n) + o_P(|\hat{\mu}|^r)}{|\hat{\mu}|^r},$$

yielding

$$\text{sign}(\hat{\mu}_n) |\hat{\mu}_n|^r = \frac{1 + o_P(|\hat{\mu}|^r)}{\text{sign}(\hat{\mu}_n) \log \sqrt{n} - L_n} = \frac{1 + o_P(|\hat{\mu}|^r)}{\text{sign}(\hat{\mu}_n) \log \sqrt{n}} \sum_{k=0}^{\infty} \left(\frac{\text{sign}(\hat{\mu}_n) L_n}{\log \sqrt{n}} \right)^k,$$

whenever $|L_n| < \log \sqrt{n}$, and the probability for this tends to 1. Since $\hat{\mu}_n = o_P(1)$ we conclude

$$|\hat{\mu}_n|^r = \frac{1}{\log \sqrt{n}} + \frac{\text{sign}(\hat{\mu}_n) L_n}{(\log \sqrt{n})^2} + o_P \left(\frac{1}{(\log \sqrt{n})^2} \right).$$

By symmetry, this yields Assertion (b). It further yields

$$\begin{aligned} \hat{\mu}_n &= \text{sign}(\hat{\mu}_n) \left(\frac{1}{\log \sqrt{n}} + \frac{\text{sign}(\hat{\mu}_n) L_n}{(\log \sqrt{n})^2} + o_P \left(\frac{1}{(\log \sqrt{n})^2} \right) \right)^{\frac{1}{r}} \\ &= \frac{\text{sign}(\hat{\mu}_n)}{(\log \sqrt{n})^{\frac{1}{r}}} \left(1 + \frac{1}{r} \frac{\text{sign}(\hat{\mu}_n) L_n}{\log \sqrt{n}} + o_P \left(\frac{1}{\log \sqrt{n}} \right) \right). \end{aligned}$$

With the definition of L_n above, we thus obtain Assertion (c). □