Bernoulli **22**(4), 2016, 2521–2547
DOI: 10.3150/15-BEJ736

Methods for improving estimators of truncated circular parameters

KANIKA* and SOMESH KUMAR**

*Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur-721302, West Bengal, India. E-mail: *10ma90r04@iitkgp.ac.in; **smsh@maths.iitkgp.ernet.in*

In decision theoretic estimation of parameters in Euclidean space $\mathbb{R}^p$, the action space is chosen to be the convex closure of the estimand space. In this paper, the concept has been extended to the estimation of circular parameters of distributions having support as a circle, torus or cylinder. As directional distributions are of curved nature, existing methods for distributions with parameters taking values in $\mathbb{R}^p$ are not immediately applicable here. A circle is the simplest one-dimensional Riemannian manifold. We employ concepts of convexity, projection, etc., on manifolds to develop sufficient conditions for inadmissibility of estimators for circular parameters. Further invariance under a compact group of transformations is introduced in the estimation problem and a complete class theorem for equivariant estimators is derived. This extends the results of Moors [*J. Amer. Statist. Assoc.* **76** (1981) 910–915] on $\mathbb{R}^p$ to circles. The findings are of special interest to the case when a circular parameter is truncated. The results are implemented to a wide range of directional distributions to obtain improved estimators of circular parameters.

Keywords: admissibility; convexity; directional data; invariance; projection; truncated estimation problem

1. Introduction

Problems of estimation when the parameter space is restricted are encountered often in practice. These restrictions arise due to prior information on parameters and they can be in the form of bounds on the range or equality/inequality constraints of several parameters. For recent developments and discussions on various aspects of estimation procedures in restricted parameter space problems, one may refer to [14,15,17,18,28] and references therein. Frequently in practical applications, we assume the random observations taking values in Euclidean spaces. However, it sometimes may be more useful to represent them on circles/spheres/cylinders. In such cases, we employ directional distributions. For instance, mortality data due to a specific disease may be better represented as circular data to study the seasonal pattern of the disease. There are numerous situations in biological, meteorological, astronomical applications, where directional data (circular/axial/spherical) arises [1,8,19]. However, little attention has been paid to problems of estimating directional parameters under constraints. Rueda, Fernández and Peddada [25] considered the estimation of the circular parameters under order restrictions. There are situations when the parameter may lie on an arc of the circle. For example, the peak of mortality rates due to respiratory diseases occurs during November to February.

One major consequence of placing restrictions on the natural parameter space is that estimators derived using standard concepts of maximum likelihood, minimaxity, invariance, etc., become inadmissible. However, existing methods developed for Euclidean spaces $\mathbb{R}^p$ are not di-

1350-7265

rectly applicable to directions which are represented to lie on unit hypersphere with the center at origin $\mathbb{S}_{p-1} = \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2} = 1\}$. Topological properties of $\mathbb{S}_{p-1}$ depend on the differential geometry of embedding $\mathbb{S}_{p-1}$ in $\mathbb{R}^p$ as $\mathbf{x} \to \mathbf{x}/\|\mathbf{x}\|$ with $\mathbf{x} \in \mathbb{R}^p$. We need to suitably modify techniques available for $\mathbb{R}^p$ to improve standard estimators of directions. In this paper, we consider the case of $p = 2$ and denote by $\mathbb{S}$ a unit circle. The elements of $\mathbb{S}$ can be specified by corresponding angles with respect to an arbitrary choice of zero direction and orientation. Let us define by $\mathbb{T} = [0, 2\pi)$ the space of amplitudes of the unit vectors in $\mathbb{S}$. The point on $\mathbb{S}$ corresponding to an angle $\alpha \in \mathbb{T}$ is $(\cos\alpha, \sin\alpha)^T$ and the angle corresponding to a point $\mathbf{x} = (x_1, x_2)^T \in \mathbb{S}$ is $\mathsf{atan}(x_2/x_1)$, where the function $\mathsf{atan}(\cdot)$ is given in (2.2). Note that $\mathbb{S}$ and $\mathbb{T}$ are isomorphic.

Let random variable $\mathbf{Z} \in \mathfrak{Z}$ have an unknown probability distribution $\mathsf{P}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} \in \Omega$. Let the family $\{\mathsf{P}_{\boldsymbol{\nu}}, \boldsymbol{\nu} \in \Omega\}$ be dominated by a measure η. The support $\mathfrak{Z}$ may be circle $\mathbb{T}$, torus $\mathbb{T}^k$ or cylinder $\mathbb{R} \times \mathbb{T}$, etc.; however, the estimand $h(\boldsymbol{\nu})$ is circular, where $h(\cdot)$ is a measurable function from Ω into $\mathbb{T}$. The problem of estimating $h(\boldsymbol{\nu})$ is considered under a circular loss function:

$$\mathsf{L}(\boldsymbol{\nu}, \delta) = 1 - \cos\big(h(\boldsymbol{\nu}) - \delta\big). \tag{1.1}$$

In the case when the parameter space is a subspace of $\mathbb{R}^p$ and the loss function is an increasing function of Euclidean distance, the action space is chosen as a convex closure of the range of estimand. Unlike this well-known result, it is demonstrated in Section 3 that analogous result does not necessarily hold for circular parameter $h(\boldsymbol{\nu})$ under the loss L.

One of the major contributions was of Moors [22] (see [21], Chapter 3, also) to the estimation problem of truncated parameters of a unknown family of distributions on $\mathbb{R}^p$ dominated by a σ-finite measure. Estimators (except the constant ones) taking value near the boundaries of action space with the positive probability turn out to be inadmissible with respect to the squared loss function under certain conditions on the transformation group. He considered invariance under a finite group with measure preserving elements such that induced transformations of the action space satisfy the linearity property and group of these transformations is commutative. Under this scenario, he constructed a subspace of the original action space and proved that any invariant estimator taking values outside this new action space with the positive probability is inadmissible and dominated by its projection on the new action space. Later, Moors and van Houwelingen [23] relaxed conditions of measure preserving and commutativity. Without dropping these conditions, Kumar and Sharma [16] generalized the result of Moors [22] to a locally compact group such that induced transformations on the action space are affine (stronger condition than linear property) and loss function is an increasing function of Euclidean distance. Along with these ideas, an analogous theory for circular parameter $h(\boldsymbol{\nu})$ is developed in Section 4.

The outline of the paper is as follows. The concepts of distance formulae, convexity, closure of a set and projections play a prominent role here. Section 2 provides the mathematical background of these concepts for $\mathbb{T}$. In Section 3, we consider the estimation of circular parameter $h(\boldsymbol{\nu})$ when it is restricted to lie on an arc of circle and estimation space $\mathcal{A}$ is chosen as the convex closure of $h(\Omega)$. A complete class result for this estimation problem is obtained under certain conditions. Then the result is illustrated for several directional distributions. In Section 4, we introduce invariance under a compact group $\mathcal{G}$ in the estimation problem such that induced transformations on $\mathcal{A}$ satisfy the circular property. Sufficient condition for inadmissibility of an $\mathcal{G}$-equivariant estimator is obtained. Applications of this result are demonstrated for both unrestricted or re-

stricted estimation problems. For restricted estimation problems, improved estimators obtained in Section 3 are further improved using the result of Section 4.

2. Definitions and preliminary results

Before we embark on estimation problem, we introduce some preliminary results in this section. For a subset $A \subset \mathbb{T}$, Lebesgue measure, interior, convex hull, convex closure and boundary of A are denoted by $l(A)$, $\mathsf{int}(A)$, $\mathsf{conv}(A)$, $\mathsf{cc}(A)$ and $\mathsf{bd}(A)$, respectively.

2.1. Convexity

It is more convenient to deal with polar coordinates than Cartesian coordinates when observations lie on a unit circle. We summarize the concept of convexity for the circle $\mathbb{S}$ and then adopt it for the space $\mathbb{T}$.

A geodesic ([27], page 15), on Riemannian manifold generalizes the line in Euclidean space. In the context of Riemannian manifold $\mathbb{M}$ equipped with Riemannian metric, a subset $A \subset \mathbb{M}$ is convex if minimal geodesic with the end points in A belongs to A. In the case of $\mathbb{S}$, great circles are geodesics. Minimal geodesic between any two points on $\mathbb{S}$ is unique unless points are antipodal (diametrically opposite). Some concepts of convexity on $\mathbb{S}$ were introduced in [4], Section 9.1. Here, we use convexity and strong convexity as given below.

Definition 2.1 (Convex). *A set $A \subset \mathbb{S}$ is convex if for any two points in A, there exists a minor arc of great circle lying entirely in A joining them.*

By convention, this definition allows antipodal points to lie in convex sets. Every segment of a semicircle is a convex subset of $\mathbb{S}$.

Definition 2.2 (Strongly convex). *A subset A of $\mathbb{S}$ is strongly convex if A is convex and does not contain antipodal points.*

In the case of $\mathbb{R}^p$, convex hull of any subset is the collection of all possible weighted arithmetic mean of points in that subset. Analogously for a set $A \subset \mathbb{S}$, the convex hull of A is the smallest convex set (not necessarily strong convex) containing A, that is, it consists of

$$(w_1\mathbf{x}_1 + \cdots + w_n\mathbf{x}_n)/\|w_1\mathbf{x}_1 + \cdots + w_n\mathbf{x}_n\| \tag{2.1}$$

for all nonnegative weights $w_1, \ldots, w_n$ such that $\sum_{i=1}^n w_i = 1$ and for all $\mathbf{x}_1, \ldots, \mathbf{x}_n \in A$ provided the norm in the denominator is nonzero. The convex hull of two antipodal points of $\mathbb{S}$ does not exist ([3], Section 2.3). Polar form of (2.1) is discussed in Section 2.2.

Extension of a fundamental theorem of Carathéodory for $\mathbb{R}^2$ to $\mathbb{S}$ can be stated as below (see [3]).

Lemma 2.1. *Each point in the convex hull of a set $A \subset \mathbb{S}$ can be expressed as normalized weighted arithmetic mean of at most 2 points of A.*

To extend definitions of convex and strong convex sets to the space $\mathbb{T}$, we need certain subsets of $\mathbb{T}$. For $\alpha, \beta \in \mathbb{T}$, definitions for sets of type I, J, K and K_1 are stated only for intervals of form (α, β). They may be extended to intervals of other forms $(\alpha, \beta]$, $[\alpha, \beta)$ and $[\alpha, \beta]$. For $\alpha \leq \beta$, let

$$I(\alpha, \beta) = (\alpha, \beta),$$

$$J(\alpha, \beta) = [0, \alpha) \cup (\beta, 2\pi),$$

$$K(\alpha, \beta) = \begin{cases} I(\alpha, \beta), & \text{if } 0 \leq (\beta - \alpha) < \pi; \\ I(\alpha, \beta) \text{ or } J(\alpha, \beta), & \text{if } (\beta - \alpha) = \pi; \\ J(\alpha, \beta), & \text{if } \pi < (\beta - \alpha) < 2\pi. \end{cases}$$

Sets $I[\alpha, \beta]$ and $J[\alpha, \beta]$ contain all the angles corresponding to an arc joining two points $(\cos\alpha, \sin\alpha)^T$ and $(\cos\beta, \sin\beta)^T$ in the positive and the negative directions, respectively. Moreover, $K[\alpha, \beta]$ contains angles corresponding to the minor arc joining them.

Remark 2.1. Although in definition of I-type set, β is not allowed to take value 2π, intervals $(\alpha, 2\pi)$ and $[\alpha, 2\pi)$ can be expressed as $J(0, \alpha)$ and $J(0, \alpha]$, respectively, for all $\alpha \in \mathbb{T}$. For any $\alpha \in \mathbb{T}$, $J(\alpha, \alpha) = \mathbb{T} - \{\alpha\}$ and $J(\alpha, \alpha] = J[\alpha, \alpha) = J[\alpha, \alpha] = \mathbb{T}$.

For defining sets of type I, J and K, we have taken $\alpha \leq \beta$. The following definition for K_1-type subsets of $\mathbb{T}$ does not have this restriction.

$$K_1(\alpha, \beta) = \begin{cases} K(\alpha, \beta), & \text{if } \alpha \leq \beta, \\ K(\beta, \alpha), & \text{if } \alpha > \beta. \end{cases}$$

Note that set $K(\alpha, \beta)$ is isomorphic to $I(0, \gamma)$ with $\gamma \in [0, \pi]$. For $0 \leq (\beta - \alpha) \leq \pi$ and $\pi < (\beta - \alpha) < 2\pi$, $K(\alpha, \beta)$ can be transformed to $I(0, \beta - \alpha)$ and $I(0, 2\pi - \beta + \alpha)$ using rotation by angles $2\pi - \alpha$ and $2\pi - \beta$, respectively. Extending this argument, we have the following result.

Lemma 2.2. *Sets $K_1(\alpha, \beta)$, $K_1(\alpha, \beta]$, $K_1[\alpha, \beta)$ and $K_1[\alpha, \beta]$ with $\alpha, \beta \in \mathbb{T}$ are isomorphic to $I(0, \gamma)$, $I(0, \gamma]$, $I[0, \gamma)$ and $I[0, \gamma]$, respectively, with $0 \leq \gamma \leq \pi$.*

For studying topological properties, we consider the metric space $(\mathbb{T}, d)$ with the following definition of metric:

$$d(\alpha, \beta) = 1 - \cos(\alpha - \beta), \qquad \alpha, \beta \in \mathbb{T}.$$

Here, $\sqrt{2d}$ simply returns lengths of chord between points $(\cos\alpha, \sin\alpha)^T$ and $(\cos\beta, \sin\beta)^T$, respectively. Consider the following classes of subsets of $\mathbb{T}$:

$$\mathfrak{C}_1 = \{\mathbb{T}\} \cup \big\{K_1(\alpha, \beta), K_1(\alpha, \beta], K_1[\alpha, \beta), K_1[\alpha, \beta] : \alpha, \beta, \in \mathbb{T}\big\},$$

$$\mathfrak{C}_2 = \big\{K_1(\alpha, \beta), K_1(\alpha, \beta], K_1[\alpha, \beta) : \alpha, \beta \in \mathbb{T}\big\} \cup \big\{K_1[\alpha_1, \beta_1] : \alpha_1, \beta_1 \in \mathbb{T} \text{ and } |\alpha_1 - \beta_1| \neq \pi\big\},$$

$$\mathfrak{C}_3 = \mathfrak{C}_1 - \mathfrak{C}_2 = \{\mathbb{T}\} \cup \big\{K_1[\alpha, \beta] : \alpha, \beta, \in \mathbb{T} \text{ and } |\alpha - \beta| = \pi\big\},$$

$$\mathfrak{C}_4 = \{\varnothing, \mathbb{T}\} \cup \big\{I[\alpha, \beta], J[\alpha, \beta] : \alpha, \beta, \in \mathbb{T} \text{ and } \alpha \leq \beta\big\},$$

$$\mathfrak{C}_5 = \mathfrak{C}_1 \cap \mathfrak{C}_4 = \{\varnothing, \mathbb{T}\} \cup \big\{K_1[\alpha, \beta] : \alpha, \beta \in \mathbb{T} \text{ and } \alpha \leq \beta\big\}.$$

Remark 2.2. Classes $\mathfrak{C}_1$, $\mathfrak{C}_2$ and $\mathfrak{C}_5$ consist of all convex, strongly convex and closed convex subsets of $\mathbb{T}$, respectively. Elements of $\mathfrak{C}_3$ and $\mathfrak{C}_4$ are closed subsets of $\mathbb{T}$. Moreover, sets belonging to $\mathfrak{C}_3$ (except $\mathbb{T}$) and $\mathfrak{C}_4$ are corresponding to any minor arc and any arc on the unit circle, respectively.

2.2. Circular mean direction

Since the arithmetic mean is not a suitable measure of central tendency for the angular data, the circular mean direction is used ([12], page 13). The weighted circular mean direction of the observations $\phi_1, \ldots, \phi_n$ belonging to $\mathbb{T}$ with weights $w_1, \ldots, w_n$, $w_i \geq 0$ $(i = 1, \ldots, n)$ such that $\sum_{i=1}^{n} w_i = 1$ is defined as

$$\bar{\phi}_w = \mathsf{atan}\left(\frac{\sum_{i=1}^{n} w_i \sin\phi_i}{\sum_{i=1}^{n} w_i \cos\phi_i}\right),$$

with the following definition of $\mathsf{atan}(\cdot)$

$$\mathsf{atan}\left(\frac{s}{c}\right) = \begin{cases} \tan^{-1}(s/c), & \text{if } c > 0,\ s \geq 0; \\ \pi + \tan^{-1}(s/c), & \text{if } c < 0; \\ \pi/2, & \text{if } c = 0,\ s > 0; \\ 2\pi + \tan^{-1}(s/c), & \text{if } c \geq 0,\ s < 0; \\ \text{not defined}, & \text{if } s = 0,\ c = 0, \end{cases} \tag{2.2}$$

where $\tan^{-1}(\cdot)$ is the standard inverse tangent function taking values in $(-\pi/2, \pi/2)$. The definition of $\mathsf{atan}(\cdot)$ function ensures the following property.

Lemma 2.3. *For all $\phi \in \mathbb{T}$ and $a \in \mathbb{R}$, we have*

$$\mathsf{atan}\left(\frac{\tan\phi + a}{1 - a\tan\phi}\right) = \left\{\phi + \tan^{-1}(a)\right\} \bmod (2\pi).$$

Note that $\bar{\phi}_w$ is polar form of (2.1) if ϕ_i is corresponding angle to $\mathbf{x}_i \in \mathbb{S}$ for all $i = 1, \ldots, n$. The following proposition proves that convex combination (weighted circular mean direction) of finite collection of the points in convex subset of $\mathbb{T}$ is again in that subset.

Proposition 2.1. *Let $A \in \mathfrak{C}_1$ (convex), $\phi_1, \ldots, \phi_n \in A$ and $w_1, \ldots, w_n$ be nonnegative weights with $\sum_{i=1}^{n} w_i = 1$. Then weighted circular mean direction $\bar{\phi}_w$ of these observations belongs to A, if it is defined.*

Circular mean direction of a circular random variable θ is defined as

$$\mathsf{CE}(\theta) = \mathsf{atan}(\mathsf{E}\sin\theta / \mathsf{E}\cos\theta).$$

An elementary result given in [7], page 74, states that if random variable $\mathbf{X}$ lies in convex subset of $\mathbb{R}^p$ with probability one, $\mathsf{E}(\mathbf{X})$ lies in the same subset. An analogous result for $\mathbb{T}$ is given below.

Proposition 2.2. *If $A \in \mathfrak{C}_1$ (convex) and θ is a random angle such that $\Pr(\theta \in A) = 1$, the mean direction $\mathsf{CE}(\theta) \in A$ if $\mathsf{CE}(\theta)$ exists. Furthermore, if $A \in \mathfrak{C}_2$ (strongly convex), $\mathsf{CE}(\theta)$ always exists. If $A \in \mathfrak{C}_3$ (convex but not strongly), $\mathsf{CE}(\theta)$ does not necessarily exist.*

2.3. Projection

The concept of projection in $\mathbb{R}^p$ is adopted to define projections of angles in $\mathbb{T}$.

Definition 2.3. *The projection of an angle $\phi \in \mathbb{T}$ on a nonempty set $A \in \mathfrak{C}_4$ (closed) is defined to be the unique point $\phi_0 \in A$ such that*

$$d(\phi, \phi_0) = \inf_{\psi \in A} d(\phi, \psi).$$

The case when $A = \mathbb{T}$ is trivial. For $\alpha, \beta \in \mathbb{T}$, let $\gamma = (\alpha + \beta)/2$. If A is of from $I[\alpha, \beta]$, ϕ_0 is given by

$$\phi_0 = \begin{cases} \phi, & \text{if } \phi \in I[\alpha, \beta]; \\ \alpha, & \text{if } \phi \in K_1(\alpha, \pi + \gamma]; \\ \beta, & \text{if } \phi \in K_1(\beta, \pi + \gamma); \end{cases} \quad \text{or} \quad \begin{cases} \phi, & \text{if } \phi \in I[\alpha, \beta]; \\ \alpha, & \text{if } \phi \in K_1(\alpha, \pi + \gamma); \\ \beta, & \text{if } \phi \in K_1(\beta, \pi + \gamma]. \end{cases}$$

Note that the two definitions are equivalent except when $\phi = \pi + \gamma$. For $\phi = \pi + \gamma$, first and second ones yield $\phi_0 = \alpha$ and $\phi_0 = \beta$, respectively. This is so because $d(\alpha, \pi + \gamma) = d(\beta, \pi + \gamma)$. If A is the form of $J[\alpha, \beta]$, ϕ_0 is given by

$$\phi_0 = \begin{cases} \phi, & \text{if } \phi \in J[\alpha, \beta]; \\ \alpha, & \text{if } \phi \in I(\alpha, \gamma]; \\ \beta, & \text{if } \phi \in I(\gamma, \beta); \end{cases} \quad \text{or} \quad \begin{cases} \phi, & \text{if } \phi \in J[\alpha, \beta]; \\ \alpha, & \text{if } \phi \in I(\alpha, \gamma); \\ \beta, & \text{if } \phi \in I[\gamma, \beta). \end{cases}$$

Once again the two definitions are equivalent except when $\phi = \gamma$.

Let A be a closed convex subset of $\mathbb{R}^p$ and $\mathbf{x} \in \mathbb{R}^p$. The projection $\mathbf{x}_0$ of $\mathbf{x} \notin A$ on A satisfies

$$\|\mathbf{x}_0 - \mathbf{y}\| < \|\mathbf{x} - \mathbf{y}\| \qquad \text{for all } \mathbf{y} \in A.$$

An analogous statement holds only for specific closed convex subsets of $\mathbb{T}$. The following result can be easily proved using geometrical arguments.

Lemma 2.4. *Let ϕ_0 be the projection of an angle $\phi \notin A$ on a set $A \in \mathfrak{C}_5$ (closed convex). The inequality*

$$d(\phi_0, \psi) < d(\phi, \psi) \qquad \textit{for all } \psi \in A$$

holds iff $A = I[\alpha, \beta]$ with $\beta < \alpha + (2/3)\pi$ or $A = J[\alpha, \beta]$ with $\beta > \alpha + (4/3)\pi$, that is, $l(A) < (2/3)\pi$. Moreover, if $l(A) = (2/3)\pi$ and $\mathsf{bd}(A) = \{b_1, b_2\}$, the above inequality remains strict for $\psi \in \mathsf{int}(A)$ and at least one of $\psi = b_i$ $(i = 1, 2)$.

For remaining sets in $\mathfrak{C}_5$, the above result holds for the expected values under certain conditions on the distribution of random variable θ (Lemma 2.5).

If the distribution of a circular random variable θ is symmetric about ψ, the density of θ with respect to any measure (measure is always finite as $\mathbb{T}$ is a compact space) would be a function of $\cos(\theta - \psi)$. Further, if this distribution is unimodal, mean direction and mode coincide. Let us denote by $f(\theta|\psi) = f(\cos(\theta - \psi))$ the density of a symmetric unimodal distribution with mode ψ. Now consider the mixture of two unimodals $f(\cdot|\psi)$ and $f(\cdot|\psi + \pi)$. This mixture would necessarily not be bimodal. Let θ have a mixture distribution with probability density $\varepsilon f(\theta|\psi) + (1 - \varepsilon) f(\theta|\psi + \pi)$, $\varepsilon \in [0, 1]$. For this mixture distribution, define

$$\zeta(t) = f'(t)/f'(-t), \qquad t \in [-1, 1] \tag{2.3}$$

with $t = \cos(\theta - \psi)$. Maxima and minima of $\zeta(t)$ are denoted by $\zeta_{\max}$ and $\zeta_{\min}$, respectively. Distribution of θ would be unimodal with modes ψ and $\psi + \pi$ for $\varepsilon \in [\{1 + \zeta_{\min}\}^{-1}, 1]$ and $\varepsilon \in [0, \{1 + \zeta_{\max}\}^{-1}]$, respectively. For remaining values of ε, it would be bimodal.

Lemma 2.5. *Suppose that θ is a continuous circular random variable whose distribution is symmetric about one of its mode ψ, where ψ belongs to $A \in \mathfrak{C}_5$ (closed convex) with $l(A) \in (2\pi/3, \pi]$ such that $\Pr(\theta \notin A) > 0$. Let θ_0 be the projection of θ on A. Then*

$$\mathsf{E}_{\psi}^{\theta}\left\{d(\theta_0, \psi)\right\} < \mathsf{E}_{\psi}^{\theta}\left\{d(\theta, \psi)\right\}$$

if distribution of θ satisfies one of the following conditions:

(C1) *distribution is unimodal with mode ψ*;
(C2) *distribution is mixture with probability density $\varepsilon f(\cdot|\psi) + (1 - \varepsilon) f(\cdot|\psi + \pi)$, where $\varepsilon \geq 1/2$ and $\zeta(\cdot)$ defined in (2.3) is an increasing function.*

It may be noted that the condition (C2) implies (C1) for $\varepsilon \in [\{1 + \zeta(-1)\}^{-1}, 1]$.

Remark 2.3. Convexity of density function $f(t)$ in $t \in [-1, 1]$ yields increasing nature of the function $\zeta(t)$.

3. Improving estimators in restricted parameter spaces

In Euclidean spaces, the action space is chosen as a convex closure of the estimand space since estimators outside this space with the positive probability are dominated by their projections on it. An analogous result stated below for estimating the circular parameter $h(\boldsymbol{\nu})$ is an immediate consequence of Lemmas 2.4 and 2.5.

Theorem 3.1. *Let estimand be $h(\boldsymbol{\nu}) \in \Omega_1 = h(\Omega) \subset \mathbb{T}$ and the loss function be L defined in (1.1). Denote the estimation space by $\mathcal{A} = \mathsf{cc}(\Omega_1)$. Any estimator $\delta(\mathbf{Z})$ satisfying $\Pr_{\boldsymbol{\nu}}(\delta(\mathbf{Z}) \notin \mathcal{A}) > 0$ for some $\boldsymbol{\nu} \in \Omega$ is inadmissible and dominated by the projection of $\delta(\mathbf{Z})$ on $\mathcal{A}$ if either of the following conditions holds*:

(C3) $l(\mathcal{A}) \leq (2/3)\pi$;
(C4) *distribution of $\delta(\mathbf{Z})$ is symmetric about $h(\boldsymbol{\nu})$ and with respect to Lebesgue measure, it satisfies one of the conditions* (C1) *and* (C2) *with $\psi = h(\boldsymbol{\nu})$.*

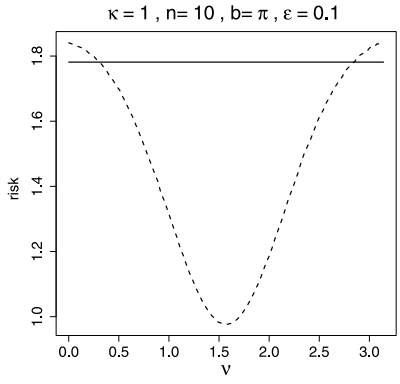

Figure 1. Risk plot (a) $\bar{\theta}$ (straight line) and (b) projection of $\bar{\theta}$ on $[0, \pi]$ (dotted line) under the loss L.

For the sake of clarity, the estimation space $\mathcal{A}$ can be called the action space only when its Lebesgue measure is less than or equal to $2\pi/3$.

Remark 3.1. Note that when estimand $h(\boldsymbol{\nu})$ is forced to lie on an arc of semicircle, $\mathcal{A}$ is strictly a subset of $\mathbb{T}$. If $h(\boldsymbol{\nu})$ does not take value on a semicircle, $\mathcal{A} = \mathbb{T}$ ([3], Theorem 7).

Conditions given in Theorem 3.1 for the inadmissibility of an estimator are sufficient but not necessary. Suppose that θ has a mixture distribution which is generated from distributions $\mathrm{CN}(\nu, \kappa)$ and $\mathrm{CN}(\nu + \pi, \kappa)$ with probabilities ε and $(1 - \varepsilon)$, respectively, where $\nu \in [0, \pi]$, $\varepsilon = 0.1$ and $\kappa = 1$. Based on the random sample of size $n = 10$, risk functions of the sample mean direction $\bar{\theta}$ (straight line) and the projection of $\bar{\theta}$ on $[0, \pi]$ (dotted line) under the loss L are plotted in Figure 1. Density of CN distribution and $\bar{\theta}$ are defined in the next subsection. It can be seen that for end points of $\nu \in [0, \pi]$, $\bar{\theta}$ is not improved by its projection. This demonstration refutes the result stated in Theorem 3.1 for an arbitrary estimator in case $l(\mathcal{A}) > (2/3)\pi$.

If an estimator $\delta(\mathbf{Z})$ has a distribution with mixture probability density $\varepsilon f(\delta(\mathbf{z})|h(\boldsymbol{\nu})) + (1 - \varepsilon) f(\delta(\mathbf{z})|h(\boldsymbol{\nu}) + \pi)$, its mean direction is given by

$$\mathsf{CE}\big(\delta(\mathbf{Z})\big) = \begin{cases} h(\boldsymbol{\nu}) + \pi, & \text{if } \varepsilon < 1/2; \\ \text{undefined}, & \text{if } \varepsilon = 1/2; \\ h(\boldsymbol{\nu}), & \text{if } \varepsilon > 1/2. \end{cases}$$

Therefore, when $\varepsilon < 1/2$, $\delta(\mathbf{Z})$ can be treated as an estimator for $h(\boldsymbol{\nu}) + \pi$.

Although Theorem 3.1 is based on a condition (C4) satisfied by the distribution of an estimator when $l(\mathcal{A}) > (2/3)\pi$, examining the distribution of the estimator can be a complex exercise. We try to simplify these conditions for specific cases.

Consider the problem of estimating the location parameter $\nu \in \mathbb{T}$ of a circular random variable θ under the loss L which is invariant under a rotation group

$$\mathcal{G}_1 = \big\{ g_\alpha : g_\alpha(\theta) = (\theta + \alpha) \bmod (2\pi) \big\}. \tag{3.1}$$

Under L, an $\mathcal{G}_1$-equivariant estimator for ν based on a random sample $\theta_1, \ldots, \theta_n$ satisfies

$$\delta(\theta_1, \ldots, \theta_n) = \theta_1 + \xi(\theta_2 - \theta_1, \ldots, \theta_n - \theta_1),$$

where ξ is an arbitrary statistic whose distribution is free from ν. This indicates that distribution of an $\mathcal{G}_1$-equivariant estimator is of the same nature as θ. Using this fact, we deduce the following result from Theorem 3.1.

Corollary 3.1. *If θ is a continuous circular random variable whose distribution is symmetric about $\nu \in \Omega_1$ and satisfies one of the conditions* (C1) *and* (C2) *with $\psi = \nu$ such that $\mathcal{A} = \mathsf{cc}(\Omega_1)$ is K_1-type, that is, $l(\mathcal{A}) \leq \pi$, any $\mathcal{G}_1$-equivariant estimator $\delta(\theta)$ lying outside $\mathcal{A}$ with the positive probability is inadmissible and dominated by its projection on $\mathcal{A}$ under loss* L.

A similar result can be extended to the torus $\mathbb{T}^k = \mathbb{T} \times \cdots \times \mathbb{T}$. A distribution on $\mathbb{T}^k$ can be specified as that of k circular random variables, that is, k-tuple vector $\mathbf{Z} = (\theta_1, \ldots, \theta_k)$ taking values on $\mathfrak{Z} \subset \mathbb{T}^k$. Suppose that all k components are independently distributed and each component has a common location parameter ν. This estimation problem is invariant under a group $\mathcal{G}_2$ given by

$$\mathcal{G}_2 = \big\{\mathbf{g}_\alpha = (g_{1\alpha}, \ldots, g_{k\alpha}) : g_{i\alpha}(\theta_i) = (\theta_i + \alpha) \bmod (2\pi)\big\}. \tag{3.2}$$

Problem of estimating ν can also be thought as multisample problem of estimating common ν. Therefore, we can also draw random samples of different sizes from different components of $\mathbf{Z}$ as components are independently distributed. As in Corollary 3.1, we deduce the following result from Theorem 3.1.

Corollary 3.2. *Let all components of a random variable $\mathbf{Z} = (\theta_1, \ldots, \theta_k)$ taking value on $\mathbb{T}^k$ be independently distributed. If each component has a common location parameter $\nu \in \Omega_1$ and satisfies with respect to Lebesgue measure one of the conditions* (C1) *and* (C2) *with $\psi = \nu$, any $\mathcal{G}_2$-equivariant estimator $\delta(\mathbf{z})$ lying outside $\mathcal{A} = \mathsf{cc}(\Omega_1)$ with the positive probability is inadmissible and dominated by its projection on $\mathcal{A}$ under loss* L *when $l(\mathcal{A}) \leq \pi$.*

Corollaries 3.1 and 3.2 enable us to improve various estimators available in the literature for the circular location ν of several directional distributions. Apart from the maximum likelihood estimator (MLE) δ_{ml}, the following estimators for ν have been proposed on the basis of a random sample $\theta_1, \ldots, \theta_n$.

(E1) (Watson [29], page 135) Sample mean direction $\bar{\theta}$ minimizes $\sum_{i=1}^n d(\theta_i, \alpha)$ over $\alpha \in \mathbb{T}$ and is obtained as $\bar{\theta} = \mathsf{atan}(\sum_{i=1}^n \sin\theta_i / \sum_{i=1}^n \cos\theta_i)$.

(E2) (Mardia and Jupp [19], page 167) Circular median δ_{cm} minimizes $\sum_{i=1}^n d_1(\theta_i, \alpha)$ over $\alpha \in \mathbb{T}$, where $d_1(\alpha, \beta) = \pi - |\pi - |\alpha - \beta||$ for $\alpha, \beta \in \mathbb{T}$.

(E3) (He and Simpson [10]) L_1-estimator δ_{l1} minimizes $\sum_{i=1}^n \{d(\theta_i, \alpha)\}^{1/2}$ over $\alpha \in \mathbb{T}$.

(E4) (Ducharme and Milasevic [5]) Normalized spatial median $\delta_{\text{nsm}} = (\alpha_2^*/\alpha_1^*)$, where (α_1^*, α_2^*) is the solution of

$$\min_{(\alpha_1,\alpha_2)\in\mathbb{R}^2} \sum_{i=1}^{n} \{1 + \alpha_1^2 + \alpha_2^2 - 2(\alpha_1 \cos\theta_i + \alpha_2 \sin\theta_i)\}^{1/2}.$$

(E5) (Neeman and Chang [24], Tsai [26]) Circular Wilcoxon estimator δ_{cw} minimizes $\sum_{i=1}^{n} R_i d_1(\theta_i, \alpha)$ over $\alpha \in \mathbb{T}$, where R_i $(i = 1, \ldots, n)$ is the rank of $\sin(\theta_i - \alpha)$ amongst $\sin(\theta_1 - \alpha), \ldots, \sin(\theta_n - \alpha)$.

Except δ_{cm}, all other estimators are proposed in their Cartesian forms. All the above mentioned estimators are either M-estimator or restricted M-estimator or R-estimator.

Remark 3.2. Due to lack of preference for the zero direction and orientation in the definition of $\mathbb{T}$, rotation-equivariant estimators for circular parameters are preferred. All the above-mentioned estimators enjoy the property of $\mathcal{G}_1$-equivariance when support is $\mathbb{T}$. When support is $\mathbb{T}^k$, these estimators based on sample values $\theta_{11}, \ldots, \theta_{1n_1}, \ldots, \theta_{k1}, \ldots, \theta_{kn_k}$ are also $\mathcal{G}_2$-equivariant. This equivariance property is used in the following section to derive improved estimators.

3.1. Applications of Theorem 3.1

Theorem 3.1 is applicable to a wide variety of estimation problems for directional distributions. In this section, we consider various examples where Theorem 3.1 leads to improvement over traditional estimators.

Example 3.1 (Unimodal distributions on circle $\mathbb{T}$). A circular normal distribution $\text{CN}(\nu, \kappa)$ is defined by the following density:

$$f_2(\theta; \nu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \nu)}, \qquad \theta, \nu \in \mathbb{T}, \kappa > 0,$$

where I_v is the modified Bessel function of the first kind and order v. It is known a priori that $\nu \in \Omega_1$ such that Ω_1 is an arc of the circle. Without loss of generality, we can assume that $\Omega_1 = [0, b]$, where $b \in \mathbb{T}$. The estimation space for ν is

$$\mathcal{A} = \begin{cases} [0, b], & \text{if } b \le \pi; \\ \mathbb{T}, & \text{if } b > \pi. \end{cases}$$

The unrestricted MLE of ν is $\delta_{\text{ml}} = \bar{\theta}$. The circular normal distribution is the only rotationally symmetric distribution for which MLE of the mean direction ν is the sample mean direction $\bar{\theta}$. Maximization of likelihood function over Ω_1 yields the restricted MLE δ_{rml} as

$$\delta_{\text{rml}} = \begin{cases} \bar{\theta}, & \text{if } \bar{\theta} \in [0, b]; \\ b, & \text{if } \bar{\theta} \in (b, \pi + b/2); \\ 0, & \text{if } \bar{\theta} \in [\pi + b/2, 2\pi); \end{cases} \quad \text{or} \quad \begin{cases} \bar{\theta}, & \text{if } \bar{\theta} \in [0, b]; \\ b, & \text{if } \bar{\theta} \in (b, \pi + b/2]; \\ 0, & \text{if } \bar{\theta} \in (\pi + b/2, 2\pi). \end{cases} \tag{3.3}$$

At $\bar{\theta} = \pi + b/2$, δ_{rml} can take two values. Since $\Pr(\bar{\theta} = \pi + b/2) = 0$, both estimators are equivalent. Note that when $b \le \pi$, δ_{rml} is also the projection of $\bar{\theta}$ on $\mathcal{A}$. Corollary 3.1 yields that δ_{rml} improves $\bar{\theta}$ under the loss L when $b \le \pi$. When $b > \pi$, the projection of $\bar{\theta}$ on $\mathcal{A}$ is the same as $\bar{\theta}$. In a similar way, improvements over all other estimators, δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} (as defined in (E2), (E3), (E4), (E5)) can be obtained from Corollary 3.1 and Remark 3.2 for ν when ν is restricted to $[0, b]$ and $b \le \pi$.

Other well-known symmetric unimodal distributions are wrapped Cauchy $\text{WC}(\nu, \rho)$, wrapped normal $\text{WN}(\nu, \rho)$ and cardioid $C(\nu, \rho)$ with the following probability densities in terms of $t = \cos(\theta - \nu)$,

$$f_3(t) = \frac{(1-\rho^2)}{(1+\rho^2-2\rho t)}, \qquad \rho \in (0,1);$$

$$f_4(t) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{i=1}^{\infty} \rho^{i^2} T_i(t), \qquad \rho \in (0,1);$$

$$f_5(t) = (2\pi)^{-1} + \pi^{-1}\rho t, \qquad |\rho| < 1/2,$$

respectively, where T_v is a Chebyshev polynomial of first kind of order v. Jones and Pewsey [13] proposed a family of symmetric unimodal distributions on $\mathbb{T}$ whose densities are provided in terms of t as

$$f_6(t) = \frac{\{\cosh(\kappa\psi)\}^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))}\left\{1 + \tanh(\kappa\psi)t\right\}^{1/\psi}, \qquad \kappa > 0, \psi \in \mathbb{R},$$

where P_v is the associated Legendre function of the first kind of degree v and order 0. Here, we exclude the case of $\kappa = 0$ since it yields the uniform distribution on $\mathbb{T}$. Circular normal, wrapped Cauchy and cardioid distributions are contained in this family corresponding to $\psi = 0, -1$ and 1, respectively.

Another general family of symmetric unimodal distributions on $\mathbb{T}$ contains wrapped α-stable distributions with densities of the following form in terms of t ([19], page 52)

$$f_7(t) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{i=1}^{\infty} \rho^{i^\alpha} T_i(t), \qquad \rho \in (0,1), \alpha = (0,1) \cup (1,2].$$

For $\alpha = 2$, it yields wrapped normal distribution.

Since the distributions corresponding to densities f_6 and f_7 satisfy condition (C1) , Corollary 3.1 and Remark 3.2 yield improvements over all estimators $\bar{\theta}$, δ_{ml}, δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} for the mean direction ν when ν is restricted to $[0, b]$ such that $b \le \pi$.

Example 3.2 (Mixture distributions on $\mathbb{T}$). Let random variable θ be generated from $\text{CN}(\nu, \kappa)$ and $\text{CN}(\nu + \pi, \kappa)$ with probabilities ε and $(1 - \varepsilon)$, respectively. For this distribution, $\zeta(t) = e^{2\kappa t}$ which is increasing in t. Thus, Corollary 3.1 and Remark 3.2 yield that all the estimators, $\bar{\theta}$, δ_{ml}, δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw}, for $\nu \in \Omega_1$ are improved by their projections on $\mathcal{A}$ if $l(\mathcal{A}) \le \pi$ and $\varepsilon \ge 1/2$. For the values of $\varepsilon < 1/2$, improvements are possible when $l(\mathcal{A}) \le (2/3)\pi$.

Similar improvements are possible when we mix either two wrapped normal or two distributions with density f_6 for same value of $\psi \leq 1$ with different mean directions ν and $\nu + \pi$. To apply Corollary 3.1, we must show that the corresponding function $\zeta(t)$ is an increasing function in t. For a general density $f_6(t)$, derivative of $\zeta(t)$ with respect to t is obtained as

$$\zeta'(t) = \frac{2(1-\psi)}{\psi} \frac{\tanh(\kappa\psi)}{\{1+\tanh(\kappa\psi)t\}^2} \left\{ \frac{1+\tanh(\kappa\psi)t}{1-\tanh(\kappa\psi)t} \right\}^{1/\psi}, \qquad \psi \neq 0,$$

which is always nonnegative unless $\psi > 1$.

The wrapped normal distribution can be represented by theta function ϑ_3. Using the representation of ϑ_3 in terms of infinite products ([9], page 921, equation 8.181.2), we have

$$f_4(t) = \frac{1}{2\pi} \prod_{i=1}^{\infty} \bigl(1-\rho^{2i}\bigr)\bigl(1+2t\rho^{2i-1}+\rho^{2(i-1)}\bigr).$$

Second derivative of $f_4(t)$ with respect to t is $f_4(t)\sum_{i\neq j}\xi_i(t)\xi_j(t)$, where $\xi(i) = 2\rho^{2i-1}/\{1+2t\rho^{2i-1}+\rho^{2(i-1)}\}^{-1}$. Convexity of $f_4(t)$ follows from the positiveness of $\xi_i(t)$ and increasing nature of $\zeta(t)$ follows from the convexity of $f_4(t)$ using Remark 2.3.

Example 3.3 (Distributions on $\mathbb{T}$ with k-fold rotational symmetry). This distribution is constructed by putting k copies of the original distribution end-to-end ([19], page 53). If we are given a distribution of θ which is unimodal and symmetric about ν, constructed distribution has the density $f(\cos(k(\theta-\nu_0)))$, $\theta \in \mathbb{T}$, $\nu_0 \in [0, 2\pi/k)$. Note that new distribution is k-modal, for example, k-modal circular normal distribution ([12], page 209). For $k \geq 3$, condition (C3) of Theorem 3.1 is satisfied and so estimators of ν_0 lying outside $\mathcal{A}$ with a positive probability can be improved by their projections on $\mathcal{A}$. Note that the results hold when the parameter space is full, that is, $[0, 2\pi/k)$ or restricted, that is, a subset of $[0, 2\pi/k)$. For $k = 2$, the result holds only for restricted parameter space if $\nu_0 \in \Omega_1 \subset [0, b]$ such that $b \leq (2/3)\pi$.

Example 3.4 (Distributions on torus $\mathbb{T}^k$). Suppose that all k components of $\mathbf{Z} = (\theta_1, \ldots, \theta_k) \in \mathfrak{Z} \subset \mathbb{T}^k$ are independently distributed and ith component θ_i follows $\mathrm{CN}(\nu, \kappa_i)$ with known κ_i. Consider the estimation of common $\nu \in \Omega_1$ under the loss L. Let $(\theta_{i1}, \ldots, \theta_{in_i})$ be a random sample from the ith population $(i = 1, \ldots, n)$. Suppose $\bar{\theta}_i$ denotes the sample mean direction and $R_i = \{(\sum_{j=1}^{n_i} \sin\theta_{ij})^2 + (\sum_{j=1}^{n_i} \cos\theta_{ij})^2\}^{1/2}$ denotes the sample resultant length for the sample of ith component. The MLE of ν is

$$\tilde{\theta} = \operatorname{atan}\left(\frac{\sum_{i=1}^{k} \kappa_i R_i \sin\bar{\theta}_i}{\sum_{i=1}^{k} \kappa_i R_i \cos\bar{\theta}_i} \right). \tag{3.4}$$

The conditional distribution of $\tilde{\theta}$ is again circular normal $\mathrm{CN}(\nu, R^*)$, where

$$R^* = \left\{ \left(\sum_{i=1}^{k} \kappa_i R_i \sin\bar{\theta}_i \right)^2 + \left(\sum_{i=1}^{k} \kappa_i R_i \cos\bar{\theta}_i \right)^2 \right\}^{1/2} \qquad \text{Holmquist [11].}$$

Since the distribution of R^* is dependent only on κ_i, distribution of $\tilde{\theta}$ is unimodal and symmetric about its mode ν. Corollary 3.2 yields that $\tilde{\theta}$ is dominated by its projection on $\mathcal{A}$ under the loss L if $l(\mathcal{A}) \le \pi$. Note that improved estimator of $\tilde{\theta}$ is also the restricted MLE for common ν when $l(\mathcal{A}) \le \pi$.

Using Corollary 3.2 and Remark 3.2, we can obtain improvements over other estimators $\bar{\theta}$, δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} based on random sample $\theta_{11}, \ldots, \theta_{1n_1}, \ldots, \theta_{k1}, \ldots, \theta_{kn_k}$.

Remark 3.3. This model can be further extended to the cases where distributions of independent components are not necessarily the same and satisfy conditions (C1) and (C2) with respect to Lebesgue measure, namely, distributions considered in Examples 3.1 and 3.2.

Example 3.5 (Distribution on unit sphere $\mathbb{S}_2$***).*** Point $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{S}_2$ can be specified by its geographical coordinates: colatitude $\theta = \cos^{-1}(x_3) \in [0, \pi]$ and longitude $\phi = \mathsf{atan}(x_2/x_1) \in [0, 2\pi)$. Hence, sphere $\mathbb{S}_2$ is isomorphic to $\mathbb{T}_2 = [0, \pi] \times [0, 2\pi)$. The density of Fisher distribution on the support $\mathbb{T}_2$ is given by

$$f_8(\theta, \phi; \nu_1, \nu_2, \kappa) = \frac{\kappa \sin\theta}{4\pi \sinh(\kappa)} \exp\big\{\kappa\big(\sin\theta \sin\nu_1 \cos(\phi - \nu_2) + \cos\theta \cos\nu_1\big)\big\},$$
$$(\nu_1, \nu_2) \in \mathbb{T}_2, \kappa > 0,$$

where (ν_1, ν_2) is the mean direction. On the basis of a random sample of size n, the MLE of ν_2, when κ is known/unknown and ν_1 is unknown, is given by

$$\delta_{\mathrm{ml}} = \mathsf{atan}\left(\frac{\sum_{i=1}^n \sin\theta_i \sin\phi_i}{\sum_{i=1}^n \sin\theta_i \cos\phi_i}\right).$$

It is known that mean direction is restricted to a continuous arc of hemisphere. Without loss of generality, we can assume that $\nu_1 \in [0, \pi]$ and $\nu_2 \in [0, b]$ with $b \le \pi$. We can improve the MLE of ν_2 by its projection on $[0, b]$ if $b \le (2/3)\pi$ using Theorem 3.1.

Example 3.6 (Distribution on cylinder $\mathbb{R} \times \mathbb{T}$***).*** Mardia and Sutton [20] proposed a distribution on the cylinder $\mathbb{R} \times \mathbb{T}$. Let (X, θ) have the support $\mathbb{R} \times \mathbb{T}$ where the marginal distribution of θ is $\mathrm{CN}(\nu_0, \kappa)$, $\nu_0 \in \mathbb{T}, \kappa > 0$ and conditional distribution of $X|\theta$ is a normal with mean

$$\mu_c = \mu + \sigma\rho\sqrt{\kappa}\big\{\cos(\theta - \nu) - \cos(\nu_0 - \nu)\big\}, \qquad \mu \in \mathbb{R}, \nu \in \mathbb{T}, 0 \le \rho \le 1, \sigma > 0$$

and variance $\sigma^2(1 - \rho^2)$. Based on a random sample of size n, the MLE of ν is

$$\delta_{\mathrm{ml}} = \mathsf{atan}\left(\frac{s_2}{s_3}\frac{r_{23}r_{12} - r_{13}}{r_{23}r_{13} - r_{12}}\right),$$

where for $i = 1, 2, 3$ and $j = 1, \ldots, n$, $x_{1j} = x_j$, $x_{2j} = \cos\theta_j$, $x_{3j} = \sin\theta_j$, $\bar{x}_i = \sum_j x_{ij}/n$, $s_i^2 = \sum_j (x_{ij} - \bar{x}_i)^2$; and for $i \ne k$ $(i, k = 1, 2, 3)$,

$$r_{ik} = \frac{1}{s_i s_k} \sum_j (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_k).$$

If ν is restricted to Ω_1 such that $l(\mathcal{A}) \leq (2/3)\pi$, δ_{ml} is dominated by its projection on $\mathcal{A}$ under the loss L using Theorem 3.1. Note that simulations indicate that this result is also valid for $(2/3)\pi < l(\mathcal{A}) \leq \pi$.

4. An inadmissibility result for general equivariant rules

Moors [21,22], Kumar and Sharma [16] gave a general method for obtaining improved equivariant estimators of parameters in Euclidean spaces. In this section, we extend these results for estimating circular parameters.

Let the problem of estimating $h(\boldsymbol{\nu}) \in \Omega_1$ under the loss L be invariant under a compact group $\mathcal{G}$ of measurable transformations $\mathbf{g} : \mathfrak{Z} \to \mathfrak{Z}$. There exists a finite and left (right) invariant Haar measure λ on $\mathcal{G}$ ([6], Theorem 1.5). Let $\bar{\mathcal{G}}$ and $\tilde{\mathcal{G}}$ be the groups induced by $\mathcal{G}$ on parameter space Ω and estimation space $\mathcal{A} = \mathsf{cc}(\Omega_1)$.

Lemma 4.1. *For the $\mathcal{G}$-invariant estimation problem defined above, we have*:

(i) $f(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})) = f(\mathbf{g}^{-1}(\mathbf{z})|\boldsymbol{\nu})$ *a.e. with respect to measure* η;
(ii) $h\bar{\mathbf{g}}(\boldsymbol{\nu}) = \tilde{g}h(\boldsymbol{\nu}) \bmod (2\pi)$ *for all* $\boldsymbol{\nu} \in \Omega$;
(iii) *Let* $A \in \mathfrak{C}_5$ (*closed convex*). *If* ϕ_0 *is the projection of* $\phi \in \mathbb{T}$ *on* A, *the projection of* $\tilde{g}(\phi)$ *on* $\tilde{g}(A)$ *is* $\tilde{g}(\phi_0)$.

For each $\mathbf{z} \in \mathfrak{Z}$ and $\boldsymbol{\nu} \in \Omega$, define a probability measure on $\mathcal{G}$ as

$$\tau\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big) = \frac{f(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))}{\int_{\mathcal{G}} f(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}.$$

With the help of these measures, for a fixed $\mathbf{z} \in \mathfrak{Z}$, define a function $h_{\mathbf{z}} : \Omega \to \mathcal{A}$ as

$$h_{\mathbf{z}}(\boldsymbol{\nu}) = \begin{cases} \mathsf{atan}\left(\dfrac{\int_{\mathcal{G}} \sin(\tilde{g}h(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g})}{\int_{\mathcal{G}} \cos(\tilde{g}h(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g})}\right), & \text{if } \displaystyle\int_{\mathcal{G}} f\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\,\mathrm{d}\lambda(\mathbf{g}) > 0; \\ h(\boldsymbol{\nu}), & \text{if } \displaystyle\int_{\mathcal{G}} f\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\,\mathrm{d}\lambda(\mathbf{g}) = 0. \end{cases} \tag{4.1}$$

The new estimation space $\mathcal{A}_{\mathbf{z}}$ is defined as a convex closure of $h_{\mathbf{z}}(\Omega)$, that is, $\mathcal{A}_{\mathbf{z}} = \mathsf{cc}(h_{\mathbf{z}}(\Omega))$.

Remark 4.1. In case of $\int_{\mathcal{G}} f(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}) = 0$, $\mathcal{A}_{\mathbf{z}} = \mathcal{A}$. When $\int_{\mathcal{G}} f(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}) > 0$, $h_{\mathbf{z}}(\boldsymbol{\nu})$ can be written as

$$h_{\mathbf{z}}(\boldsymbol{\nu}) = \mathsf{atan}\left(\frac{\mathsf{E}\sin(\tilde{g}h(\boldsymbol{\nu}))}{\mathsf{E}\cos(\tilde{g}h(\boldsymbol{\nu}))}\right), \tag{4.2}$$

where the expectation is taken over $\mathbf{g}$ with respect to a probability measure $\tau(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g})$. If $\Pr(\tilde{g}h(\boldsymbol{\nu}) \in \mathcal{A}) = 1$, $h_{\mathbf{z}}(\boldsymbol{\nu}) \in \mathcal{A}$ from the convexity of $\mathcal{A}$ (see Proposition 2.2). Since $\mathcal{A}_{\mathbf{z}}$ is the smallest convex set containing $h_{\mathbf{z}}(\boldsymbol{\nu})$, we conclude that $\mathcal{A}_{\mathbf{z}} \subset \mathcal{A}$.

We assume that every transformation $\tilde{g} \in \tilde{\mathcal{G}}$ satisfies circular property, that is, for a fixed α and for all $\phi \in \mathcal{A}$, $\tilde{g}(\phi)$ is either $\alpha + \phi$ or $\alpha - \phi$ such that images are in $\mathcal{A}$ itself.

Lemma 4.2. *If every induced transformation on $\mathcal{A}$ satisfies the circular property, the reduced estimation space $\mathcal{A}_{\mathbf{z}}$ satisfies $\mathcal{A}_{\mathbf{g(z)}} = \tilde{g}(\mathcal{A}_{\mathbf{z}})$ for all $\mathbf{z} \in \mathfrak{Z}$ and $\mathbf{g} \in \mathcal{G}$.*

Lemmas 4.1 and 4.2 are utilized to prove the main result of this section.

Theorem 4.1. *Consider an $\mathcal{G}$-invariant estimation problem under the loss function L with a compact group $\mathcal{G}$ such that elements of induced group $\tilde{\mathcal{G}}$ on the estimation space $\mathcal{A}$ satisfy the circular property. Any $\mathcal{G}$-equivariant estimator δ satisfying $\Pr_{\boldsymbol{\nu}}(\delta(\mathbf{Z}) \notin \mathcal{A}_{\mathbf{Z}}) > 0$ for some $\boldsymbol{\nu} \in \Omega$ is dominated by its projection on $\mathcal{A}_{\mathbf{Z}}$ provided that $l(\mathcal{A}_{\mathbf{Z}}) \leq (2/3)\pi$.*

This result is applicable to both restricted and unrestricted estimation problems as illustrated in the following subsection.

4.1. Applications of Theorem 4.1

We consider estimation of the location parameter $\nu \in \Omega_1$ of a circular random variable θ under the loss L. Let us denote by $f_\nu(\theta)$ the density of θ that would be a function of $\cos(\theta - \nu)$.

4.1.1. *Unrestricted estimation problems*

The estimation problem is invariant under the rotation group $\mathcal{G}_1$. Clearly, the induced group $\tilde{\mathcal{G}}_1$ on the estimation space $\mathcal{A} = \mathbb{T}$ is itself $\mathcal{G}_1$ and every transformation in $\mathcal{G}_1$ satisfies the circular property. Taking $\mathcal{G} = \mathcal{G}_1$, we define

$$h_\theta(\nu) = \operatorname{atan}\left(\frac{\int_0^{2\pi} \sin(\alpha+\nu) f_{\alpha+\nu}(\theta)\,\mathrm{d}\alpha}{\int_0^{2\pi} \cos(\alpha+\nu) f_{\alpha+\nu}(\theta)\,\mathrm{d}\alpha}\right) = \operatorname{atan}\left(\frac{\int_0^{2\pi} \sin\alpha f_\alpha(\theta)\,\mathrm{d}\alpha}{\int_0^{2\pi} \cos\alpha f_\alpha(\theta)\,\mathrm{d}\alpha}\right),$$

or equivalently, $h_\theta(\nu)$ is constant on $\nu \in \mathbb{T}$. Therefore the following result follows from Theorem 4.1.

Corollary 4.1. *If θ is a circular random variable with the unrestricted location parameter $\nu \in \mathbb{T}$, there is only one admissible $\mathcal{G}_1$-equivariant estimator under the loss L which is obtained as*

$$\delta_{\mathrm{ad}} = \operatorname{atan}\left(\frac{\int_0^{2\pi} \sin\alpha f_\alpha(\theta)\,\mathrm{d}\alpha}{\int_0^{2\pi} \cos\alpha f_\alpha(\theta)\,\mathrm{d}\alpha}\right),$$

where $f_\nu(\theta)$ the density of θ.

For $\mathrm{CN}(\nu, \kappa)$ distribution, $\delta_{\mathrm{ad}} = \bar{\theta}$. All other $\mathcal{G}_1$-equivariant estimators δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} for $\nu \in \mathbb{T}$ under the loss L are improved by $\bar{\theta}$ using Corollary 4.1. This result is significant in the sense that so far comparison of $\bar{\theta}$ with δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} was done only with respect to asymptotic efficiency and robustness.

4.1.2. *Restricted estimation problems*

Let the location ν be restricted to any arc of semicircle. Without loss of generality, we can assume that $\nu \in \Omega_1 = [0, b]$ with $0 < b \leq \pi$. In Example 3.1, the same restricted space estimation problem was considered and we obtained improvements over estimators $\bar{\theta}$, δ_{ml}, δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} for various distributions. Denote by $\bar{\theta}^*$, δ^*_{ml}, δ^*_{cm}, δ^*_{l1}, δ^*_{nsm} and δ^*_{cw}, the dominating estimators of $\bar{\theta}$, δ_{ml}, δ_{cm}, δ_{l1}, δ_{nsm} and δ_{cw} as their projections on $\Omega_1 = [0, b]$. Here, we can further improve upon these improved estimators.

We consider the transformation group as $\mathcal{G}_3 = \{e, g\}$ where $g(\theta) = (b - \theta) \bmod (2\pi)$ and e is the identity transformation. The estimation problem remains invariant under $\mathcal{G}_3$ and the induced group on the estimation space $\mathcal{A}$ is $\mathcal{G}_3$. Clearly, elements of $\mathcal{G}_3$ satisfy the circular property. Define the following function for a fixed random sample $\theta_1, \ldots, \theta_n$ as

$$h_{(\theta_1,\ldots,\theta_n)}(\nu) = \left\{b/2 + \tan^{-1}\big(a(\nu)\big)\right\} \bmod (2\pi),$$

where after some algebraic computations, $a(\nu)$ is derived as

$$a(\nu) = \frac{\prod_{i=1}^n f_{b-\nu}(\theta_i) - \prod_{i=1}^n f_{\nu}(\theta_i)}{\prod_{i=1}^n f_{b-\nu}(\theta_i) + \prod_{i=1}^n f_{\nu}(\theta_i)} \tan\left(\frac{b}{2} - \nu\right).$$

Since $f_\nu(\theta)$ is a function of $\cos(\theta - \nu)$, $h_{(\theta_1,\ldots,\theta_n)}(\nu)$ is symmetric about $\nu = b/2$. It is sufficient to assume that $\nu \in [0, b/2]$ to study the behaviour of the function $h_{(\theta_1,\ldots,\theta_n)}(\nu)$.

We consider distributions for which $a(\nu)$ is monotonic in $\nu \in [0, b/2]$. As $a(b/2) = 0$, monotonic nature of $a(\nu)$ is dependent on the sign of $a(0)$. The new estimation space is

$$\mathcal{A}_{(\theta_1,\ldots,\theta_n)} = \mathsf{cc}\big(h_{(\theta_1,\ldots,\theta_n)}(\Omega_1)\big) = \begin{cases} \big[b^*, b/2\big], & \text{if } a(0) < 0; \\ \{b/2\}, & \text{if } a(0) = 0; \\ \big[b/2, b^*\big], & \text{if } a(0) > 0, \end{cases}$$

where $b^* = h_{(\theta_1,\ldots,\theta_n)}(0)$. It may be noted that $l(\mathcal{A}_{(\theta_1,\ldots,\theta_n)}) \leq b/2$ which is a substantial reduction in $l(\mathcal{A}) = b$.

Since $\mathcal{G}_3 \subset \mathcal{G}_1$, all estimators $\bar{\theta}$, δ_{ml}, δ_{cm}, δ_{l1}, δ_{nsm}, δ_{cw}, $\bar{\theta}^*$, δ^*_{ml}, δ^*_{cm}, δ^*_{l1}, δ^*_{nsm} and δ^*_{cw} are also $\mathcal{G}_3$-invariant. For the distributions which satisfy the assumption of monotonicity of $a(\nu)$, all these estimators can be improved by their projections on $\mathcal{A}_{(\theta_1,\ldots,\theta_n)}$ using Theorem 4.1.

If θ follows a $\mathrm{CN}(\nu, \kappa)$ distribution, the function $a(\nu)$ is given as

$$a(\nu) = \tanh\left(\kappa r \sin\left(\tilde{\theta} - \frac{b}{2}\right) \sin\left(\frac{b}{2} - \nu\right)\right) \tan\left(\frac{b}{2} - \nu\right).$$

Monotonicity of $a(\nu)$ can be easily observed. The new estimation space is equal to

$$\mathcal{A}_{(\bar{\theta},r)} = \begin{cases} \big[b^*, b/2\big], & \text{if } \bar{\theta} \in J_1; \\ \{b/2\}, & \text{if } \bar{\theta} \in \{b/2, \pi + b/2\}; \\ \big[b/2, b^*\big], & \text{if } \bar{\theta} \in I_1, \end{cases}$$

where $b^* = b/2 + \tan^{-1}[\tan(b/2)\tanh\{\kappa r \sin(\bar{\theta} - b/2)\sin(b/2)\}]$, $J_1 = J(b/2, \pi + b/2)$ and $I_1 = I(b/2, \pi + b/2)$. Based on $\mathcal{A}_{(\bar{\theta},R)}$, an estimator δ is dominated by

$$\delta_I = \begin{cases} \delta, & \text{if } \bar{\theta} \in J_1 \text{ and } \delta \in [b^*, b/2], \text{ or, } \bar{\theta} \in I_1 \text{ and } \delta \in [b/2, b^*]; \\ b^*, & \text{if } \bar{\theta} \in J_1 \text{ and } \delta \in J(b^*, \gamma), \text{ or, } \bar{\theta} \in I_1 \text{ and } \delta \in (b^*, \gamma); \\ b/2, & \text{if } \bar{\theta} \in J_1 \text{ and } \delta \in (b/2, \gamma], \text{ or, } \bar{\theta} \in I_1 \text{ and } \delta \in J(b/2, \gamma], \text{ or, } \bar{\theta} \in \{b/2, \pi + b/2\}, \end{cases}$$

where $\gamma = \pi + b^*/2 + b/4$. In Example 3.1, the estimator δ for ν is dominated by the projection of δ on $\mathcal{A}$. Let us denote by δ^* this improved estimator. Based on $\mathcal{A}_{(\bar{\theta},R)}$, improved estimators δ_I and δ_I^* of δ and δ^*, respectively, are equivalent except when either $\delta \in [\gamma, \pi + b/2]$ and $\bar{\theta} \in J_1$ or $\delta \in [\pi + b/2, \gamma]$ and $\bar{\theta} \in I_1$. If $\delta = \bar{\theta}$, both $\bar{\theta}_I$ and $\bar{\theta}_I^*$ are equivalent and given by

$$\begin{cases} b^*, & \text{if } \begin{cases} \bar{\theta} < b^* \text{ and } \bar{\theta} \in [0, b/2), \\ \bar{\theta} > b^* \text{ and } \bar{\theta} \in (b/2, \pi + b/2), \\ \bar{\theta} \in (\pi + b/2, 2\pi), \end{cases} \\ b/2, & \text{if } \bar{\theta} = \pi + b/2, \\ \bar{\theta}, & \text{elsewhere.} \end{cases} \tag{4.3}$$

In Figure 2, we have plotted the risk functions of the MLE $\bar{\theta}$, restricted estimator δ_{rml} defined in (3.3) and improved estimator $\bar{\theta}_I$ defined in (4.3) under the loss function L. The risk values have been evaluated using simulations. For this, we have generated 100 000 samples from a $\mathrm{CN}(\nu, \kappa)$ distribution for various values of (n, κ, b). The following conclusions can be made from the numerical study.

(a) The risk function of $\bar{\theta}$ is constant for a fixed value of (n, κ). The risk functions of δ_{rml} and $\bar{\theta}_I$ are symmetric about $b/2$. For small values of κ or b, these risk functions are strictly decreasing in $\nu \in [0, b/2]$. For higher values of κ or b, behaviour is reverse.

(b) For all the values of (n, κ) and $\nu \in [0, b]$, δ_{rml} uniformly improves $\bar{\theta}$ and $\bar{\theta}_I$ uniformly improves δ_{rml} when $b \in (0, \pi]$. Risk values of δ_{rml} and $\bar{\theta}_I$ are the same when $b = \pi$ and are less than that of $\bar{\theta}$ for all $\nu \in [0, b]$.

(c) The amount of relative improvement of δ_{rml} over $\bar{\theta}$ is increasing as κ or b decreases. This is seen to be as high as 95%. Similarly, relative improvement of $\bar{\theta}_I$ over δ_{rml} is seen to be up to 75%.

Similar observations have been made for various other configurations of (n, κ, b). Simulations for various directional distributions show significant improvements. We omit details here.

Remark 4.2. For the support $\mathbb{T}$, both unrestricted and restricted estimation problems discussed in Sections 4.1.1 and 4.1.2 can be easily extendable to the case of support $\mathbb{T}^k$.

5. Concluding remarks

For estimating parameters in Euclidean spaces, with respect to the loss function as an increasing function of distance, the action space is taken to be the smallest convex set containing the estimand space. If an estimator lies outside it with a positive probability, an improvement is obtained

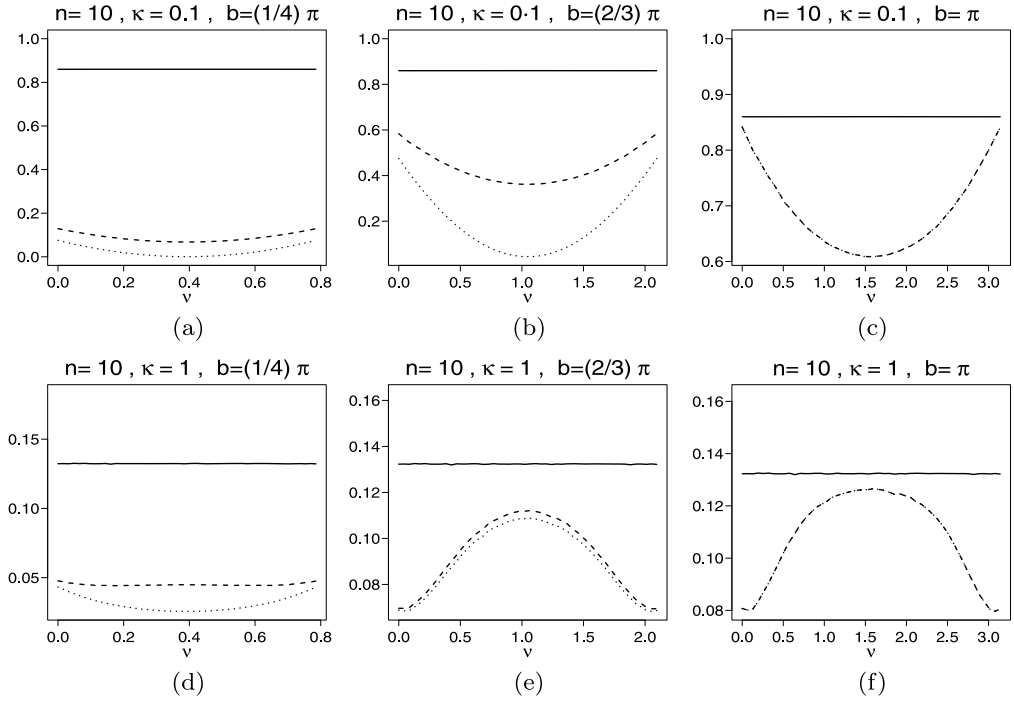

Figure 2. Risk plots of (a) $\bar{\theta}$ (straight line), (b) δ_{rml} (dashed line) and (c) $\bar{\theta}_I$ (dotted line) under the loss function L when $b \in (0, \pi]$.

by projecting this estimator on the action space. In Section 3, we have extended this concept to the estimation of circular parameters of directional distributions. The result of Theorem 3.1, is not exactly analogous to the result for Euclidean spaces. Further, [16,22] gave a new technique for improving equivariant estimators in Euclidean spaces. In Section 4, we have developed a theory to extend this technique to circular parameters. The results have been applied to various estimation problems in directional distributions. The resulting estimators are seen to show significant improvements over the usual estimators.

It would be interesting to further extend these results to parameters lying in spheres of higher dimensions.

Appendix

A.1. Proof of Proposition 2.1

The statement trivially follows when $A = \mathbb{T}$. From Lemma 2.2, it is sufficient to consider A to be of I-type. First, we assume $A = I[0, \gamma]$ with $\gamma \in (0, \pi]$. Note that for $\gamma = 0$, the proof is trivial.

Let $\{\phi_1, \ldots, \phi_n\}$ be linearly ordered sample of $n(\geq 2)$ observations taking values in A. After rotating sample by an angle $2\pi - \phi_1$, modified ordered sample $\{0, \alpha_2, \ldots, \alpha_n\}$ takes values in $I[0, \alpha_n] \subset A$ with $\alpha_i = \phi_i - \phi_1$ $(i = 2, \ldots, n)$. We have to prove

$$\bar{\alpha}_w = \operatorname{atan}\left(\frac{\sum_{i=2}^n w_i \sin\alpha_i}{w_1 + \sum_{i=2}^n w_i \cos\alpha_i}\right) \in I[0, \alpha_n],$$

where weights are $w_1, \ldots, w_n$. From (2.2), the lower bound of $\bar{\alpha}_w$ is zero. If $\alpha_n \neq \pi$, both $\sum_{i=2}^n w_i \sin\alpha_i$ and $w_1 + \sum_{i=2}^n w_i \cos\alpha_i$ cannot be zero simultaneously, that is, $\bar{\alpha}_w$ always exists. When $n = 2$ and $\alpha_2 = \pi$, $\bar{\alpha}_w$ does not exist if $w_1 = w_2$.

(i) Consider $\alpha_n < \pi/2$. Since $\tan(\cdot)$ is an increasing function in $[0, \pi/2)$ and $w_1 + \sum_{i=2}^n \cos\alpha_i$ is positive, $\bar{\alpha}_w \leq \alpha_n$ is equivalent to

$$\sum_{i=2}^{n-1} w_i \sin\alpha_i \leq \tan\alpha_n \left(w_1 + \sum_{i=2}^{n-1} w_i \cos\alpha_i\right). \tag{A.1}$$

Induction method is used to prove the inequality (A.1). For $n = 2$, (A.1) is reduced to $w_1 \tan\alpha_2 \geq 0$ which always holds. We will show that (A.1) is true for $n = k + 1$ after using it for $n = k$. Thus,

$$\tan\alpha_{k+1}\left(w_1 + \sum_{i=2}^{k} w_i \cos\alpha_i\right) \geq \tan\alpha_k \left(w_1 + \sum_{i=2}^{k} w_i \cos\alpha_i\right) = \sum_{i=1}^{k} w_i \sin\alpha_i.$$

The above steps follow since $\tan\alpha_k \leq \tan\alpha_{k+1}$ and (A.1) is assumed to be true for $n = k$.

(ii) Now consider the case when $\pi/2 < \alpha_n < \pi$. Since $w_1 + \sum_{i=2}^n w_i \cos\alpha_i < 0$ and $\tan(\cdot)$ is increasing in $(\pi/2, \pi]$, $\bar{\alpha}_w \leq \alpha_n$ is equivalent to the reverse of inequality (A.1). The proof can be completed as above.

(iii) Cases when $\alpha_n = \pi/2$ and $\alpha_n = \pi$ are straightforward.

Hence, the proposition is established for $A = K_1[\alpha, \beta]$ with $\alpha, \beta \in \mathbb{T}$. In a similar manner, it can be proved when A is of types $K_1(\alpha, \beta)$, $K_1(\alpha, \beta]$ and $K_1[\alpha, \beta)$.

A.2. Proof of Proposition 2.2

The statement is trivially true when $A = \mathbb{T}$. As in the proof of Proposition 2.1, it is enough to prove the result only for $A = [0, \gamma]$ with $\gamma \in (0, \pi]$. From (2.2), $\mathsf{CE}(\theta) \geq 0$, if it exists. Every interval is convex subset of $\mathbb{R}$. Using the fact that an analogous result is true for the expectation of random variables in $\mathbb{R}$ such as $\sin\theta$ and $\cos\theta$ ([7], page 74), we can observe ranges of both $\mathsf{E}\sin\theta$ and $\mathsf{E}\cos\theta$.

Suppose that $\Pr(\theta \in A) = 1$ and so $\Pr(\gamma - \theta \in A) = 1$. Since $\sin(\cdot)$ is nonnegative in the range $[0, \pi]$, $\Pr(\sin(\gamma - \theta) \geq 0) = 1$. Thus,

$$\mathsf{E}\sin(\gamma - \theta) \geq 0. \tag{A.2}$$

(i) Consider the case $0 < \gamma < \pi/2$. Since $\Pr(\cos\theta > 0) = 1$, $\mathsf{E}\cos\theta > 0$. Dividing both sides of (A.2) by $\cos\theta(\mathsf{E}\cos\theta)$ (term is positive with probability one), we get

$$(\mathsf{E}\sin\theta/\mathsf{E}\cos\theta) \le \tan\gamma. \tag{A.3}$$

Using facts that $\tan^{-1}(\cdot)$ is increasing in $[0, \infty)$ and $\mathsf{atan}(\cdot) = \tan^{-1}(\cdot)$ in the range $[0, \pi/2)$, we obtain $\mathsf{CE}(\theta) \le \gamma$.

(ii) Consider $\gamma = \pi/2$. Since both $\mathsf{E}\sin\theta$ and $\mathsf{E}\cos\theta$ are nonnegative and cannot be zero simultaneously, $\mathsf{CE}(\theta) \le \pi/2$.

(iii) If $\pi/2 < \gamma < \pi$, $\Pr(\cos\theta < 0) = 1$ and so $\mathsf{E}\cos\theta$ is negative. Dividing both sides of (A.2) by $\cos\theta(\mathsf{E}\cos\theta)$ (positive quantity), we again obtain (A.3). As $\tan^{-1}(\cdot)$ is increasing in the range $(-\infty, 0]$, we have

$$\tan^{-1}(\mathsf{E}\sin\theta/\mathsf{E}\cos\theta) \le \gamma - \pi.$$

Since $\mathsf{atan}(\cdot) = \tan^{-1}(\cdot) + \pi$, $\mathsf{CE}(\theta) \le \gamma$.

(iv) Now consider that $\gamma = \pi$. Both $\sin\theta \in [0, 1]$ and $\cos\theta \in [-1, 1]$ with probability one. Hence $\mathsf{CE}(\theta) \le \pi$, if it exists.

Existence of $\mathsf{CE}(\theta)$ is not confirmed only for the case (iv). Suppose that $\Pr(\theta = 0) = \Pr(\theta = \pi) = 1/2$, both $\mathsf{E}\sin\theta$ and $\mathsf{E}\cos\theta$ are zero, so $\mathsf{CE}(\theta)$ does not exist. Moreover, if $A = \mathbb{T}$ and θ follows uniform distribution on $\mathbb{T}$ with the density $(2\pi)^{-1}$, $\mathsf{CE}(\theta)$ does not exist.

A.3. Proof of Lemma 2.5

Using Lemma 2.2, it is enough to consider the case $A = [0, b]$ with $b \in (2\pi/3, \pi]$. Decompose the set A as $A_1 \cup A_2 \cup A_3$, where

$$A_1 = [0, 3b/4 - \pi/2), \quad A_2 = [3b/4 - \pi/2, b/2], \qquad A_3 = (b/2, b],$$

and complement of A as $B_1 \cup B_2$, where

$$B_1 = (b, \pi + b/2], \qquad B_2 = (\pi + b/2, 2\pi).$$

We have to show for all $\psi \in A$,

$$a(\psi) = \mathsf{E}^{\theta}_{\psi}\{u(\theta)\} = \int_{\theta \in B_1 \cup B_2} u(\theta) f_1(\theta|\psi)\,\mathrm{d}\theta > 0, \tag{A.4}$$

where $f_1(\theta|\psi)$ is the probability density of θ with respect to Lebesgue measure and $u(\theta)$ is given by

$$u(\theta) = \cos(\theta_0 - \psi) - \cos(\theta - \psi), \qquad \theta \notin A, \psi \in A,$$

with θ_0 as the projection of θ on A. According to the assumption, the distribution of θ is symmetric about ψ, therefore, we have

$$\mathsf{E}^{\theta}_{\psi}\cos(\theta_0 - \psi) = \mathsf{E}^{\theta}_{b-\psi}\cos(\theta_0 - b + \psi).$$

Hence, $a(\psi) = a(b - \psi)$ for $\psi \in A$. It is enough to prove that $a(\psi) > 0$ for $\psi \in A_1 \cup A_2$.

Next, we examine the sign of the function $u(\cdot)$. Define $u(\theta) = u_1(\theta) + u_2(\theta)$, where from definition of projection θ_0, $u_1(\theta)$ and $u_2(\theta)$ are given by

$$u_1(\theta) = \begin{cases} 2\sin\left(\dfrac{\theta - b}{2}\right)\sin\left(\dfrac{\theta + b}{2} - \psi\right), & \text{if } \theta \in B_1; \\ 0, & \text{if } \theta \in B_2; \end{cases}$$

$$u_2(\theta) = \begin{cases} 0, & \text{if } \theta \in B_1; \\ 2\sin\left(\dfrac{\theta}{2}\right)\sin\left(\dfrac{\theta}{2} - \psi\right), & \text{if } \theta \in B_2. \end{cases}$$

If $\theta \in B_1$, $(\theta - b) \in (0, \pi - b/2] \subset (0, 2\pi/3)$. Hence, sign of $u_1(\theta)$ is only dependent on that of $\sin((\theta + b)/2 - \psi)$. When $\psi \in A_1$, decompose B_1 as $B_1 = B_{11} \cup B_{12}$, where

$$B_{11} = (b, 2\pi + 2\psi - b], \qquad B_{12} = (2\pi + 2\psi - b, \pi + b/2].$$

It can be noted that

$$\left(\frac{\theta + b}{2} - \psi\right) \in \begin{cases} (b - \psi, \pi] \subset (2\pi/3, \pi], & \text{if } \theta \in B_{11}; \\ (\pi, \pi/2 + 3b/4 - \psi] \subset (\pi, 5\pi/4], & \text{if } \theta \in B_{12}. \end{cases}$$

Thus, when $\psi \in A_1$, $u_1(\theta) \geq 0$ for $\theta \in B_{11}$ and $u_1(\theta) < 0$ for $\theta \in B_{12}$. If $\theta \in B_1$ and $\psi \in A_2$, $((\theta + b)/2 - \psi) \in (\pi/3, \pi]$ and so $u_1(\theta) \geq 0$. Similarly, when $\theta \in B_2$, $\theta/2 \subset (2\pi/3, \pi)$, the sign of $u_2(\theta)$ is only dependent on that of $\sin(\theta/2 - \psi)$. If $\theta \in B_2$ and $\psi \in A_1 \cup A_2$, $(\theta/2 - \psi) \in (\pi/4, \pi)$. This implies that $u_2(\theta) > 0$ when $\psi \in A_1 \cup A_2$.

From Table A.1, it is sufficient to prove

$$\int_{\theta \in B_{12} \cup B_2} u(\theta) f_1(\theta|\psi)\, \mathrm{d}\theta > 0 \qquad \text{for } \psi \in A_1. \tag{A.5}$$

Now we examine two cases separately, when distribution of θ is unimodal and bimodal.

(i) If density $f_1(\theta|\psi)$ is either $f(\theta|\psi)$ or $\varepsilon f(\theta|\psi) + (1 - \varepsilon) f(\theta|\psi + \pi)$ with $\varepsilon \geq (1 + \zeta_{\min})^{-1}$, distribution of θ is unimodal with mode at ψ. It means that $f_1(\theta|\psi)$ is increasing in $\theta \in [\psi + \pi, 2\pi)$. Write $u(\theta) = v_1(\theta) + v_2(\theta)$, where for $i = 1, 2$, $v_i(\theta) = u_i(\theta)$, $\theta \in B_{12} \cup B_2$. Thus,

$$\frac{v_1(\theta)}{v_2(\theta)} = \begin{cases} -\infty, & \text{if } \theta \in B_{12}; \\ 0, & \text{if } \theta \in B_2. \end{cases}$$

Table A.1. Behaviour of functions $u_1(\theta)$ and $u_2(\theta)$

	A_1	A_2	$A_1 \cup A_2$
B_1	$u_1(\theta) \geq 0$ $u_1(\theta) < 0$	$u_1(\theta) \geq 0$	$u_2(\theta) = 0$
B_2	$u_1(\theta) = 0$	$u_1(\theta) = 0$	$u_2(\theta) > 0$

Therefore, both $v_1(\theta)/v_2(\theta)$ and $f_1(\theta|\psi)$ are increasing in $\theta \in B_{12} \cup B_2$. Let ϕ be a uniform distributed random variable with respect to Lebesgue measure such that $\Pr(\phi \in B_{12} \cup B_2) = 1$. As $\mathsf{E}^{\phi}\{v_2(\phi) f_1(\phi|\psi)\} > 0$, using a result of [2], Theorem 2.1, we have

$$\frac{\mathsf{E}^{\phi}\{v_1(\phi)\}}{\mathsf{E}^{\phi}\{v_2(\phi)\}} \leq \frac{\mathsf{E}^{\phi}\{v_1(\phi) f_1(\phi|\psi)\}}{\mathsf{E}^{\phi}\{v_2(\phi) f_1(\phi|\psi)\}}.$$

In order to prove (A.5), it remains to show

$$\mathsf{E}^{\phi}\big\{v_1(\phi) + v_2(\phi)\big\} > 0 \qquad \text{for all } \psi \in A_1. \tag{A.6}$$

The above holds since we have

$$(b - 2\psi)\mathsf{E}^{\phi}\big\{v_1(\phi) + v_2(\phi)\big\} = \int_{\theta \in B_{12} \cup B_2} \big\{\cos(\phi_0 - \psi) - \cos(\phi - \psi)\big\}\, \mathrm{d}\phi > 0.$$

This completes the proof when distribution of θ is unimodal.

(ii) Now we assume that θ has a mixture distribution with the probability density $f_1(\theta|\psi) = \varepsilon f(\theta|\psi) + (1 - \varepsilon) f(\theta|\psi + \pi)$ with $1/2 \leq \varepsilon < (1 + \zeta_{\min})^{-1}$, that is, the distribution of θ is bimodal. It has two modes ψ and $\psi + \pi$. Antimodes are $\psi + w$ and $\psi + 2\pi - w (= \beta)$, where $w = \cos^{-1}(\zeta^{-1}((1 - \varepsilon)/\varepsilon))$. According to condition (C2) , $w \in [\pi/2, \pi)$. We may note that $\beta \leq 2\pi$ since $\psi < \pi/2 \leq w$ as $\psi \in A_1$. This implies that $f_1(\theta|\psi)$ is increasing in $\theta \in [\beta, 2\pi)$. If $\beta \leq 2\pi + 2\psi - b$, (A.5) can be proved following the lines of the above case. When $\beta > 2\pi + 2\psi - b$, decompose the set $B_{12} \cup B_2 = (2\pi + 2\psi - b, 2\pi)$ as $C_1 \cup C_2 \cup C_3$, where

$$C_1 = (2\pi + 2\psi - b, \beta], \qquad C_2 = (\beta, 2\pi - 2w + b], \qquad C_3 = (2\pi - 2w + b, 2\pi).$$

Since the probability density $f_1(\theta|\psi)$ is symmetric about β, that is, $f_1(\theta|\psi) = f_1(2\beta - \theta|\psi)$, we have

$$\int_{\theta \in C_1} u(\theta) f_1(\theta|\psi)\, \mathrm{d}\theta = \int_{\theta \in C_2} u(2\beta - \theta) f_1(2\beta - \theta|\psi)\, \mathrm{d}\theta = \int_{\theta \in C_2} u(2\beta - \theta) f_1(\theta|\psi)\, \mathrm{d}\theta.$$

Define

$$u_3(\theta) = \begin{cases} u(\theta) + u(2\beta - \theta), & \text{if } \theta \in C_2; \\ u(\theta), & \text{if } \theta \in C_3. \end{cases}$$

It may be noted that (A.6) yields the following for $\psi \in A_1$:

$$\int_{\theta \in C_2 \cup C_3} u_3(\theta)\, \mathrm{d}\theta = \int_{\theta \in B_{12} \cup B_2} u(\theta)\, \mathrm{d}\theta \geq 0. \tag{A.7}$$

Consider the function

$$u(\theta) + u(2\beta - \theta) = 4 \sin\left(\frac{\theta - \theta_0}{2}\right) \cos\left(\frac{\theta + \theta_0}{2} - \beta\right) \sin(\beta - \psi).$$

When $\theta \in B_{12} \cup B_2$, $\sin((\theta - \theta_0)/2) > 0$. Note that $\sin(\beta - \psi) = -\sin(w) < 0$ as $w \in [\pi/2, \pi)$. The sign of $u(\theta) + u(2\beta - \theta)$ is the opposite of that of $\cos((\theta + \theta_0)/2 - \beta)$. There are three cases according to $\pi + b/2 \in C_i$, for $i = 1, 2, 3$. As in the case when distribution of θ is unimodal, we define two functions $v_1(\theta)$ and $v_2(\theta)$ for $\theta \in C_2 \cup C_3$ in all these three cases such that $v_1(\theta)/v_2(\theta)$ is increasing in $\theta \in C_2 \cup C_3$ and $v_2(\theta)$ is nonnegative for all $\theta \in C_2 \cup C_3$.

When $\pi + b/2 \in C_1$, choices are

$$v_1(\theta) = \begin{cases} 0, & \text{if } \theta \in C_2; \\ u(\theta), & \text{if } \theta \in C_3; \end{cases} \qquad v_2(\theta) = \begin{cases} u(\theta) + u(2\beta - \theta), & \text{if } \theta \in C_2; \\ 0, & \text{if } \theta \in C_3. \end{cases}$$

Now consider $\pi + b/2 \in C_2$. Decompose the interval $C_2 = C_{21} \cup C_{22}$, where

$$C_{21} = (\beta, \pi + b/2], \qquad C_{22} = (\pi + b/2, 2\pi - 2w + b].$$

In this case, we choose

$$v_1(\theta) = \begin{cases} u(\theta) + u(2\beta - \theta), & \text{if } \theta \in C_{21}; \\ 0, & \text{if } \theta \in C_{22}; \\ 0, & \text{if } \theta \in C_3; \end{cases}$$

$$v_2(\theta) = \begin{cases} 0, & \text{if } \theta \in C_{21}; \\ u(\theta) + u(2\beta - \theta), & \text{if } \theta \in C_{22}; \\ u(\theta), & \text{if } \theta \in C_3. \end{cases}$$

When $\pi + b/2 \in C_3$, decompose the interval $C_3 = C_{31} \cup C_{32}$, where

$$C_{31} = (2\pi - 2w + b, \pi + b/2], \qquad C_{32} = (\pi + b/2, 2\pi).$$

In this case, we define

$$v_1(\theta) = \begin{cases} u(\theta) + u(2\beta - \theta), & \text{if } \theta \in C_2; \\ 0, & \text{if } \theta \in C_{31}; \\ u(\theta), & \text{if } \theta \in C_{32}; \end{cases} \qquad v_2(\theta) = \begin{cases} 0, & \text{if } \theta \in C_2; \\ -u(\theta), & \text{if } \theta \in C_{31}; \\ 0, & \text{if } \theta \in C_{32}. \end{cases}$$

Since the density $f_1(\theta|\psi)$ is increasing in $\theta \in C_2 \cup C_3$, Theorem 2.1 of [2] completes the proof when distribution of θ is bimodal.

A.4. Proof of Lemma 4.1

See [22], Lemma 2, for the proof of (i). Note that the proof given in [22] utilizes the measure perseverance of the element $\mathbf{g} \in \mathcal{G}$ under the measure η, that is, $\eta(\mathbf{g}^{-1}(B)) \neq \eta(B)$ for all $B \in \mathfrak{B}(\mathfrak{Z})$, where $\mathfrak{B}(\mathfrak{Z})$ consists of Borel sets of the sample space $\mathfrak{Z}$. Later, Moors and van Houwelingen [23] relaxed this condition of measure perseverance.

For any $\mathbf{g} \in \mathcal{G}$, if δ is an $\mathcal{G}$-equivariant estimator, we have $\mathsf{L}(\bar{\mathbf{g}}(\boldsymbol{\nu}), \tilde{g}(\delta)) = \mathsf{L}(\boldsymbol{\nu}, \delta)$, that is,

$$\cos\big(h\bar{\mathbf{g}}(\boldsymbol{\nu}) - \tilde{g}(\delta)\big) = \cos\big(h(\boldsymbol{\nu}) - \delta\big), \tag{A.8}$$

for all $\boldsymbol{\nu} \in \Omega$ and $\delta \in \mathcal{A}$. Substituting $\delta = h(\boldsymbol{\nu})$, we obtain $\cos(h\bar{\mathbf{g}}(\boldsymbol{\nu}) - \tilde{g}h(\boldsymbol{\nu})) = 1$. Thus, we have $h\bar{\mathbf{g}}(\boldsymbol{\nu}) = \tilde{g}h(\boldsymbol{\nu}) \bmod (2\pi)$, for all $\nu \in \Omega$. This proves (ii). Exploiting this result, (A.8) reduces to

$$d\big(\tilde{g}h(\boldsymbol{\nu}), \tilde{g}(\delta)\big) = d\big(h(\boldsymbol{\nu}), \delta\big),$$

for all $\boldsymbol{\nu} \in \Omega$ and $\delta \in h(\Omega)$. This implies that $\tilde{g}(\cdot)$ is distance-preserving map on $\mathcal{A}$. Let ψ be the projection of $\tilde{g}(\phi)$ on $\tilde{g}(A)$. Since $\tilde{g}(\cdot)$ is injective, we have

$$d\big(\tilde{g}(\phi), \psi\big) = d(\phi, \phi_0) = d\big(\tilde{g}(\phi), \tilde{g}(\phi_0)\big).$$

From the uniqueness of projection, $\psi = \tilde{g}(\phi_0)$. This proves (iii).

A.5. Proof of Lemma 4.2

First, we show that for all $\mathbf{g} \in \mathcal{G}$ and $\boldsymbol{\nu} \in \Omega$, new estimand $h_{\mathbf{z}}(\boldsymbol{\nu})$ satisfies

$$h_{\mathbf{g}(\mathbf{z})}\bar{\mathbf{g}}(\boldsymbol{\nu}) = \tilde{g}h_{\mathbf{z}}(\boldsymbol{\nu}) \qquad \text{for all } \boldsymbol{\nu} \in \Omega. \tag{A.9}$$

Note that

$$\begin{aligned}\int_{\mathcal{G}} f\big(\mathbf{g}(\mathbf{z})|\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\,\mathrm{d}\lambda\big(\mathbf{g}^*\big) &= \int_{\mathcal{G}} f\big(\mathbf{z}|\bar{\mathbf{g}}^{-1}\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\,\mathrm{d}\lambda\big(\mathbf{g}^*\big) \qquad \big(\text{from Lemma 4.1(i)}\big)\\ &= \int_{\mathcal{G}} f\big(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu})\big)\,\mathrm{d}\lambda\big(\mathbf{g}^*\big) \qquad \big(\text{using transformation } \mathbf{g}^* \to \mathbf{g}\mathbf{g}^*\mathbf{g}^{-1}\big).\end{aligned}$$

This implies that for $\int_{\mathcal{G}} f(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*) = 0$, (A.9) follows from Lemma 4.1(ii). In the case of $\int_{\mathcal{G}} f(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*) > 0$, for any $\mathbf{g} \in \mathcal{G}$, we have

$$\begin{aligned}h_{\mathbf{g}(\mathbf{z})}\bar{\mathbf{g}}(\boldsymbol{\nu}) &= \operatorname{atan}\left(\frac{\int_{\mathcal{G}} \sin(\tilde{g}^*h\bar{\mathbf{g}}(\boldsymbol{\nu}))\tau(\mathbf{g}(\mathbf{z})|\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}{\int_{\mathcal{G}} \cos(\tilde{g}^*h\bar{\mathbf{g}}(\boldsymbol{\nu}))\tau(\mathbf{g}(\mathbf{z})|\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}\right)\\ &= \operatorname{atan}\left(\frac{\int_{\mathcal{G}} \sin(h\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}^{-1}\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}{\int_{\mathcal{G}} \cos(h\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}^{-1}\bar{\mathbf{g}}^*\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}\right)\\ &= \operatorname{atan}\left(\frac{\int_{\mathcal{G}} \sin(\tilde{g}\tilde{g}^*h(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}{\int_{\mathcal{G}} \cos(\tilde{g}\tilde{g}^*h(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}\right)\\ &= \tilde{g}\left(\operatorname{atan}\left(\frac{\int_{\mathcal{G}} \sin(\tilde{g}^*h(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}{\int_{\mathcal{G}} \cos(\tilde{g}^*h(\boldsymbol{\nu}))\tau(\mathbf{z}|\bar{\mathbf{g}}^*(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}^*)}\right)\right)\\ &= \tilde{g}h_{\mathbf{z}}(\boldsymbol{\nu})\end{aligned}$$

for all $\boldsymbol{\nu} \in \Omega$. The above equalities utilize Lemmas 4.1 and 2.3, the transformation $\mathbf{g}^* \to \mathbf{g}\mathbf{g}^*\mathbf{g}^{-1}$ and the circular property of $\tilde{g}$. Therefore, the surjection property of $\bar{\mathbf{g}}$ and (A.9) imply that

$$\tilde{g}h_{\mathbf{z}}(\Omega) = h_{\mathbf{g}(\mathbf{z})}(\Omega), \tag{A.10}$$

or equivalently, $\mathsf{cc}(\tilde{g}h_{\mathbf{z}}(\Omega)) = \mathsf{cc}(h_{\mathbf{g(z)}}(\Omega)) = \mathcal{A}_{\mathbf{g(z)}}$. Clearly, $\tilde{g}$-image of a closed convex set is again a closed convex set. Therefore, $\tilde{g}(\mathsf{cc}(h_{\mathbf{z}}(\Omega)))$ is also a closed convex and $\tilde{g}h_{\mathbf{z}}(\Omega) \subset \tilde{g}(\mathsf{cc}(h_{\mathbf{z}}(\Omega)))$. This implies that

$$\mathcal{A}_{\mathbf{g(z)}} = \mathsf{cc}\big(\tilde{g}h_{\mathbf{z}}(\Omega)\big) \subset \tilde{g}\big(\mathsf{cc}\big(h_{\mathbf{z}}(\Omega)\big)\big) = \tilde{g}(A_{\mathbf{z}})$$

as $\mathsf{cc}(\tilde{g}h_{\mathbf{z}}(\Omega))$ is the smallest convex set containing $\tilde{g}h_{\mathbf{z}}(\Omega)$. Next, we show that $\tilde{g}(A_{\mathbf{z}}) \subset \mathcal{A}_{\mathbf{g(z)}}$. As $\phi \in A_{\mathbf{z}} = \mathsf{conv}(h_{\mathbf{z}}(\Omega)) \cup \mathrm{bd}(h_{\mathbf{z}}(\Omega))$, there can be the following two cases:

(i) Suppose $\phi \in \mathsf{conv}(h_{\mathbf{z}}(\Omega))$. From Proposition 2.1, there exists $\phi_1, \phi_2 \in h_{\mathbf{z}}(\Omega)$ such that

$$\phi = \mathsf{atan}\left(\frac{w\sin\phi_1 + (1-w)\sin\phi_2}{w\cos\phi_1 + (1-w)\cos\phi_2}\right).$$

Operating $\tilde{g}$ on the both sides of the above equation and using the circular property, we get

$$\tilde{g}(\phi) = \mathsf{atan}\left(\frac{w\sin\tilde{g}(\phi_1) + (1-w)\sin\tilde{g}(\phi_2)}{w\cos\tilde{g}(\phi_1) + (1-w)\cos\tilde{g}(\phi_2)}\right).$$

Note that $\tilde{g}(\phi_1), \tilde{g}(\phi_2) \in \tilde{g}h_{\mathbf{z}}(\Omega)$. Using (A.10), both belong to $h_{\mathbf{g(z)}}(\Omega)$. Therefore, $\tilde{g}(\phi) \in \mathsf{conv}(h_{\mathbf{g(z)}}(\Omega)) \subset \mathcal{A}_{\mathbf{g(z)}}$.

(ii) Any $\phi \in \mathrm{bd}(h_{\mathbf{z}}(\Omega))$ is the limit point of a series of points $\{\phi_n\}$ with $\phi_n \in \mathsf{conv}(h_{\mathbf{z}}(\Omega))$. Since $\tilde{g}$ is circular so is continuous, $\tilde{g}(\phi)$ is the limit point of the series $\{\tilde{g}(\phi_n)\}$ in $\mathsf{conv}(\tilde{g}h_{\mathbf{z}}(\Omega)) = \mathsf{conv}(h_{\mathbf{g(z)}}(\Omega))$. Hence, $\tilde{g}(\phi) \in \mathrm{bd}(h_{\mathbf{g(z)}}(\Omega)) \subset \mathcal{A}_{\mathbf{g(z)}}$.

A.6. Proof of Theorem 4.1

Since risk of an equivariant estimator $\delta(\mathbf{Z})$ is constant on the orbits of $\boldsymbol{\nu}$ ([7], page 149), risk of the $\mathcal{G}$-equivariant estimator $\delta(\mathbf{Z})$ under the loss function L satisfies $\mathsf{R}(\boldsymbol{\nu}, \delta) = \int_{\mathcal{G}} \mathsf{R}(\bar{\mathbf{g}}(\boldsymbol{\nu}), \delta)\, \mathrm{d}\lambda(\mathbf{g})$. Using this, the risk of $\delta(\mathbf{Z})$ is given by

$$\begin{aligned}\mathsf{R}(\boldsymbol{\nu}, \delta) &= \int_{\mathcal{G}}\int_{\mathfrak{Z}} \big\{1 - \cos\big(\delta(\mathbf{z}) - h\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\big\} f\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\, \mathrm{d}\eta(\mathbf{z}))\, \mathrm{d}\lambda(\mathbf{g}) \\ &= \int_{\mathfrak{Z}}\int_{\mathcal{G}} \big\{1 - \cos\big(\delta(\mathbf{z}) - \tilde{g}h(\boldsymbol{\nu})\big)\big\} f\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\, \mathrm{d}\lambda(\mathbf{g})\, \mathrm{d}\eta(\mathbf{z}).\end{aligned}$$

In the above step, we utilize the interchange in order of integration and Lemma 4.1(ii). Note that $\delta_0(\mathbf{g}(\mathbf{z}))$ is the projection of $\delta(\mathbf{g}(\mathbf{z}))$ on $\mathcal{A}_{\mathbf{g(z)}}$, that is, $\delta_0(\mathbf{g}(\mathbf{z}))$ is the projection of $\tilde{g}(\delta(\mathbf{z}))$ on $\tilde{g}(\mathcal{A}_{\mathbf{z}})$ from invariance of $\delta(\mathbf{z})$ and Lemma 4.2. From Lemma 4.1(iii), $\delta_0(\mathbf{g}(\mathbf{z})) = \tilde{g}(\delta_0(\mathbf{z}))$, or equivalently, $\delta_0(\mathbf{z})$ is also $\mathcal{G}$-equivariant estimator. Therefore, the above risk expression is also valid for $\delta_0(\mathbf{z})$. The difference $\mathsf{R}(\boldsymbol{\nu}, \delta) - \mathsf{R}(\boldsymbol{\nu}, \delta_0)$ is given by

$$\begin{aligned}u &= \int_{\delta(\mathbf{z})\notin\mathcal{A}_{\mathbf{z}}}\int_{\mathcal{G}} \big\{\cos\big(\delta_0(\mathbf{z}) - \tilde{g}h(\boldsymbol{\nu})\big) - \cos\big(\delta(\mathbf{z}) - \tilde{g}h(\boldsymbol{\nu})\big)\big\} f\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\, \mathrm{d}\lambda(\mathbf{g})\, \mathrm{d}\eta(\mathbf{z}) \\ &= \int_{\delta(\mathbf{z})\notin\mathcal{A}_{\mathbf{z}}} \left[\mathsf{E}^{\mathbf{g}}\big\{\cos\big(\delta_0(\mathbf{z}) - \tilde{g}h(\boldsymbol{\nu})\big) - \cos\big(\delta(\mathbf{z}) - \tilde{g}h(\boldsymbol{\nu})\big)\big\} \int_{\mathcal{G}} f\big(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\big)\, \mathrm{d}\lambda(\mathbf{g})\right] \mathrm{d}\eta(\mathbf{z}),\end{aligned}$$

if $\int_{\mathcal{G}} f(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g}) > 0$, where the expectation is taken over $\mathbf{g}$ with respect to a probability measure $\tau(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu}))\,\mathrm{d}\lambda(\mathbf{g})$. Using the representation of $h_{\mathbf{z}}(\boldsymbol{\nu})$ given in (4.2) and denoting by $v_{\mathbf{z}}(\boldsymbol{\nu}) = [\{\mathsf{E}\sin(\tilde{g}h(\boldsymbol{\nu}))\}^2 + \{\mathsf{E}\sin(\tilde{g}h(\boldsymbol{\nu}))\}^2]^{1/2}$, we have

$$\sin h_{\mathbf{z}}(\boldsymbol{\nu}) = \mathsf{E}\sin\bigl(\tilde{g}h(\boldsymbol{\nu})\bigr)/v_{\mathbf{z}}(\boldsymbol{\nu}),$$
$$\cos h_{\mathbf{z}}(\boldsymbol{\nu}) = \mathsf{E}\cos\bigl(\tilde{g}h(\boldsymbol{\nu})\bigr)/v_{\mathbf{z}}(\boldsymbol{\nu}),$$

the risk difference is given by

$$u = \int_{\delta(\mathbf{z})\notin\mathcal{A}_{\mathbf{z}}} v_{\mathbf{z}}(\boldsymbol{\nu})\bigl\{\cos\bigl(\delta_0(\mathbf{z}) - h_{\mathbf{z}}(\boldsymbol{\nu})\bigr) - \cos\bigl(\delta(\mathbf{z}) - h_{\mathbf{z}}(\boldsymbol{\nu})\bigr)\bigr\} \int_{\mathcal{G}} f\bigl(\mathbf{z}|\bar{\mathbf{g}}(\boldsymbol{\nu})\bigr)\,\mathrm{d}\lambda(\mathbf{g})\,\mathrm{d}\eta(\mathbf{z}).$$

If $l(\mathcal{A}_{\mathbf{z}}) \le (2/3)\pi$, the above integrand is positive since $\cos(\delta(\mathbf{z}) - h_{\mathbf{z}}(\boldsymbol{\nu})) \le \cos(\delta_0(\mathbf{z}) - h_{\mathbf{z}}(\boldsymbol{\nu}))$ from Lemma 2.4. This completes the proof.

Acknowledgements

The authors are thankful to the two referees and an associate editor for their constructive comments and suggestions which have substantially improved the paper.

References

[1] Batschelet, E. (1981). *Circular Statistics in Biology*. London: Academic Press. MR0659065
[2] Bhattacharya, C.G. (1984). Two inequalities with an application. *Ann. Inst. Statist. Math.* **36** 129–134. MR0752011
[3] Buss, S.R. and Fillmore, J.P. (2001). Spherical averages and applications to spherical splines and interpolation. *ACM Trans. Graph.* **20** 95–126.
[4] Danzer, L., Grünbaum, B. and Klee, V. (1963). Helly's theorem and its relatives. In *Proc. Sympos. Pure Math., Vol. VII* 101–180. Providence, RI: Amer. Math. Soc. MR0157289
[5] Ducharme, G.R. and Milasevic, P. (1987). Spatial median and directional data. *Biometrika* **74** 212–215. MR0885936
[6] Eaton, M.L. (1989). *Group Invariance Applications in Statistics. Regional Conference Series in Probability and Statistics*. Hayward, CA: IMS.
[7] Ferguson, T.S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. New York: Academic Press. MR0215390
[8] Fisher, N.I. (1993). *Statistical Analysis of Circular Data*. Cambridge: Cambridge Univ. Press. MR1251957
[9] Gradshteyn, I.S. and Ryzhik, I.M. (1965). *Table of Integrals, Series and Products*. New York: Academic Press.
[10] He, X. and Simpson, D.G. (1992). Robust direction estimation. *Ann. Statist.* **20** 351–369. MR1150348
[11] Holmquist, B. (1991). Estimating and testing the common mean direction of several von Mises–Fisher populations with known concentrations. *Statistics* **22** 369–378. MR1108133
[12] Jammalamadaka, S.R. and SenGupta, A. (2001). *Topics in Circular Statistics* **5**. River Edge, NJ: World Scientific. MR1836122

[13] Jones, M.C. and Pewsey, A. (2005). A family of symmetric distributions on the circle. *J. Amer. Statist. Assoc.* **100** 1422–1428. MR2236452
[14] Kubokawa, T. (2004). Minimaxity in estimation of restricted parameters. *J. Japan Statist. Soc.* **34** 229–253. MR2116757
[15] Kubokawa, T. (2005). Estimation of bounded location and scale parameters. *J. Japan Statist. Soc.* **35** 221–249. MR2328426
[16] Kumar, S. and Sharma, D. (1992). An inadmissibility result for affine equivariant estimators. *Statist. Decisions* **10** 87–97. MR1165706
[17] Marchand, E. and Strawderman, W.E. (2004). Estimation in restricted parameter spaces: A review. In *A Festschrift for Herman Rubin*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **45** 21–44. Beachwood, OH: IMS. MR2126884
[18] Marchand, É. and Strawderman, W.E. (2012). A unified minimax result for restricted parameter spaces. *Bernoulli* **18** 635–643. MR2922464
[19] Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR1828667
[20] Mardia, K.V. and Sutton, T.W. (1978). A model for cylindrical variables with applications. *J. Roy. Statist. Soc. Ser. B* **40** 229–233.
[21] Moors, J.J.A. (1985). Estimation in truncated parameter space. Ph.D. dissertation, Tilburg Univ., The Netherlands.
[22] Moors, J.J.A. (1981). Inadmissibility of linearly invariant estimators in truncated parameter spaces. *J. Amer. Statist. Assoc.* **76** 910–915. MR0650903
[23] Moors, J.J.A. and van Houwelingen, J.C. (1993). Estimation of linear models with inequality restrictions. *Stat. Neerl.* **47** 185–198. MR1243854
[24] Neeman, T. and Chang, T. (2001). Rank score statistics for spherical data. *Contemp. Math.* **287** 241–254.
[25] Rueda, C., Fernández, M.A. and Peddada, S.D. (2009). Estimation of parameters subject to order restrictions on a circle with application to estimation of phase angles of cell cycle genes. *J. Amer. Statist. Assoc.* **104** 338–347. MR2504381
[26] Tsai, M.-T. (2009). Asymptotically efficient two-sample rank tests for modal directions on spheres. *J. Multivariate Anal.* **100** 445–458. MR2483430
[27] Udrişte, C. (1994). *Convex Functions and Optimization Methods on Riemannian Manifolds*. *Mathematics and Its Applications* **297**. Dordrecht: Kluwer Academic. MR1326607
[28] van Eeden, C. (2006). *Restricted Parameter Space Estimation Problems*. New York: Springer. MR2265239
[29] Watson, G.S. (1983). *Statistics on Spheres*. New York: Wiley. MR0709262

Received January 2015