Journal of Multivariate Analysis 178 (2020) 104600

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

A Central Limit Theorem for extrinsic antimeans and estimation of Veronese–Whitney means and antimeans on planar Kendall shape spaces

Yunfan Wang *, Vic Patrangenaru, Ruite Guo

Department of Statistics, Florida State University, United States of America

ARTICLE INFO

Article history:
Received 15 March 2019
Received in revised form 12 February 2020
Accepted 13 February 2020
Available online 3 March 2020

AMS 2020 subject classifications:
primary 62R30
62H35
62G05
secondary 62F12
62P10
62R99

Keywords:
Complex projective space
Extrinsic antimean
Kendall planar shape space
Nonparametric bootstrap
Random object
Statistics on manifolds
Veronese Whitney embedding

ABSTRACT

This article is concerned with random objects in the complex projective space $\mathbb{C}P^{k-2}$. It is shown that the Veronese–Whitney (VW) antimean, which is the extrinsic antimean of a random point on $\mathbb{C}P^{k-2}$ relative to the VW-embedding, is given by the point on $\mathbb{C}P^{k-2}$ represented by the eigenvector corresponding to the smallest eigenvalue of the expected mean of the VW-embedding of the random point, provided this eigenvalue is simple. We also derive a CLT for extrinsic sample antimeans, and an asymptotic χ^2-distribution of an appropriately studentized statistic, based on the extrinsic antimean, which in the particular case of a VW-embedding is then used to construct nonparametric bootstrap confidence regions for the VW-antimean planar Kendall shape. Simulations studies and an application to medical imaging are illustrating the proposed methodology.

1. Introduction

To date, Object Data Analysis (ODA) is the most inclusive type of data analysis, as far as sampling from distributions on metric spaces is concerned. Early examples of object spaces were spaces of directions [28], direct similarity shape spaces [14], axial spaces [2,9], and Stiefel manifolds [13]. In the infinite dimensional case, ODA leads to a nonlinear extension of functional data analysis [21, Chapter 11].

Fréchet [10] noticed that for higher complexity data, such as the shape of a random contour, numbers or vectors do not provide a meaningful representation. To investigate these kinds of data he introduced the notion of elements, nowadays called objects. Fréchet's visionary concepts, were nevertheless difficult to handle computationally during his lifetime. It took many decades, until such data became the bread and butter of modern Statistics, including Image Analysis. Nowadays, various types of shapes of configurations extracted from digital images are represented as points on projective shape spaces [19,23], on affine shape spaces [24,27], or on Kendall shape spaces [8,14]. To analyze the mean and variance

* Corresponding author.
E-mail address: yw15e@my.fsu.edu (Y. Wang).

https://doi.org/10.1016/j.jmva.2020.104600

of the random object X on a smooth object space $\mathscr{M}$ with a metric ρ, Fréchet defined what we call now the Fréchet function given by

$$\mathscr{F}(p) = E(\rho^2(p, X)) = \mathscr{F}(p) = \int \rho^2(p, x)Q(dx), \tag{1}$$

where $Q = P_X$ is the probability measure on $\mathscr{M}$, associated with X. If $(\mathscr{M}, \rho)$ is complete, the minimizers of the Fréchet function form the Fréchet mean set. In general, if $\rho = \rho_g$ is the geodesic distance associated with a Riemannian structure g on a manifold $\mathscr{M}$, there are no necessary and sufficient conditions for the existence of a unique minimizer of $\mathscr{F}$ in (1), [21, Chapter 4]. Therefore, with the possible exception of complete flat Riemannian manifolds or complete simply connected Riemannian manifolds of negative curvature, it is advisable to consider mainly the case when $\rho = \rho(j)$ is the "chord" distance on $\mathscr{M}$ induced by the Euclidean distance in $\mathbb{R}^N$ via an embedding $j : \mathscr{M} \to \mathbb{R}^N$, that is

$$\rho(p_1, p_2) = \|j(p_2) - j(p_1)\| = d(j(p_1), j(p_2)), \forall (p_1, p_2) \in \mathscr{M}^2, \tag{2}$$

where $\| \cdot \|$ is the Euclidean norm, and d is the corresponding Euclidean distance in $\mathbb{R}^N$. The Fréchet function becomes

$$\mathscr{F}(p) = \int_{\mathscr{M}} \|j(x) - j(p)\|^2 Q(dx). \tag{3}$$

The spread of data on a compact manifold differs from data spread on a Euclidean space. Indeed, given the finiteness of a probability measure, numerical data is spread thin outside a ball of sufficiently large radius; this is not the case with data on a compact manifold, that can be even uniformly spread. Test for uniformity of data on compact homogeneous spaces are well documented [1]. Therefore the mean of a distribution on a compact manifold is not interpretable as the "center" of data, as no such center exists, especially since the topology of such a manifold is nontrivial. Additional location parameters are therefore necessary to describe such data. In this paper it is assumed that $(\mathscr{M}, \rho)$ is a compact metric space, therefore the Fréchet function associated with the embedding j is bounded, and its extreme values are attained at two sets of points on $\mathscr{M}$: the lower bound – at the extrinsic mean set, and the upper bound – at the extrinsic antimean set [22]. In case the extrinsic mean set has one point only, that point is the extrinsic mean of X, and it is labeled $\mu_{j,E}(Q)$, or simply μ_E, when j and Q are known. If the extrinsic antimean set, has one point only, that point is the extrinsic antimean of X, and it is labeled $\alpha\mu_{j,E}(Q)$, or simply $\alpha\mu_E$, when j and Q are known.

Remark 1. The name of the new parameter, extrinsic antimean, is inspired from the case of a probability measure on a round sphere, when the extrinsic antimean is the antipodal point of the extrinsic mean. Recall that a d-dimensional sphere, is a set of points in the $d + 1$ dimensional Euclidean space, that are equidistant from a fixed point, known as center of the sphere. The d-dimensional sphere admits infinitely many Riemannian structures. The round sphere is the d-dimensional sphere with the Riemannian structure induced by the scalar product of the ambient $d+1$ Euclidean space. This is an extreme case, given that the round sphere of dimension $d, d \geq 2$ is the only simply connected compact d dimensional Riemannian manifold that admits a group of isometries of dimension $\frac{d(d+1)}{2}$, the highest possible degree of symmetry [17].

Also, given $X_1, \ldots, X_n$ i.i.d.r.o.'s from Q, their extrinsic sample mean (set) is the extrinsic mean (set), of the empirical distribution $\hat{Q}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$; their extrinsic sample antimean (set) is the extrinsic antimean (set) of $\hat{Q}_n$ [21].

Recall that a vector $y \in \mathbb{R}^N$ for which there is a unique point $p \in \mathscr{M}$, such that

$$d(y, j(\mathscr{M})) = \inf_{x\in\mathscr{M}} \|y - j(x)\| = d(y, j(p)), \tag{4}$$

is said to be j-nonfocal. The point $j(p)$ in (4) is the projection of y on $j(\mathscr{M})$. Thus if P_j, is the projection defined on the set of j-nonfocal points, then $P_j(y) = j(p)$. A vector $y \in \mathbb{R}^N$ for which there is a unique point $p \in \mathscr{M}$, such that

$$\sup_{x\in\mathscr{M}} \|y - j(x)\| = d(y, j(p)), \tag{5}$$

is said to be αj-nonfocal (see [25]); and the point $j(p)$ in (5), is the farthest projection of y on $j(\mathscr{M})$. Thus, if $P_{F,j}$ is the farthest projection defined on the set of αj-nonfocal points, and (p, y) are as in (5), then $P_{F,j}(y) = j(p)$.

The extrinsic mean of X, exists if and only if the mean vector μ of $j(X)$ is a j-nonfocal point, and in this case the extrinsic mean is $\mu_E = j^{-1}(P_j(\mu))$; the extrinsic antimean of X, exists if and only if μ is an αj-nonfocal point, and in this case the extrinsic antimean is $\alpha\mu_E = j^{-1}(P_{F,j}(\mu))$.

Remark 2. Note that on a compact manifold, the extrinsic mean, when it exists, lacks the interpretability of center of the data, as such "center" of data exists on a non-Euclidean compact space, when the data is not confined to a small neighborhood, as no such center exists, especially when the topology of that manifold is nontrivial. For example any distribution with a nonzero probability density w.r.t. the uniform measure on the complex projective plane $\mathbb{C}P^2$ has no center, due to the nontrivial homology group $H^2(\mathbb{C}P^2) = \mathbb{Z}$. Indeed this is an algebraic topology obstruction insuring that the complement of a point on $\mathbb{C}P^2$, does not have the topology of a disk [26]; therefore one cannot talk about its center.

Remark 3. In general there are other helpful features of the extrinsic antimean as a useful population parameter, including that (i) it may differentiate between two distributions having the same extrinsic mean, (ii) there are distributions having an extrinsic antimean, that do not have an extrinsic mean, and, (iii) there are families of distributions with a sticky mean, and varying non sticky antimeans. Examples of distributions enjoying any of the properties (i)–(iii), can be easily imagined, since the existence of the extrinsic mean, or antimean is determined by the relative position of the mean of the distribution related to the embedding of the object space in the ambient space. In the 1D case, such examples are given graphically in [25].

This paper is specialized in random objects on the complex projective space $\mathbb{C}P^{k-2}$, which is a compact symmetric space of dimension $2k-4$, that serves as manifold representation of the space of direct similarity shapes of k-ads (k labeled points that are not all identical) in the Euclidean plane [14]. Recall that the complex projective space $\mathbb{C}P^q$ is the set of all one dimensional linear subspaces of $\mathbb{C}^{q+1}$ and can also be represented as the set orbits of group action of S^1 on the $2q+1$ dimensional sphere $S^{2k+1} = \{z \in \mathbb{C}^{k+1}, \|z\| = 1\}$ given by scalar multiplications with complex numbers of modulus 1. The complex projective space has a convenient embedding in the space of $(k-1)\times(k-1)$ self adjoint complex matrices, that was introduced in shape analysis by Kent [15], and is often known as the Veronese–Whitney embedding. The VW mean sets, and VW means and the VW antimean sets and VW antimeans are extrinsic mean sets, and extrinsic means, respectively extrinsic antimean sets, and extrinsic antimeans of random points on $\mathbb{C}P^{k-2}$, relative to the VW embedding. In Section 2 we give explicit formulas for VW means and VW antimeans, and their sample counterparts. That result is used to give two examples of sample VW means and sample VW antimeans computations on planar Kendall shape spaces, two synthetic, and one data driven. In anticipation of those results, note that the action of S^1 on S^{2k+1} leaves the volume measure invariant, therefore if f is density function on S^{2q+1}, relative to the volume measure on the sphere, with the property that $f(\zeta z) = f(z), \forall \zeta \in S^1, \forall z \in S^{2q+1}$, one may define a density function $\tilde{f}$ on $\mathbb{C}P^q$ with respect to the corresponding quotient volume measure ν on $\mathbb{C}P^q$, via $\tilde{f}([z]) = f(z)$. Let H_q be the set of $q\times q$ Hermitian (selfadjoint) complex matrices. Recall that a complex Bingham distributed random object Y on $\mathbb{C}P^q$ has the probability density, with respect to volume measure on $\mathbb{C}P^q$ of the form $f_A, A \in H_q$, where $f_A([z]) \propto \exp(z^*Az)$, where z^* is the complex conjugate of the transpose of z [16]. One can show that if a random object X has the complex Bingham distribution f_A, then the VW mean set, respectively the VW mean set of X are the sets $\{[z], z \in V_{\lambda_{max}}\}$, respectively $\{[z], z \in V_{\lambda_{min}}\}$, where V_λ is the eigenspace of $\mathbb{C}^{q+1}$ corresponding to the eigenvalue λ of A. Therefore, if the eigenvalue λ_{max} has multiplicity two or higher, while λ_{min} is a simple eigenvalue, the random object X does not have a VW mean, and has a VW antimean.

In Section 3 we derive confidence regions for VW antimeans. Here, the first part is dedicated to a central limit theorem for sample extrinsic antimeans, and to a key asymptotic result for the norm of the studentized version of the tangential component of the sample extrinsic antimean (Theorem 2). In Section 3.2 we introduce the concept of VW anticovariance matrix, which is necessary in Theorem 3, to estimate the VW antimean based on a large sample of observations, which we apply to the midface anatomic landmark configurations in Section 2. For an improved coverage error in the VW mean estimation and of the VW antimean estimation, in Corollary 2, we give a nonparametric pivotal bootstrap based confidence regions for the planar VW-antimean, with a computational example for the first simulated shape data in Section 2. An interesting question is finding a relationship between simultaneous confidence intervals and the confidence regions for VW antimeans given in Corollaries 1 and 2. Additional nonparametric antimean inference topics that parallel nonparametric analysis of extrinsic means, are discussed in upcoming papers [18,20].

2. VW antimean

Direct similarity shapes of k-ads (set of k labeled points at least two of which are distinct) in the Euclidean space, were introduced by Kendall [14], who showed that in the 2D case, these shapes can be represented as points on a complex projective space $\mathbb{C}P^{k-2}$, which is the set of equivalence classes $[z]$ of vectors $z \in \mathbb{C}^{k-1}\backslash\{0_{k-1}\}$, with respect to the equivalence $z \simeq \lambda z, \lambda \in \mathbb{C}\backslash\{0\}$. A standard shape analysis method, initiated by Kent [15], consists in using the now so called Veronese–Whitney (VW) embedding of $\mathbb{C}P^{k-2}$ in the space of $(k-1)\times(k-1)$ self adjoint complex matrices, to represent shape data in a Euclidean space. This VW embedding $j_k\colon \mathbb{C}P^{k-2} \to S(k-1,\mathbb{C})$, where $S(k-1,\mathbb{C})$ is the space of $(k-1)\times(k-1)$ Hermitian matrices with the Euclidean distance (see proof of Theorem 1), is given by

$$j_k([z]) = zz^*,\ \ z^*z = 1, \tag{6}$$

where the $*$-operator, means taking the transpose followed by complex conjugation. This embedding is a $SU(k-1)$ equivariant embedding, where $SU(k-1)$ is the special unitary group $(k-1)\times(k-1)$ matrices of determinant 1, since $j_k([Az]) = Aj_k([z])A^*, \forall A \in SU(k-1)$. The corresponding extrinsic mean (set), respectively extrinsic antimean (set) of a random shape X on $\mathbb{C}P^{k-2}$ is called the VW mean (set), respectively VW antimean (set). The VW mean, respectively VW antimean, when they exist, will be labeled as μ_{VW}, respectively $\alpha\mu_{VW}$, and, given independent identically distributed random objects (i.i.d.r.o.'s) $X_1, \ldots, X_n$ from a distribution Q on $\mathbb{C}P^{k-2}$, their sample counterparts, when they exist, will be denoted by $\bar{X}_{VW}$, respectively by $a\bar{X}_{VW}$. A probability distribution Q on $\mathbb{C}P^{k-2}$ is said to be VW nonfocal, respectively αVW nonfocal, if Q is j_k-nonfocal, respectively αj_k-nonfocal. Necessary and sufficient conditions for a probability distribution to be VW-nonfocal are given in ([21, p.174]). The following result provides necessary and sufficient conditions for αVW nonfocality. A related result (for real projective spaces) can be found in [22].

Theorem 1. *Let $Q = P_{[Z]}, Z^*Z = 1$, be a probability distribution on $\mathbb{C}P^{k-2}$ and let $\{x_i = [z_i], \|z_i\| = 1, i \in \{1, \ldots, n\}\}$ be a random sample from Q.*
(i) Q is αVW nonfocal iff λ_1, the smallest eigenvalue of $E[ZZ^*]$ is simple, and in this case $\alpha\mu_{VW} = [\nu]$, where ν is an eigenvector of $E[ZZ^*]$ corresponding to λ_1, with $\| \nu \| = 1$.
(ii) If the empirical $\hat{Q}_n = \sum_{i=1}^n \frac{1}{n}\delta_{x_i}$ is αVW nonfocal, the sample VW antimean $a\bar{X}_{VW} = [m]$, where m is an eigenvector of norm 1 of $J = \frac{1}{n}\sum_{i=1}^n z_i z_i^*$, $\|z_i\| = 1, i \in \{1, \ldots, n\}$, corresponding to the smallest eigenvalue of J, provided this eigenvalue has multiplicity one.

Proof. (i) The squared Euclidean distance induced from $\mathbb{C}^{(k-1)^2}$ on the space $S(k-1, \mathbb{C})$ of Hermitian matrices is $d^2(A, B) = Tr((A - B)(A - B)^*) = Tr((A - B)^2)$. A random object $X = [Z]$ on $\mathbb{C}P^{k-2}$, with $Z^*Z = 1$, has the associated Fréchet function

$$\mathscr{F}([z]) = E(Tr(ZZ^* - zz^*)^2),\ z^*z = 1. \tag{7}$$

The matrix $A = E(ZZ^*)$ is positive semidefinite, having the eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_{k-1}$, and can be represented as $A = B\Lambda B^*$, where $B \in SU(k-1)$ and $\Lambda = \text{Diagonal}(\lambda_1, \lambda_2, \ldots, \lambda_{k-1})$. From (7), we get $\mathscr{F}([z]) = Tr(A^2) + 1 - 2Tr(z^*Az)$, thus $\mathscr{F}$ is maximized iff $[z] \to Tr(z^*Az)$ is minimized, or $[v] \to Tr(v^*\Lambda v)$ is minimized, where $v = Bz$. Note that $v^*v = 1$, and if $v^T = (v_1, \ldots, v_{k-1})$, then $Tr(v^*\Lambda v) = \sum_{a=1}^{k-1} \lambda_a |v_a|^2 \ge \lambda_1 = \mathscr{F}([u_1])$, where u_1 is an eigenvector of A corresponding to the eigenvalue λ_1. If λ_1 has multiplicity two or higher, $[v_1]$ and $[v_2]$ yield distinct points of minimum of $\mathscr{F}$, thus Q is not α VW nonfocal, proving (i), by contradiction. Part (ii) follows, by taking the empirical distribution, with the expectation A from Part (i) being given by $J = E(\hat{Q}_n) = n^{-1}\sum_{r=1}^n z_r z_r^*$.

In the supplemental material we ran two simulations, and, using the VW embedding of a complex projective space (manifold representation of a planar Kendall shape space), we compared the behavior of VW means and VW antimeans for those data sets of landmark configurations.

2.1. Application

We are interested to determine how concentrated is the bootstrap distribution of the sample VW antimeans around the sample VW antimean, in the case of shapes of landmark configurations extracted from medical imaging outputs. Our data consists of shapes for a group of eighth midface labeled anatomical landmarks from X-rays of skulls of eight year old and fourteen year-old North American children (36 boys and 26 girls), known as the University School Study data (see [7, p. 400–405]). For immediate access of the data see http://life.bio.sunysb.edu/morph/data/Book-UnivSch.txt). Each child's skull was imaged twice, at age 8 and next at age 14. The data set, collected to study anatomical changes during children growth, represents coordinates of eight craniofacial landmarks, whose names and position on the skull are given in [7, p. 400–405]. In [6] only part of this data set (boys only) was used. Here, we use only the University School Study data, recorded at age 14, for both sexes. For graphic displays of planar Kendall shape data, a useful tool for "removing location" of a k-ad, is the multiplication by a Helmert sub-matrix H, consisting in the last $(k-1) \times k$ rows of a full Helmert matrix. The full Helmert matrix HF, commonly used in Statistics, is a square $k \times k$ orthogonal matrix with its first row equal to $1/\sqrt{k}\mathbf{1}_k^T$, having the remaining rows orthogonal to the first row, with an increasing number of nonzero entries, as in (8). We drop the first row of HF so that the resulting matrix H does not depend on the original location of the configuration [8]. The jth row of the Helmert sub-matrix H is given by

$$(h_j, \ldots, h_j, -jh_j, 0, \ldots, 0),\ h_j = \{j(j+1)\}^{-1/2}. \tag{8}$$

One way of graphically representing planar k-ad shape data, or sample VW means or sample VW antimeans, is via a Helmert registration, where for each shape, regarded as equivalence class of a labeled landmark configuration, one selects a unique representative, by the same rule, equally applying to all shape observations, as follows: apply to each configuration of k landmarks, as point in $\mathbb{C}^k$, a Helmert sub-matrix in (8), and obtain a vector in $\mathbb{C}^{k-1}$ of norm 1, whose $k-1$ components are affixes of points in the Euclidean plane, in a representation that "removes location and scaling". In our application, the registered coordinates are displayed in Fig. 1. The shape variable is valued in a Kendall space of planar octads, $\mathbb{C}P^6$ (real dimension $= 12$), and is displayed in Fig. 2.

In Fig. 3 is displayed an icon of the sample VW mean of the Helmertized data, in a spherical representation (the corresponding vector $z \in \mathbb{C}^7$ is of norm one). Like with the simulated data in the Appendix, one may notice that the VW mean icon, has a fairly close shape to the shapes of the Helmertized cranial landmark data configurations.

Note that the sample size for the boys data only or girls data only is too small to be reliable in the estimation of the VW-mean and VW-antimean shape based on a large sample result. Therefore we computed the nonparametric bootstrap distribution of the sample VW means shapes in MATLAB, that we ran for 500 random re-samples, for the combined data. An icon of the Helmertized spherical representation of the bootstrap distribution of the sample VW means is displayed in Fig. 4. Note that the bootstrap distribution of the sample VW means is very concentrated around the sample VW mean in Fig. 3, as theoretically predicted.

As for the sample VW antimean, its icon is shown in Fig. 5. Since the sample VW antimean is on average far from the shape data, as expected, the relative location of the landmarks in the sample VW antimean icon looks quite different from those in the configurations of the original anatomical landmark data.

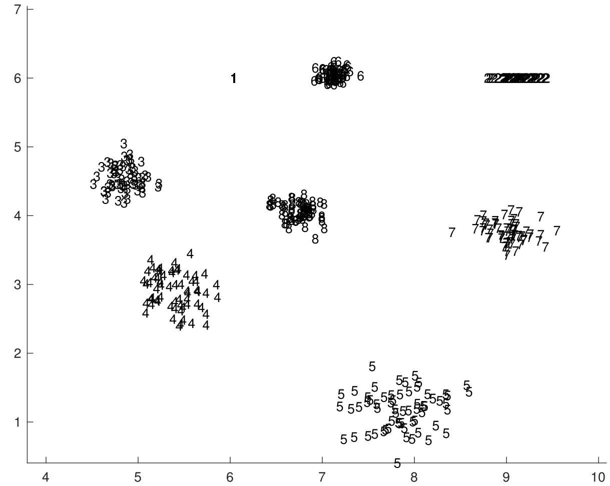

Fig. 1. Coordinates of eight mid-sagittal landmark configurations on midface children cranial data recorded from X-rays (University School Study data).

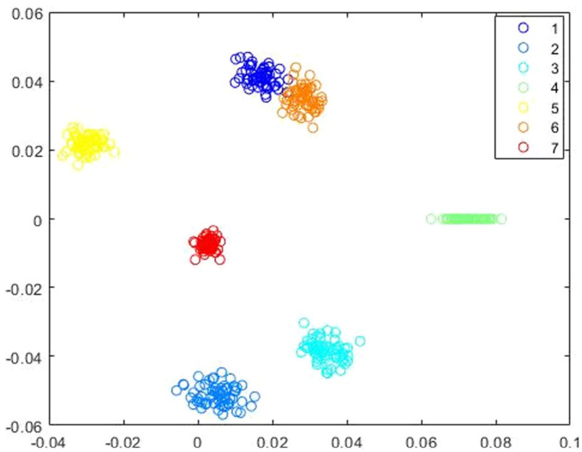

Fig. 2. Coordinates of landmark configurations in Fig. 1 transformed via the Helmert sub-matrix H.

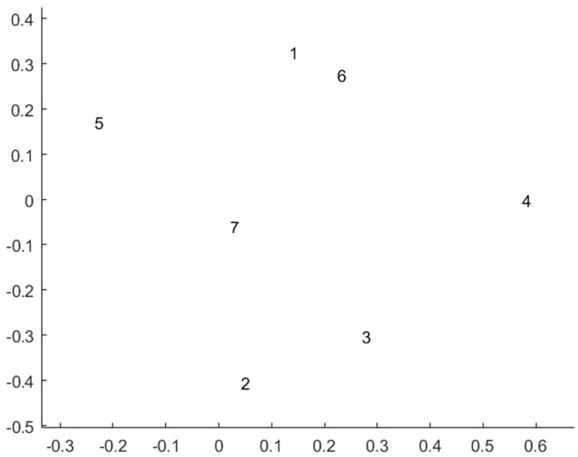

Fig. 3. Icon of the sample VW mean shape of the midface cranial landmark configuration data in Fig. 1 after applying the Helmert sub-matrix transformation H.

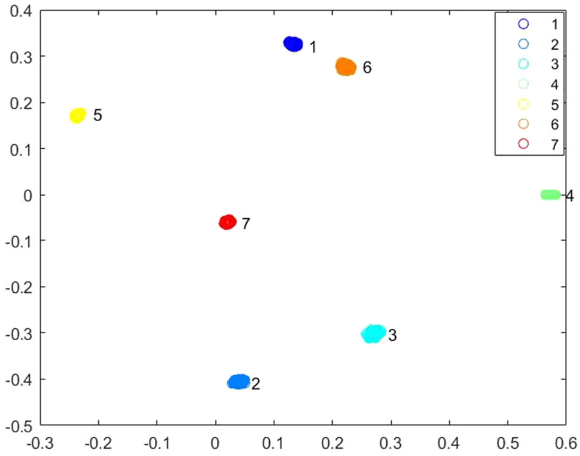

Fig. 4. Icons of 500 nonparametric bootstrap sample VW means for the midface cranial landmark configuration data in Fig. 1 after applying the Helmert sub-matrix transformation.

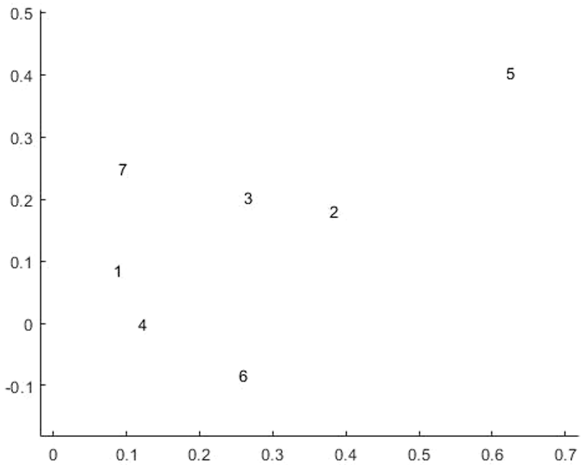

Fig. 5. Icon of sample VW antimean shape of midface cranial landmark configuration data in Fig. 1, after applying the Helmert sub-matrix transformation.

We computed the nonparametric bootstrap distribution using MATLAB, that we ran again for 500 random re-samples. A spherical representation of the bootstrap distribution of the sample VW antimeans in Helmetrized coordinates is displayed in Fig. 6. Here again, the icons of configurations for bootstrap distribution of the sample VW antimeans have a similar look with the one of the VW sample antimean, however it is not as concentrated around the registered icon of the sample VW antimean, as in the case of the bootstrap distribution of the VW means, partially due to computational rounding errors for eigenvectors associated with the smallest eigenvalue of J^*. It turned out that the smallest and the second smallest eigenvalue are clustered near one another, a fact that resulted in the eigenvectors being more sensitive to re-sampling with repetition. That is why the bootstrap VW antimeans are not as concentrated as the bootstrap VW means. The standard affine embedding: $\mathbb{C}^{k-2} \to \mathbb{C}P^{k-2}$ is $(z^1, \dots, z^{k-2}) \to [z^1 : \dots : z^{k-2} : 1]$, leads to the notion of affine coordinates of a projective point

$$p = [z^1 : \dots : z^{k-1}],\ z^{k-1} \neq 0 \tag{9}$$

to be defined as

$$(w^1, w^2, \dots, w^{k-2}) = (\frac{z^1}{z^{k-1}}, \dots, \frac{z^{k-2}}{z^{k-1}}). \tag{10}$$

Using simultaneous complex confidence intervals [6] for the affine coordinates of the VW antimean, we obtain the following results: w^1 : [-0.1804 − 0.1808i 0.0549 + 0.1365i], w^2 : [0.4913 − 0.2301i 0.6136 − 0.0747i], w^3 : [0.4455 + 0.0385i 0.5885 + 0.2288i], w^4: [0.1344 − 0.1923i 0.2346 − 0.0748i], w^5: [0.2376 − 0.4823i 0.5682 − 0.1533i], w^6: [-0.2752 + 0.2558i 0.1936 + 0.8011i].

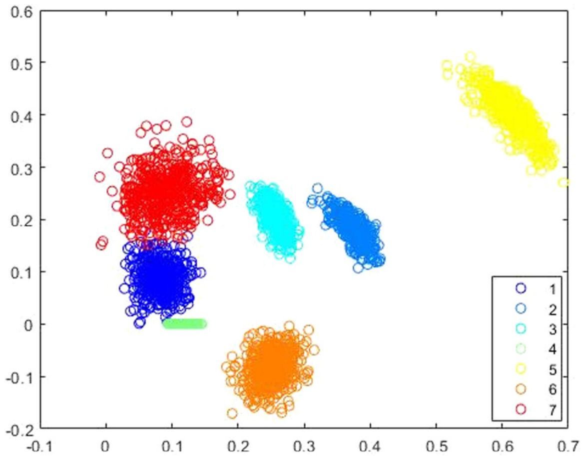

Fig. 6. Icons of 500 nonparametric bootstrap sample VW antimeans for the midface cranial landmark configuration data in Fig. 1 after applying the Helmert sub-matrix transformation.

3. Asymptotic distribution of extrinsic sample antimeans

In this section we will discuss the asymptotic distribution of sample antimeans in axial data analysis and in planar shape analysis, after a review of a Central Limit Theorem for extrinsic sample antimeans. As a notational matter, in general, if $k : \mathscr{P} \to \mathscr{Q}$ is a differentiable function between two manifolds, $d_x k$ is the differential of the function k evaluated at the point $x \in \mathscr{P}$.

3.1. Central limit theorem for extrinsic sample antimeans

In preparation, we are using the large sample distribution for extrinsic sample antimeans given in [22].

Assume j is an embedding of a d-dimensional manifold $\mathscr{M}$ such that $j(\mathscr{M})$ is closed in $\mathbb{R}^N$, and $Q = P_X$ is a αj-nonfocal probability measure on $\mathscr{M}$ such that $j(Q)$ has finite moments of order 2. Let μ and Σ be the mean and covariance matrix of $j(Q)$ regarded as a probability measure on $\mathbb{R}^N$. Let $\mathscr{F}$ be the set of αj-focal points of $j(\mathscr{M})$, and let $P_{F,j} : \mathscr{F}^c \to j(\mathscr{M})$ be the farthest projection on $j(\mathscr{M})$. $P_{F,j}$ is differentiable at μ and has the differentiability class of $j(\mathscr{M})$ around any αj nonfocal point. In order to evaluate the differential $d_\mu P_{F,j}$ we consider a special orthonormal frame field that will ease the computations.

A local frame field $p \to (e_1(p), \ldots, e_k(p))$, defined on an open neighborhood $U \subseteq \mathbb{R}^N$ is adapted to the embedding j if it is an orthonormal frame field and $\forall x \in j^{-1}(U)$, $e_r(j(x)) = d_x j(f_r(x))$, $r \in \{1, \ldots, d\}$, where $(f_1, \ldots, f_d)$ is a local frame field on $\mathscr{M}$, and $f_r(x)$ is the value of the local vector field f_r at x.

Let $e_1, \ldots, e_N$ be the canonical basis of $\mathbb{R}^N$ and assume $(e_1(p), \ldots, e_N(p))$ is an adapted frame field around $P_{F,j}(\mu) = j(\alpha\mu_E)$. Then $d_\mu P_{F,j}(e_b) \in T_{P_{F,j}(\mu)} j(\mathscr{M})$ is a linear combination of $e_1(P_{F,j}(\mu)), \ldots, e_d(P_{F,j}(\mu))$:

$$d_\mu P_{F,j}(e_b) = \sum_{a=1}^{d} (d_\mu P_{F,j}(e_b)) \cdot e_a(P_{F,j}(\mu)) e_a(P_{F,j}(\mu)) \tag{11}$$

where $d_\mu P_{F,j}$ is the differential of $P_{F,j}$ at μ. By the delta method, $n^{1/2}(P_{F,j}(\overline{j(X)}) - P_{F,j}(\mu))$ converges weakly to $N_N(0_N, \alpha\Sigma_\mu)$, where $\overline{j(X)} = \frac{1}{n}\sum_{i=1}^{n} j(X_i)$ and

$$\alpha\Sigma_\mu = [\sum_{a=1}^{d} d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu)) e_a(P_{F,j}(\mu))]_{b\in\{1,\ldots,N\}} \times \Sigma[\sum_{a=1}^{d} d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu)) e_a(P_{F,j}(\mu))]^T_{b\in\{1,\ldots,N\}}. \tag{12}$$

Here Σ is the covariance matrix of $j(X_1)$ w.r.t the canonical basis $e_1, \ldots, e_N$.

The asymptotic distribution $N_N(0_N, \alpha\Sigma_\mu)$ is degenerate and the support of this distribution is on $T_{P_{F,j}} j(\mathscr{M})$, since the range of $d_\mu P_{F,j}$ is $T_{P_{F,j(\mu)}} j(\mathscr{M})$. Note that $d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu)) = 0$ for $a \in \{d+1, \ldots, N\}$.

The tangential component $tan(v)$ of $v \in \mathbb{R}^N$, w.r.t the basis $e_a(P_{F,j}(\mu)) \in T_{P_{F,j(\mu)}} j(\mathscr{M})$, $a \in \{1, \ldots, d\}$ is given by

$$tan(v) = [e_1(P_{F,j}(\mu))^T v, \ldots, e_d(P_{F,j}(\mu))^T v]^T. \tag{13}$$

Then, the random vector $(d_{\alpha\mu_E}j)^{-1}(tan(P_{F,j}(\overline{(j(X))}) - P_{F,j}(\mu))) = \sum_{a=1}^{d} \overline{X}_j^a f_a$ has the following covariance matrix w.r.t the basis $f_1(\alpha\mu_E), \ldots, f_d(\alpha\mu_E)$:

$$\begin{aligned}\alpha\Sigma_{j,E} &= e_a(P_{F,j}(\mu))^T \alpha\Sigma_\mu e_b(P_{F,j}(\mu))_{1\le a,b\le d} \\ &= [d_\mu P_{F,j}(e_b)\cdot e_a(P_{F,j}(\mu))]_{a\in\{1,\ldots,d\}}\ \Sigma\ [d_\mu P_{F,j}(e_b)\cdot e_a(P_{F,j}(\mu))]^T_{a\in\{1,\ldots,d\}},\end{aligned} \tag{14}$$

which is the anticovariance matrix of the random object X. Similarly, given i.i.d.r.o.'s $X_1, \ldots, X_n$ from Q, we define the sample anticovariance matrix $aS_{j,E,n}$ as the anticovariance matrix associated with the empirical distribution $\hat{Q}_n$.

If, in addition, rank $\alpha\Sigma_\mu = d$, then $\alpha\Sigma_{j,E}$ is invertible and if we define the j-standardized antimean vector

$$\overline{Z}_{j,n} =: n^{\frac{1}{2}} \alpha\Sigma_{j,E}{}^{-\frac{1}{2}} (\bar{X}_j^1, \ldots, \bar{X}_j^d)^T, \tag{15}$$

using basic large sample theory results, including a generalized Slutsky's lemma ([21, p.65]), one has:

Theorem 2. *Assume $\{X_r\}_{r\in\{1,\ldots,n\}}$ is a random sample from the αj-nonfocal distribution Q, and let $\mu = E(j(X_1))$. Let $(e_1(p), e_2(p), \ldots, e_N(p))$ be an orthonormal frame field adapted to j. Then* (i) *the tangential component at the extrinsic antimean $\alpha\mu_E$ of $d_{\alpha\mu_E}j^{-1}tan_{P_{F,j}(\mu)}(P_{F,j}(\overline{(j(X))}) - P_{F,j}(\mu))$ has asymptotically a multivariate normal distribution in the tangent space to M at $\alpha\mu_E$ with mean 0_d and covariance matrix $n^{-1}\alpha\Sigma_{j,E}$,* (ii) *if $\alpha\Sigma_{j,E}$ is nonsingular, the j-standardized antimean vector $\overline{Z}_{j,n} = \alpha\Sigma_{j,E}^{-\frac{1}{2}} tan_{P_{F,j}(\mu)}(P_{F,j}(\overline{(j(X))}) - P_{F,j}(\mu))$ converges weakly to a random vector with a $N_d(0_d, I_d)$ distribution, and* (iii) *under the assumptions of* (ii)

$$\|(aS_{j,E,n})^{-\frac{1}{2}} tan_{P_{F,j}(\mu)}(P_{F,j}(\overline{(j(X))}) - P_{F,j}(\mu))\|^2 \to_d \chi^2_d. \tag{16}$$

Remark 4. In Theorem 2, no assumption on the finiteness of certain moments of the distribution Q is needed, since the manifold $\mathscr{M}$ is compact, thus all moments of a distribution with support on $j(\mathscr{M})$ are finite.

3.2. VW anticovariance matrices for data in $\mathbb{C}P^{k-2}$

A complex vector $z = (z^1, z^2, \ldots, z^{k-1})$ of norm 1 corresponding to a given configuration of k landmarks, with the last landmark removed, after centering this configuration, with the identification described in [5], can be displayed in the Euclidean plane (complex line) with the superscripts as labels.

Lemma 1. *Assume $x_r = [z_r]$, $z_r^* z_r = 1, r \in \{1, \ldots, n\}$ is a random sample from a α VW-nonfocal probability measure Q with a non-degenerate VW anticovariance matrix on $\mathbb{C}P^{k-2}$. If $\hat{\lambda}_a, a \in \{2, \ldots, k-1\}$ are eigenvalues of J in Theorem 1, in their increasing order and $m_a,\ a \in \{2, \ldots, k-1\}$ are corresponding linearly independent unit eigenvectors. Then the sample VW-anticovariance matrix $aS_{VW,n}$, as a complex matrix, has the entries*

$$(aS_{VW,n})_{ab} = n^{-1}(\hat{\lambda}_a - \hat{\lambda}_1)^{-1}(\hat{\lambda}_b - \hat{\lambda}_1)^{-1} \sum_{r=1}^{n} (m_a \cdot z_r)(m_b \cdot z_r)^* |m_1 \cdot z_r|^2, a, b \in \{2, \ldots, k-1\}. \tag{17}$$

Proof. The proof is based on the equivariance of the VW embedding j_k: $\mathbb{C}P^{k-2} \to S(k-1, \mathbb{C})$, with respect to actions of the special unitary group $SU(k-1)$, of non-negative semi-definite self-adjoint complex matrices [5]. First we need to assume that $J := \frac{1}{n}\sum_{r=1}^{n} z_r z_r^*$ is a diagonal matrix, the smallest eigenvalue corresponding complex eigenvector of norm 1 of J is a simple root of the characteristic polynomial over $\mathbb{C}$, with $m_1 = e_1$. The tangent space $T_{[m_1]}j_k(\mathbb{C}P^{k-2})$ has an orthobasis $m'_a = ie_a,\ a \in \{2, \ldots, k-1\}$, where $m_a = e_a$ are eigenvector corresponding to the largest eigenvalue. Here we define a path $\eta_z(t) = [cos\ tm_1 + sin\ tz]$, where z is orthogonal to $m_1 \in \mathbb{C}^{k-1}$. $T_{P_{F,j_k}(J)}j_k(\mathbb{C}P^{k-2})$ is generated by the vectors tangent to such paths $\eta_z(t)$ at $t = 0$. Such a vector, has the form $zm_1^* + m_1z^*$, as a matrix in $S(k-1, \mathbb{C})$. Thus we take $z = m_a,\ a \in \{2, \ldots, k-1\}$, or $z = im_a,\ a \in \{2, \ldots, k-1\}$ based on the eigenvectors of J are orthogonal w.r.t. the complex scalar product. We normalize these vectors to have unit length to obtain the orthonormal frame.

$$e_a(P_{F,j_k}(J)) = d_{[m_1]}j(m_a) = 2^{-1/2}(m_a m_1^* + m_1 m_a^*),$$

$$e'_a(P_{F,j_k}(J)) = d_{[m_1]}j(m_a) = i2^{-1/2}(m_a m_1^* + m_1 m_a^*).$$

As we assume J being diagonal, in this case $m_a = e_a,\ e_a(P_{F,j_k}(J)) = 2^{-1/2}E_1^a$ and $e'_a(P_{F,j_k}(J)) = 2^{-1/2}F_1^a$, where E_a^b has the positions (a, b) and (b, a) that are equal to 1 and all other entries zero, and F_a^b has all the positions (a, b) and (b, a) that are equal to i, respectively $-i$ and other entries zero. We have the differential $d_J P_{F,j_k}(E_a^b) = d_J P_{F,j_k}(F_a^b) = 0,\ \forall 1 < a \le b \le k-1$, and

$$d_J P_{F,j_k}(E_1^a) = (\hat{\lambda}_a - \hat{\lambda}_1)^{-1} e_a(P_{F,j_k}(J)),$$

$$d_J P_{F,j_k}(F_1^a) = (\hat{\lambda}_a - \hat{\lambda}_1)^{-1} e'_a(P_{F,j_k}(J)).$$

We evaluate the VW sample anticovariance matrix $aS_{E,n}$ in formula (25) in [22] using the real scalar product in $S(k-1, \mathbb{C})$, namely, $U \cdot V = ReTr(UV^*)$. Note that,

$$\begin{aligned} &d_J P_{F,j_k}(E_1^b) \cdot e_a(P_{F,j_k}(J)) = (\hat{\lambda}_a - \hat{\lambda}_1)^{-1} \delta_{ba}, \\ &d_J P_{F,j_k}(E_1^b) \cdot e'_a(P_{F,j_k}(J)) = 0 \end{aligned} \tag{18}$$

and

$$\begin{aligned} &d_J P_{F,j_k}(F_1^b) \cdot e'_a(P_{F,j_k}(J))^T = (\hat{\lambda}_a - \hat{\lambda}_1)^{-1} \delta_{ba}, \\ &d_J P_{F,j_k}(F_1^b) \cdot e_a(P_{F,j_k}(J)) = 0. \end{aligned} \tag{19}$$

Thus we may regard $aS_{VW,n}$ as a complex matrix,and in this particular case, we obtain the following entries of this matrix:

$$(aS_{VW,n})_{ab} = n^{-1}(\hat{\lambda}_a - \hat{\lambda}_1)^{-1}(\hat{\lambda}_b - \hat{\lambda}_1)^{-1} \sum_{r=1}^{n} (e_a \cdot z_r)(e_b \cdot z_r)^* |m_1 \cdot z_r|^2, \tag{20}$$

thus proving (17) when J is diagonal. The general case follows by equivariance. $\square$

Next we consider the statistic

$$T(a\bar{X}_{VW}, \alpha\mu_{VW}) = n\|(aS_{VW,n})^{-1/2} tan(P_{F,j_k}(\overline{j_k(X)}) - P_{F,j_k}(\mu_{j_k(X_1)}))\|^2$$

given in [22], in our context of i.i.d.r.o.'s on a complex projective space, to get:

Theorem 3. *Let $X_r = [Z_r]$, $Z_r^* Z_r = 1$, $r \in \{1, \dots, n\}$, be i.i.d. random objects from a αVW nonfocal probability measure Q on $\mathbb{C}P^{k-2}$. Then the random variable T given by*

$$T([m], [\nu]) = n[(m \cdot \nu_a)_{a \in \{2,\dots,k-1\}}](aS_{VW,n})^{-1}[(m \cdot \nu_a)_{a \in \{2,\dots,k-1\}}]^* \tag{21}$$

where $[m] = a\bar{X}_{VW}$ and $\nu_2, \dots, \nu_{k-1}$ yield an orthogonal basis over $\mathbb{C}$ in the tangent space $T_{[\nu_1]}\mathbb{C}P^{k-2}$, has asymptotically a χ^2_{2k-4} distribution.

Proof. Since the VW embedding j_k is by definition isometric, and $(\nu_2, \dots, \nu_{k-1}, \nu_2^*, \dots \nu_{k-1}^*)$ is an orthogonal basis in the tangent space $T_{[\nu_1]}\mathbb{C}P^{k-2}$, the first elements of the adapted orthogonal moving frame are $e_a(P_{F,j_k}(\mu)) = (d_{[\nu_1]}j_k)(\nu_a)$, $e_a^*(P_{F,j_k}(\mu)) = (d_{[\nu_1]}j_k)(\nu_a^*)$. Then the ath tangential component of $P_{F,j_k}(j_k([m])) - P_{F,j_k}(\mu_{j(X_1)})$ w.r.t. this basis of $T_{P_{F,j_k}(\mu)}\mathbb{C}P^{k-2}$ equals up to a sign to the component of $m - \nu_1$ w.r.t. the orthobasis $\nu_2, \dots, \nu_{k-1}$ in $T_{[\nu_1]}\mathbb{C}P^{k-2}$, which is $\nu_a^T m$; and the a^*th tangential components are given by $\nu_a^{*T} m$, and together(in complex multiplication) they yield the complex vector $[(m \cdot \nu_a)_{a=\{2,\dots,k-1\}}]$. The result follows by taking $[m] = P_{F,j_k}(\overline{j_k(X)}) = j_k(a\bar{X}_{VW})$ $\square$

Remark 5. Note that the original VW embedding used in Applied Statistics, was for the complex projective space $\mathbb{C}P^{k-2}$ case [15], given its interpretation as a Kendall planar shape space. We added the Veronese embedding of $\mathbb{R}P^{m-1}$, without specifying the VW subscript in the extrinsic sample anticovariance matrix, to distinguish it from the complex case, and also, for future considerations, since the space of projective shapes of k-ads in general position can be represented as a product of real projective spaces [21, p.149].

We derive large sample confidence regions for the VW-antimean planar Kendall shape. Assume $\chi^2_{s,\beta}$ is a positive number, such that $Pr(U > \chi^2_{s,\beta}) = \beta$, where U has a χ^2_s distribution. Given that the extrinsically studentized tangential component of $P_{F,j_k}(j_k([m])) - P_{F,j_k}(\mu_{j(X_1)})$ yields an error of order $O(n^{-\frac{1}{2}})$, its norm gives an error of order $O(n^{-1})$, leading to the following result.

Corollary 1. *Assume $x_r = [z_r]$, $z_r^* z_r = 1$, $r \in \{1, \dots, n\}$, is a random sample from a αVW nonfocal probability measure Q on $\mathbb{C}P^{k-2}$. An asymptotic $(1-\beta)$—confidence region for $\alpha\mu_{VW}(Q) = [\nu]$ is given by $R_\beta(X) = \{[\nu] : T([m], [\nu]) \le \chi^2_{2k-4,\beta}\}$, where $T([m], [\nu])$ is given in (21). If Q has a nonzero absolutely continuous component w.r.t. the volume measure on $\mathbb{C}P^{k-2}$, then the coverage error of $R_\beta(X)$ is of order $O(n^{-1})$.*

Example 1. In regards with the application in Section 2, the sample size $n = 62$, is large even if the complex projective space dimension is $d = 14$, since the factor behind the shape variable considered is over time environmental adaptation (essentially a univariate variable), and the only restriction is having non-degenerate VW-sample covariance or VW-sample anticovariance matrices, which mathematically is $n-d-1$ being large in the univariate statistics sense, say $n-d-1 > 40$, an inequality that holds true when $n = 62$ and $d = 14$. From Corollary 1, one may construct large sample confidence regions for the VW-antimean. Using Eq. (20), one obtains the sample VW-anticovariance matrix $aS_{VW,n}$, given below:

$$\begin{pmatrix} 0.44+0.00i & 0.08+0.06i & 0.05+0.09i & -0.10-0.05i & -0.03+0.04i & -0.01-0.02i \\ 0.08-0.06i & 0.02+0.00i & 0.02+0.01i & -0.03+0.01i & 0.00+0.01i & -0.01-0.00i \\ 0.05-0.09i & 0.02-0.01i & 0.03+0.00i & -0.02+0.02i & 0.00+0.01i & -0.01+0.00i \\ -0.10+0.05i & -0.03-0.01i & -0.02-0.02i & 0.03+0.00i & 0.00-0.01i & 0.01+0.00i \\ -0.03-0.04i & 0.00-0.01i & 0.00-0.01i & 0.00+0.01i & 0.01+0.00i & -0.00+0.00i \\ -0.01+0.02i & -0.01+0.00i & -0.01-0.00i & 0.01-0.00i & -0.00-0.00i & 0.00+0.00i \end{pmatrix}$$

Note that the sample VW antimean $a\bar{X}_{VW} = [m]$, where m^T, in spherical coordinates, is given by the row matrix.

$$\begin{pmatrix} 0.12+0.03i & 0.41-0.054i & 0.33+0.03i & 0.10-0.06i & 0.74+0.00i & 0.17-0.21i & 0.21+0.16i. \end{pmatrix}$$

When the sample size is small, the coverage error could be quite large, and a nonparametric bootstrap analogue of Theorem 3 is preferred [12]. We recall the steps that one takes to obtain a bootstrapped statistic from a pivotal statistic. If $\{X_r\}_{r\in\{1,\ldots,n\}}$ is a random sample from the unknown distribution Q, and $\{X_r^*\}_{r\in\{1,\ldots,n\}}$ is a random sample from the empirical $\hat{Q}_n$, conditionally given $\{X_r\}_{r\in\{1,\ldots,n\}}$, then the statistic $T(X*, \hat{Q}_n)$ is obtained from $T(X, Q)$, by substituting $X_1^*, \ldots.., X_n^*$ for $X_1, \ldots..X_n$, for $\overline{j_k(X)}$ for μ and $\overline{j_k(X)}^*$ for $\overline{j_k(X)}$. Typically the number of re-samples M should be at least 1000, however often times, a few hundreds of re-samples suffice to derive confidence regions. The following result follows from Theorem 2.2. in [18]. Since $\mathbb{C}P^{k-2}$ is compact, all the moments of $j_k(X)$ are finite, which along with an assumption of a nonzero absolutely continuous component, suffices to ensure an Edgeworth expansion up to order $O(n^{-2})$ of the pivotal statistic $T(a\bar{X}_{VW}, \alpha\mu_{VW})$ [3,4,11]). We then obtain the following result:

Corollary 2. *Let $X_r = [Z_r]$, $Z_r^* Z_r = 1$, $r \in \{1, \ldots, n\}$, be a i.i.d.r.o.'s from a αVW-nonfocal distribution Q on $\mathbb{C}P^{k-2}$, such that X_1 has a nonzero absolutely continuous component w.r.t. the volume measure on $\mathbb{C}P^{k-2}$. If j_k is the VW embedding, and the restriction of the covariance matrix of $j_k(X_1)$ to $T_{[\nu]}j_k(\mathbb{C}P^{k-2})$ is non-degenerate, where $\alpha\mu_E(Q) = [\nu]$ be the extrinsic antimean of Q. For a bootstrap re-sample $\{X_r^*\}_{r\in\{1,\ldots,n\}}$ from the given sample, consider the matrix $J^* := n^{-1}\sum Z_r^* Z_r^{**}$. Let $(m_a^*)_{a\in\{1,\ldots,k-1\}}$ be the unit complex eigenvectors, corresponding to the eigenvalues $(\hat{\lambda}_a^*)_{a\in\{1,\ldots,k-1\}}$ in increasing order. Let $(aS_{VW,n})^*$ be the matrix obtained from $aS_{VW,n}$ by substituting all the entries with $*$-entries. Then,* (i) *bootstrap distribution function of*

$$T([m]^*, [m]) = n[(m_1^* \cdot m_a)_{a\in\{2,\ldots,k-1\}}](aS_{VW,n}^*)^{-1}[(m_1^* \cdot m_a)_{a\in\{2,\ldots,k-1\}}]^* \tag{22}$$

approximates the true distribution function of $T([m], [\nu])$ given in Theorem 3 with an error of order $O_P(n^{-2})$ and (ii) *a nonparametric bootstrap $(1-\beta)$—confidence region for $\alpha\mu_{VW}(Q) = [\nu]$ is given by $R_\beta^*(X) = \{[\nu] : T([m], [\nu]) \le T^*\}$, where $T([m], [\nu])$ is given in (21), and T^* is the $1-\beta$ quantile of the bootstrap distribution of $T([m]^*, [m])$. If Q has a nonzero absolutely continuous component w.r.t. the volume measure on $\mathbb{C}P^{k-2}$, then the coverage error of $R_\beta^*(X)$ is of order $O(n^{-2})$.*

Based on the simulated data displayed in Fig. 1 in the supplementary material, we re-sampled for 1000 times, and we got the .95 quantile cutoff value from Corollary 2 to be $T^* = 8.4785$.

CRediT authorship contribution statement

Yunfan Wang: Data curation, Formal analysis, Investigation, Project administration, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Vic Patrangenaru:** Conceptualization, Methodology, Supervision, Writing - original draft, Writing - review & editing. **Ruite Guo:** Methodology, Writing - review & editing.

Acknowledgments

The authors are indebted to the editor, the associate editor who handled the ms and the two referees for their constructive and valuable comments, that helped us significantly improve the contents, presentation and adjustments to the journal style of the paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2020.104600.

References

[1] R.J. Beran, Testing for uniformity on a compact homogeneous space, J. Appl. Probab. 5 (1968) 77–195.
[2] R. Beran, N.I. Fisher, Nonparametric comparison of mean directions or mean axes, Ann. Statist. 26 (1998) 472–493.
[3] R.N. Bhattacharya, J.K. Ghosh, On the validity of the formal Edgeworth expansion, Ann. Statist. 6 (1978) 434–451.
[4] R.N. Bhattacharya, L. Lin, V. Patrangenaru, A Course in Mathematical Statistics and Large Sample Theory, Springer, Statistics Series, New York, 2016.

[5] R.N. Bhattacharya, V. Patrangenaru, Large sample theory of intrinsic and extrinsic sample means on manifolds-Part I, Ann. Statist. 31 (2003) 1–29.
[6] R.N. Bhattacharya, V. Patrangenaru, Large sample theory of intrinsic and extrinsic sample means on manifolds-Part II, Ann. Statist. 33 (2005) 1211–1245.
[7] F.L. Bookstein, Morphometric Tools for Landmark Data: Geometry and Biology, Cambridge University Press, 1997.
[8] I.L. Dryden, K.V. Mardia, Statistical Shape Analysis with Applications in R. Second Edition, in: Wiley Series in Probability and Statistics, John Wiley and Sons, Ltd., Chichester, 2016.
[9] N. Fisher, P. Hall, B. Jing, A. Wood, Improved pivotal methods for constructing confidence regions with directional data, J. Amer. Statist. Assoc. 91 (1996) 435, 1062–1070.
[10] M. Fréchet, Les élements aléatoires de nature quelconque dans un espace distancié, Ann. Inst. H. Poincaré 10 (1948) 215–310.
[11] P. Hall, Theoretical comparison of bootstrap confidence intervals. With a discussion and a reply by the author, Ann. Statist. 16 (1988) 927–985.
[12] P. Hall, The Bootstrap and Edgeworth Expansion, Springer-Verlag, New York, 1992.
[13] H. Hendriks, Z. Landsman, Mean location and sample mean location on manifolds: asymptotics, tests, confidence regions, J. Multivariate Anal. 67 (1998) 227–243.
[14] D.G. Kendall, Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces, Bulletin of the London Mathematical Society, Oxford University Press, 1992.
[15] J. Kent, New directions in shape analysis, in: The Art of Statistical Science, A Tribute to G.S. Watson, in: Wiley Ser. Probab. Math. Statist., Wiley, Chichester, 1992, pp. 115–127.
[16] J. Kent, The complex Bingham distribution and shape analysis, J. R. Stat. Soc. Ser. B Stat. Methodol. 56 (1994) 285–299.
[17] H.T. Ku, L. Mann, J.L. Sicks, J. Su, Degree of symmetry of a product manifold, Trans. Amer. Math. Soc. 146 (1969) 133–149.
[18] H. Lee, V. Patrangenaru, Anti-MANOVA on compact manifolds with applications to 3D projective shape analysis, 2019, arXiv:1909.00320.
[19] K.V. Mardia, V. Patrangenaru, Directions and projective shapes, Ann. Statist. 33 (2005) 1666–1699.
[20] V. Patrangenaru, Y. Deng, Extrinsic regression and anti-regression on projective shape manifolds, Methodol. Comput. Appl. Probab. (2019) Revision submitted.
[21] V. Patrangenaru, L. Ellingson, Nonparametric Statistics on Manifolds and Their Applications to Object Data Analysis, CRC-Chapman and Hall, 2015.
[22] V. Patrangenaru, R. Guo, K.D. Yao, Nonparametric inference for location parameters via Fréchet functions, in: Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, 2016, pp. 254–262.
[23] V. Patrangenaru, X. Liu, S. Sugathadasa, Nonparametric 3D projective shape estimation from pairs of 2D images - I, in memory of W.P. Dayawansa, J. Multivariate Anal. 101 (2010) 11–31.
[24] V. Patrangenaru, K.V. Mardia, Affine shape analysis and image analysis, in: Proceedings of the Leeds Annual Statistics Research Workshop, Leeds University Press, 2003, pp. 57–62.
[25] V. Patrangenaru, K. Yao, R. Guo, Extrinsic means and antimeans, in: R. Cao, W. González Manteiga, J. Romo (Eds.), Nonparametric Statistics, Vol. 175, Springer Proceedings in Mathematics and Statistics, 2016, pp. 161–178.
[26] E.H. Spanier, Algebraic Topology. Corrected Reprint, Springer-Verlag, New York-Berlin, 1981.
[27] M.S. Sugathadasa, Affine and Projective Shape Analysis with Applications (Ph.D. dissertation), Texas Tech University, Department of Mathematics, 2006.
[28] G.S. Watson, Statistics on Spheres, University of Arkansas Lecture Notes in the Mathematical Sciences, Wiley-Interscience, 1983.