Extrinsic Means and Antimeans

Vic Patrangenaru[1,*], K. David Yao[2,†], Ruite Guo[1,‡]
Florida State University
[1] Department of Statistics, [2] Department of Mathematics

December 28, 2015

1 Introduction

Fréchet (1948) noticed that for data analysis purposes, in case a list of numbers would not give a meaningful representation of the individual observation under investigation, it is helpful to measure not just vectors, but more complicated features, he used to call "elements", and are nowadays called *objects*. As examples he mentioned "the shape of an egg taken at random from a basket of eggs". A natural way of addressing this problem consists of regarding a *random object* X as a random point on a complete metric space $(\mathcal{M}, \rho)$ that often times has a manifold structure (see Patrangenaru and Ellingson (2015)). Important examples of objects that arise from electronic image data are shapes of configurations extracted from digital images, or from medical imaging outputs. For such data, the associated object considered are points on Kendall shape spaces (see Kendall (1984), Dryden and Mardia (1998)), or on affine shape spaces (see Patrangenaru and Mardia (2003), Sughatadasa (2006)), on projective shape spaces (see Mardia and Patrangenaru (2005), Patrangenaru et al. (2010)). Other examples of object spaces are spaces of axes (see Fisher et al. (1996), Beran and Fisher (1998)), spaces of directions (see Watson (1982)) and spaces of trees (see Billera et al. (2001), Wang and Marron (2007), Hotz et al. (2013)). The afore mentioned object spaces have a structure of *compact symmetric spaces* (see Helgasson (2001)), however, the use of a Riemannian distance on a symmetric space for the goal of mean data analysis, including for regression with a response on a symmetric space, is a statistician choice, as opposed to being imposed by the nature of the data.
Therefore for practical purposes, in this paper we consider object spaces provided with a "chord" distance associated with the embedding of an object space into a numerical space, and the statistical analysis performed relative to a chord distance is termed *extrinsic data analysis*. The expected square distance from the random object X to an arbitrary point p defines what we call the *Fréchet function* associated with X :

$$\mathcal{F}(p) = \mathbb{E}(\rho^2(p, X)), \tag{1.1}$$

and its minimizers form the *Fréchet mean set*. When ρ is the "chord" distance on $\mathcal{M}$ induced by the Euclidean distance in $\mathbb{R}^N$ via an embedding $j : \mathcal{M} \to \mathbb{R}^N$, the Fréchet function becomes

$$\mathcal{F}(p) = \int_{\mathcal{M}} \|j(x) - j(p)\|_0^2 Q(dx), \tag{1.2}$$

where $Q = P_X$ is the probability measure on $\mathcal{M}$, associated with X. In this case the Fréchet mean set is called the *extrinsic mean set* (see Bhattacharya and Patrangenaru (2003)), and if we have a unique point in the extrinsic mean set of X, this point is the *extrinsic mean* of X, and is labeled $\mu_E(X)$ or simply μ_E. Also,

[*] Research supported by NSA-MSP- H98230-14-1-0135 and NSF-DMS-1106935
[†] Research supported by NSA-MSP- H98230-14-1-0135
[‡] Research supported by NSF-DMS-1106935

given $X_1, \dots, X_n$ i.i.d random objects from Q, their *extrinsic sample mean (set)* is the extrinsic mean (set) of the empirical distribution $\hat{Q}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$.
In this paper we will assume that $(\mathcal{M}, \rho)$ is a compact metric space, therefore the Fréchet function is bounded, and its extreme values are attained at points on $\mathcal{M}$. We are now introducing a **new location parameter** for X.

DEFINITION 1.1. *The set of maximizers of the Fréchet function, is called the* extrinsic antimean set*. In case the extrinsic antimean set has one point only, that point is called* **extrinsic antimean** *of* X*, and is labeled* $\alpha\mu_{j,E}(Q)$*, or simply* $\alpha\mu_E$*, when* j *is known.*

The remainder of the paper is concerned with geometric descriptions, explicit formulas and computations of extrinsic means and antimeans. Simple inference problems for extrinsic means and antimeans are also investigated. The paper ends with a discussion on future directions in extrinsic antimean analysis.

2 Geometric description of the extrinsic antimean

Let $(\mathcal{M}, \rho_0)$ be a compact metric space, where ρ_0 is the chord distance via the embedding $j : \mathcal{M} \to \mathbb{R}^N$, that is

$$\rho_0(p_1, p_2) = \|j(p_1) - j(p_2)\| = d_0(j(p_1), j(p_2)), \forall (p_1, p_2) \in \mathcal{M}^2,$$

where d_0 is the Euclidean distance in $\mathbb{R}^N$.

REMARK 2.1. *Recall that a point $y \in \mathbb{R}^N$ for which there is a unique point $p \in \mathcal{M}$ satisfying the equality,*

$$d_0(y, j(\mathcal{M})) = \inf_{x \in \mathcal{M}} \|y - j(x)\|_0 = d_0(y, j(p))$$

is called j-nonfocal. A point which is not j-nonfocal is said to be j-focal. And if y is a j-nonfocal point, its projection on $j(\mathcal{M})$ is the unique point $j(p) = P_j(y) \in j(\mathcal{M})$ with $d_0(y, j(\mathcal{M})) = d_0(y, j(p))$.

With this in mind we now have the following definition.

DEFINITION 2.1. *(a) A point $y \in \mathbb{R}^N$ for which there is a unique point $p \in \mathcal{M}$ satisfying the equality,*

$$\sup_{x \in \mathcal{M}} \|y - j(x)\|_0 = d_0(y, j(p)) \tag{2.1}$$

is called αj-nonfocal. A point which is not αj-nonfocal is said to be αj-focal.

(b) If y is an αj-nonfocal point, its projection on $j(\mathcal{M})$ is the unique point $z = P_{F,j}(y) \in j(\mathcal{M})$ with $\sup_{x \in \mathcal{M}} \|y - j(x)\|_0 = d_0(y, j(p))$.

For example if we consider the unit sphere S^m in $\mathbb{R}^{m+1}$, with the embedding given by the inclusion map $j : S^m \to \mathbb{R}^{m+1}$, then the only αj-focal point is 0_{m+1}, the center of this sphere; this point also happens to be the only j-focal point of S^m.

DEFINITION 2.2. *A probability distribution Q on $\mathcal{M}$ is said to be αj-nonfocal if the mean μ of $j(Q)$ is αj-nonfocal.*

The figures below illustrate the extrinsic mean and antimean of distributions on a complete metric space $\mathcal{M}$ where the distributions are j-nonfocal and also αj-nonfocal.

THEOREM 2.1. *Let μ be the mean vector of $j(Q)$ in $\mathbb{R}^N$. Then the following hold true:*

(i) The extrinsic antimean set is the set of all points $x \in \mathcal{M}$ such that $\sup_{p \in \mathcal{M}} \|\mu - j(p)\|_0 = d_0(\mu, j(x))$.

(ii) If $\alpha\mu_{j,E}(Q)$ exists, then μ is αj-nonfocal and $\alpha\mu_{j,E}(Q) = j^{-1}(P_{F,j}(\mu))$.

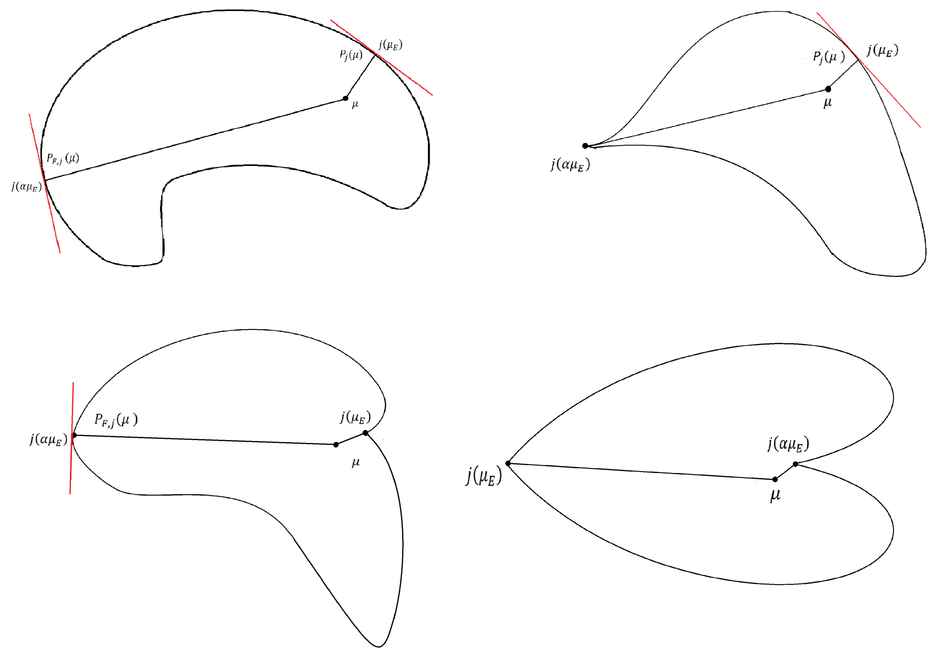

Figure 1: Extrinsic mean and extrinsic antimean on a 1-dimensional topological manifold (upper left: regular mean and antimean, upper right: regular mean and sticky antimean, lower left: sticky mean and regular antimean, lower right : sticky mean and antimean

Proof. For part (i), we first rewrite the following expression;

(2.2) $$\|j(p) - j(x)\|_0^2 = \|j(p) - \mu\|_0^2 - 2\,\langle\, j(p) - \mu,\ \mu - j(x)\,\rangle + \|\mu - j(x)\|_0^2$$

Since the manifold is compact, μ exists, and from the definition of the mean vector we have

(2.3) $$\int_{\mathcal{M}} j(x)\, Q(dx) = \int_{\mathbb{R}^N} yj(Q)(dy) = \mu.$$

From equations (2.3), (2.2) it follows that

(2.4) $$\mathcal{F}(p) = \|j(p) - \mu\|_0^2 + \int_{\mathbb{R}^N} \|\mu - y\|_0^2 j(Q)(dy)$$

Then, from (2.4),

(2.5) $$\sup_{p\in\mathcal{M}} \mathcal{F}(p) = \sup_{p\in\mathcal{M}} \|j(p) - \mu\|_0^2 + \int_{\mathbb{R}^N} \|\mu - y\|_0^2\, j(Q)(dy)$$

This then implies that the antimean set is the set of points $x \in \mathcal{M}$ with the following property;

(2.6) $$\sup_{p\in\mathcal{M}} \|j(p) - \mu\|_0 = \|j(x) - \mu\|_0.$$

For Part (ii) if $\alpha\mu_{j,E}(Q)$ exists, then $\alpha\mu_{j,E}(Q)$ is the unique point $x \in \mathcal{M}$, for which equation (2.6) holds true, which implies that μ is αj-nonfocal and $j(\alpha\mu_{j,E}(Q)) = P_{F,j}(\mu)$. $\square$

DEFINITION 2.3. *Let $x_1, ..., x_n$ be random observations from a distribution Q on a compact metric space $(\mathcal{M},\ \rho)$, then their extrinsic sample antimean set, is the set of maximizers of the Fréchet function $\hat{\mathcal{F}}_n$ associated with the empirical distribution $\hat{Q}_n = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$, which is given by*

$$\hat{\mathcal{F}}_n(p) = \frac{1}{n}\sum_{i=1}^n \|j(x_i) - j(p)\|_0^2 \tag{2.7}$$

If $\hat{Q}_n$ has an extrinsic antimean, its extrinsic antimean is called extrinsic sample antimean, and it is denoted $a\bar{X}_{j,E}$.

THEOREM 2.2. *Assume Q is an αj-nonfocal probability measure on the manifold $\mathcal{M}$ and $X = \{X_1, ..., X_n\}$ are i.i.d random objects from Q. Then,*

(a) If $\overline{j(X)}$ is αj-nonfocal, then the extrinsic sample antimean is given by $a\bar{X}_{j,E} = j^{-1}(P_{F,j}(\overline{j(X)}))$.

(b) The set $(\alpha F)^c$ of αj-nonfocal points is a generic subset of $\mathbb{R}^N$, and if $\alpha\mu_{j,E}(Q)$ exists, then the extrinsic sample antimean $a\bar{X}_{j,E}$ is a consistent estimator of $\alpha\mu_{j,E}(Q)$.

Proof. (Sketch). (a) Since $\overline{j(X)}$ is αj-nonfocal the result follows from Theorem 2.1, applied to the empirical $\hat{Q}_n$, therefore $j(a\bar{X}_{j,E}) = P_{F,j}(\overline{j(X)})$.
(b) All the assumptions of the SLLN are satisfied, since $j(\mathcal{M})$ is also compact, therefore the sample mean estimator $\overline{j(X)}$ is a strong consistent estimator of μ, which implies that for any $\varepsilon > 0$, and for any $\delta > 0$, there is sample size $n(\delta, \varepsilon)$, such that $\mathbb{P}(\|\overline{j(X)} - \mu\| > \delta) \le \varepsilon, \forall n > n(\delta, \varepsilon)$. By taking a small enough $\delta > 0$, and using a continuity argument for $P_{F,j}$, the result follows. □

REMARK 2.2. *For asymptotic distributions of the extrinsic sample antimeans see Patrangenaru et al.(2016).*

3 VW antimeans on $\mathbb{R}P^m$

In this section we consider the case of a probability measure Q on the real projective space $\mathcal{M} = \mathbb{R}P^m$, which is the set of axes (1-dimensional linear subspaces) of $\mathbb{R}^{m+1}$. Here the points in $\mathbb{R}^{m+1}$ are regarded as $(m+1) \times 1$ vectors. $\mathbb{R}P^m$ can be identified with the quotient space $S^m/\{x, -x\}$; it is a compact homogeneous space, with the group $SO(m+1)$ acting transitively on $(\mathbb{R}P^m, \rho_0)$, where the distance ρ_0 on $\mathbb{R}P^m$ is induced by the chord distance on the sphere S^m. There are infinitely many embeddings of $\mathbb{R}P^m$ in a Euclidean space, however for the purpose of two sample mean or two sample antimean testing, it is preferred to use an embedding j that is compatible with two transitive group actions of $SO(m+1)$ on $\mathbb{R}P^m$, respectively on $j(\mathbb{R}P^m)$, that is

$$j(T \cdot [x]) = T \otimes j([x]),\ \ \forall\, T \in SO(m+1), \forall\, [x] \in \mathbb{R}P^m,\ \ where\ T \cdot [x] = [Tx]. \tag{3.1}$$

Such an embedding is said to be *equivariant* (see Kent (1992)). For computational purposes, the equivariant embedding of $\mathbb{R}P^m$ that was used so far in the axial data analysis literature is the Veronese-Whitney (VW) embedding $j : \mathbb{R}P^m \to S_+(m+1, \mathbb{R})$, that associates to an axis the matrix of the orthogonal projection on this axis (see Patrangenaru and Ellingson(2015) and references therein):

$$j([x]) = xx^T, \|x\| = 1, \tag{3.2}$$

Here $S_+(m+1, \mathbb{R})$ is the set of nonnegative definite symmetric $(m+1) \times (m+1)$ matrices, and in this case

$$T \otimes A = TAT^T,\ \ \forall\, T \in SO(m+1), \forall\, A \in S_+(m+1, \mathbb{R}) \tag{3.3}$$

REMARK 3.1. *Let $N = \frac{1}{2}(m+1)(m+2)$. The space $\mathbb{E} = (S(m+1, \mathbb{R}),\ \langle\ ,\ \rangle_0)$ is an N-dimensional Euclidean space with the scalar product given by $\langle A,\ B\rangle_0 = Tr(AB),\ where\ \ A,\ B\ \in S(m+1, \mathbb{R})$. The associated norm $\|\cdot\|_0$ and Euclidean distance d_0 are given by respectively by $\|C\|_0^2 = \langle C,\ C\rangle_0$ and $d_0(A, B) = \|A - B\|_0$, $\forall C, A,\ B\ \in\ S(m+1, \mathbb{R})$.*

With the notation in Remark 3.1 we have

$$\mathcal{F}([p]) = \|j([p]) - \mu\|_0^2 + \int_{\mathcal{M}} \|\mu - j([x])\|_0^2 \, Q(d[x]), \tag{3.4}$$

and $\mathcal{F}([p])$ is maximized (minimized) if and only if $\|j([p]) - \mu\|_0^2$ is maximized (minimized) as a function of $[p] \in \mathbb{R}P^m$.

From Patrangenaru and Ellingson (2015, Chapter 4) and definitions therein, recall that the extrinsic mean $\mu_{j,E}(Q)$ of a j- nonfocal probability measure Q on $\mathcal{M}$ w.r.t. an embedding j, when it exists, is given by $\mu_{j,E}(Q) = j^{-1}(P_j(\mu))$ where μ is the mean of $j(Q)$. In the particular case when $\mathcal{M} = \mathbb{R}P^m$, and j is the VW embedding, P_j is the projection on $j(\mathbb{R}P^m)$ and $P_j : S_+(m+1, \mathbb{R}) \backslash \mathcal{F} \to j(\mathbb{R}P^m)$, where $\mathcal{F}$ is the set of j-focal points of $j(\mathbb{R}P^m)$ in $S_+(m+1, \mathbb{R})$. For the VW embedding, $\mathcal{F}$ is the set of matrices in $S_+(m+1, \mathbb{R})$ whose largest eigenvalues are of multiplicity at least 2. The projection P_j assigns to each nonnegative definite symmetric matrix A with highest eigenvalue of multiplicity 1, the matrix vv^T, where v is a unit eigenvector of A corresponding to its largest eigenvalue.

Furthermore, the VW mean of a random object $[X] \in \mathbb{R}P^m, [X^T X] = 1$ is given by $\mu_{j,E}(Q) = [\gamma(m+1)]$ and $(\lambda(a), \gamma(a)),\ a = 1, .., m+1$ are eigenvalues and unit eigenvectors pairs (in increasing order of eigenvalues) of the mean $\mu = E(XX^T)$. Similarly, the VW sample mean is given by $\bar{x}_{j,E} = [g(m+1)]$ where $(d(a), g(a)),\ a = 1, \ldots, m+1$ are eigenvalues and unit eigenvectors pairs (in increasing order of eigenvalues) of the sample mean $J = \frac{1}{n}\sum_{i=1}^n x_i x_i^T$ associated with the sample $([x_i])_{i=\overline{1,n}}$, on $\mathbb{R}P^m$, where $x_i^T x_i = 1, \forall i = \overline{1,n}$.

Based on (3.4), we get similar results in the case of an αj-nonfocal probability measure Q :

PROPOSITION 3.1. *(i) The set of αVW-nonfocal points in $S_+(m+1, \mathbb{R})$, is the set of matrices in $S_+(m+1, \mathbb{R})$ whose smallest eigenvalue has multiplicity 1.*

(ii) The projection $P_{F,j} : (\alpha F)^c \to j(\mathbb{R}P^m)$ assigns to each nonnegative definite symmetric matrix A, of rank 1, with a smallest eigenvalue of multiplicity 1, the matrix $j([\nu])$, where $\|\nu\| = 1$ and ν is an eigenvector of A corresponding to that eigenvalue.

We now have the following;

PROPOSITION 3.2. *Let Q be a distribution on $\mathbb{R}P^m$.*

(a) The VW-antimean set of a random object $[X], X^T X = 1$ on $\mathbb{R}P^m$, is the set of points $p = [v] \in V_1$, where V_1 is the eigenspace corresponding to the smallest eigenvalue $\lambda(1)$ of $E(XX^T)$.

(b) If in addition $Q = P_{[X]}$ is αVW-nonfocal, then

$$\alpha\mu_{j,E}(Q) = j^{-1}(P_{F,j}(\mu)) = [\gamma(1)]$$

where $(\lambda(a), \gamma(a)),\ a = 1, .., m+1$ are eigenvalues in increasing order and the corresponding unit eigenvectors of $\mu = E(XX^T)$.

(c) Let $[x_1], \ldots, [x_n]$ be observations from a distribution Q on $\mathbb{R}P^m$, such that $\overline{j(X)}$ is αVW-nonfocal. Then the VW sample antimean of $[x_1], \ldots, [x_n]$ is given by

$$a\overline{x}_{j,E} = j^{-1}(P_{F,j}(\overline{j(x)})) = [g(1)]$$

where $(d(a), g(a))$ are the eigenvalues in increasing order and the corresponding unit eigenvectors of $J = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T$, where $x_i^T x_i = 1, \forall i = \overline{1,n}$.

4 Two-sample test for VW means and antimeans projective shapes in 3D

Recall that the space $P\Sigma_3^k$ of projective shapes of 3D k-ads in $\mathbb{R}P^3$, $([u_1],...,[u_k])$, with $k > 5$, for which $\pi = ([u_1],\ldots,[u_5])$ is a projective frame in $\mathbb{R}P^3$, is homeomorphic to the manifold $(\mathbb{R}P^3)^q$ with $q = k-5$ (see Patrangenaru et al.(2010)). Also recall that a Lie group , is a manifold $\mathcal{G}$, that has an additional group structure $\odot : \mathcal{G}\times\mathcal{G}\to\mathcal{G}$ with the inverse map $\iota : \mathcal{G}\to\mathcal{G}, \iota(g) = g^{-1}$, such that both operations $\odot$ and ι are differentiable functions between manifolds.
Note that S^3 regarded as set of quaternions of unit norm has a Lie group structure inherited from the quaternion multiplication, which yields a Lie group structure on $\mathbb{R}P^3$. This multiplicative structure turns the $(\mathbb{R}P^3)^q$ into a product Lie group $(\mathcal{G},\odot_q)$ where $\mathcal{G} = (\mathbb{R}P^3)^q$ (see Crane and Patrangenaru (2011), Patrangenaru et al. (2014)). For the rest of this section $\mathcal{G}$ refers to the Lie group $(\mathbb{R}P^3)^q$. The VW embedding $j_q : (\mathbb{R}P^3)^q \to (S_+(4,\mathbb{R}))^q$ (see Patrangenaru et al. (2014)), is given by

$$j_q([x_1],\ldots,[x_q]) = (j([x_1]),\ldots,j([x_q])), \tag{4.1}$$

with $j : \mathbb{R}P^3 \to S_+(4,\mathbb{R})$ the VW embedding given in (3.2), for $m = 3$ and j_q is also an equivariant embedding w.r.t. the group $(S_+(4,\mathbb{R}))^q$.
Given the product structure, it turns out that the VW mean μ_{j_q} of a random object $Y = (Y^1,\ldots,Y^q)$ on $(\mathbb{R}P^3)^q$ is given by

$$\mu_{j_q} = (\mu_{1,j},\cdots,\mu_{q,j}), \tag{4.2}$$

where, for $s = \overline{1,q},\ \mu_{s,j}$ is the VW mean of the marginal Y^s.
Assume $Y_a,\ a = 1,2$ are random objects with the associated distributions $Q_a = P_{Y_a},\ a = 1,2$ on $\mathcal{G} = (\mathbb{R}P^3)^q$. We now consider the two sample problem for VW means and separately for VW-antimeans for these random objects.

4.1 Hypothesis testing for VW means

Assume the distributions $Q_a, a = 1,2$ are in addition VW-nonfocal. We are interested in the hypothesis testing problem:

$$H_0 :\ \mu_{1,j_q} = \mu_{2,j_q} \text{ vs. } H_a :\ \mu_{1,j_q} \neq \mu_{2,j_q}, \tag{4.3}$$

which is equivalent to testing the following

$$H_0 :\ \mu_{2,j_q}^{-1}\odot_q \mu_{1,j_q} = 1_{(\mathbb{R}P^3)^q} \text{ vs. } H_a :\ \mu_{2,j_q}^{-1}\odot_q \mu_{1,j_q} \neq 1_{(\mathbb{R}P^3)^q} \tag{4.4}$$

1. Let $n_+ = n_1 + n_2$ be the total sample size, and assume $\lim_{n_+\to\infty}\frac{n_1}{n_+}\to\lambda\in(0,1)$. Let φ be the log chart defined in a neighborhood of $1_{(\mathbb{R}P^3)^q}$ (see Helgason (2001)), with $\varphi(1_{(\mathbb{R}P^3)^q}) = 0$. Then, under H_0

   $$n_+^{1/2}\ \varphi(\bar{Y}_{j_q,n_2}^{-1}\odot_q \bar{Y}_{j_q,n_1}) \to_d \mathcal{N}_{3q}(0_{3q},\Sigma_{j_q}) \tag{4.5}$$

   Where Σ_{j_q} depends linearly on the extrinsic covariance matrices Σ_{a,j_q} of Q_a.

2. Assume in addition that for $a = 1,2$ the support of the distribution of $Y_{a,1}$ and the VW mean μ_{a,j_q} are included in the domain of the chart φ and $\varphi(Y_{a,1})$ has an absolutely continuous component and finite moment of sufficiently high order. Then the joint distribution

   $$V = n_+{}^{\frac{1}{2}}\varphi(\bar{Y}_{j_q,n_2}^{-1}\odot_q \bar{Y}_{j_q,n_1}) \tag{4.6}$$

   can be approximated by the bootstrap joint distribution of

$$V^* = {n_+}^{1/2}\ \varphi(\bar{Y^*}_{j_q,n_2}^{-1} \odot_q \bar{Y}^*_{j_q,n_1})$$

From Patrangenaru et al.(2010), recall that given a random sample from a distribution Q on $\mathbb{R}P^m$, if J_s, $s = 1, \ldots, q$ are the matrices $J_s = n^{-1}\sum_{r=1}^{n} X_r^s (X_r^s)^T$, and if for $a = 1, \ldots, m+1,\ d_s(a)$ and $g_s(a)$ are the eigenvalues in increasing order and corresponding unit eigenvectors of J_s, then the VW sample mean $\bar{Y}_{j_q,n}$ is given by

$$\bar{Y}_{j_q,n} = ([g_1(m+1)], \ldots, [g_q(m+1)]). \tag{4.7}$$

REMARK 4.1. *Given the high dimensionality, the VW sample covariance matrix is often singular. Therefore, for nonparametric hypothesis testing, Efron's nonpivotal bootstrap is preferred. For nonparametric bootstrap methods see eg. Efron(1982). For details, on testing the existence of a mean change 3D projective shape, when sample sizes are not equal, using nonpivotal bootstrap, see Patrangenaru et al. (2014).*

4.2 Hypothesis testing for VW antimeans

Unlike in the previous subsection, we now assume that for $a = 1, 2, Q_a$ are αVW-nonfocal. We are now interested in the hypothesis testing problem:

$$H_0 :\ \alpha\mu_{1,j_q} = \alpha\mu_{2,j_q} \text{ vs. } H_a :\ \alpha\mu_{1,j_q} \neq \alpha\mu_{2,j_q}, \tag{4.8}$$

which is equivalent to testing the following

$$H_0 :\ \alpha\mu_{2,j_q}^{-1} \odot_q \alpha\mu_{1,j_q} = 1_{(\mathbb{R}P^3)^q} \text{ vs. } H_a :\ \alpha\mu_{2,j_q}^{-1} \odot_q \alpha\mu_{1,j_q} \neq 1_{(\mathbb{R}P^3)^q} \tag{4.9}$$

1. Let $n_+ = n_1 + n_2$ be the total sample size, and assume $\lim_{n_+\to\infty} \frac{n_1}{n_+} \to \lambda \in (0, 1)$. Let φ be the log chart defined in a neighborhood of $1_{(\mathbb{R}P^3)^q}$ (see Helgason (2001)), with $\varphi(1_{(\mathbb{R}P^3)^q}) = 0_{3q}$. Then, from Patrangenaru et al. (2016), it follows that under H_0

$$n_+^{1/2}\ \varphi(a\bar{Y}_{j_q,n_2}^{-1} \odot_q a\bar{Y}_{j_q,n_1}) \to_d \mathcal{N}_{3q}(0_{3q}, \tilde{\Sigma}_{j_q}), \tag{4.10}$$

   for some covariance matrix $\tilde{\Sigma}_{j_q}$.

2. Assume in addition that for $a = 1, 2$ the support of the distribution of $Y_{a,1}$ and the VW antimean $\alpha\mu_{a,j_q}$ are included in the domain of the chart φ and $\varphi(Y_{a,1})$ has an absolutely continuous component and finite moment of sufficiently high order. Then the joint distribution

$$aV = {n_+}^{\frac{1}{2}}\varphi(a\bar{Y}_{j_q,n_2}^{-1} \odot_q a\bar{Y}_{j_q,n_1}) \tag{4.11}$$

   can be approximated by the bootstrap joint distribution of

$$aV^* = {n_+}^{1/2}\ \varphi(a\bar{Y^*}_{j_q,n_2}^{-1} \odot_q a\bar{Y}^*_{j_q,n_1})$$

Now, from Proposition 3.2, we get the following result that is used for the computation of the VW sample antimeans.

PROPOSITION 4.1. *follows that given a random sample from a distribution Q on $\mathbb{R}P^m$, if $J_s, s = 1, \ldots, q$ are the matrices $J_s = n^{-1}\sum_{r=1}^{n} X_r^s (X_r^s)^T$, and if for $a = 1, \ldots, m+1, d_s(a)$ and $g_s(a)$ are the eigenvalues in increasing order and corresponding unit eigenvectors of J_s, then the VW sample antimean $a\bar{Y}_{j_q,n}$ is given by*

$$a\bar{Y}_{j_q,n} = ([g_1(1)], \ldots, [g_q(1)]). \tag{4.12}$$

5 Two sample test for lily flowers data

In this section we will test for the existence of 3D mean projective shape change to differentiate between two lily flowers. We will use pairs of pictures of two flowers for our study.
Our data sets consist of two samples of digital images. The first one consist of 11 pairs of pictures of a single lily flower. The second has 8 pairs of digital images of another lily flower.

Figure 2: Flower Sample 1

Figure 3: Flower Sample 2

We will recover the 3D projective shape of a spatial k-ad (in our case $k = 13$) from the pairs of images, which will allow us to test for mean 3D projective shape change detection.
Flowers belonging to the genus Lilum have three petals and three petal-like sepals. It may be difficult to distinguish the lily petals from the sepals. Here all six are referred to as *tepals*. For our analysis we selected 13 anatomic landmarks, 5 of which will be used to construct a projective frame. In order to conduct a proper analysis we recorded the same labeling of landmarks and kept a constant configuration for both flowers.
The tepals where labeled 1 through 6 for both flowers. Also the six *stamens* (male part of the flower) ,were labeled 7 through 12 starting with the stamen that is closely related to tepal 1 and continuing in the same fashion. The landmarks were placed at the tip of the *anther* of each of the six stamens and in the center of the *stigma* for the *carpel* (the female part).

Figure 4: Landmarks for Flower 1 and Flower 2

For 3D reconstructions of k-ads we used the reconstruction algorithm in Ma et al (2004). The first 5 of our

13 landmarks were selected to construct our projective frame π . To each projective point we associated its projective coordinate with respect to π. The projective shape of the 3D k-ad, is then determined by the 8 projective coordinates of the remaining landmarks of the reconstructed configuration.
We tested for the VW mean change, since $(\mathbb{R}P^3)^8$ has a Lie group structure (Crane and Patrangenaru (2011)). Two types of VW mean changes were considered: one for cross validation, and the other for comparing the VW mean shapes of the two flowers.
Suppose Q_1 and Q_2 are independent r.o.'s, the hypothesis for their mean change is

$$H_0 : \mu_{1,j_8}^{-1} \odot_8 \mu_{2,j_8} = 1_{(\mathbb{R}P^3)^8}$$

Given φ, the affine chart on this Lie group, $\varphi(1_{(\mathbb{R}P^3)^8}) = 0_{24}$, we compute the bootstrap distribution

$$D_* = \varphi((\bar{Y}^*_{1,j_8,11})^{-1} \odot_8 \bar{Y}^*_{2,j_8,8})$$

We fail to reject H_0, if all simultaneous confidence intervals contain 0, and reject it otherwise. We construct 95% simultaneous nonparametric bootstrap confidence intervals. We will then expect to fail to reject the null, if we have 0 in all of our simultaneous confidence intervals.

5.1 Results for comparing the two flowers

We would fail to reject our null hypothesis

$$H_0 : \mu_{1,j_8}^{-1} \odot_8 \mu_{2,j_8} = 1_{(\mathbb{R}P^3)^8}$$

if all of our 24 confidence intervals would contain the value 0.

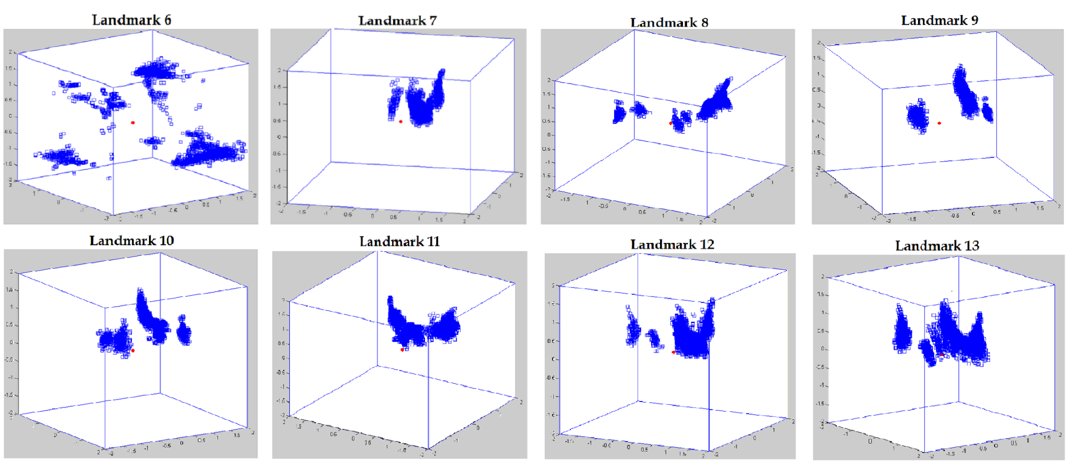

Figure 5: Bootstrap Projective Shape Marginals for lily Data

Simultaneous Confidence Intervals for lily's landmarks 6 to 9				
	LM6	LM7	LM8	LM9
x	$(0.609514, 1.638759)$	$(0.320515, 0.561915)$	$(-0.427979, 0.821540)$	$(0.055007, 0.876664)$
y	$(-0.916254, 0.995679)$	$(-0.200514, 0.344619)$	$(-0.252281, 0.580393)$	$(-0.358060, 0.461555)$
z	$(-1.589983, 1.224176)$	$(0.177687, 0.640489)$	$(0.291530, 0.831977)$	$(0.213021, 0.883361)$

Simultaneous Confidence Intervals for lily's landmarks 10 to 13				
	LM10	LM11	LM12	LM13
x	$(0.060118, 0.822957)$	$(0.495050, 0.843121)$	$(0.419625, 0.648722)$	$(0.471093, 0.874260)$
y	$(-0.346121, 0.160780)$	$(-0.047271, 0.253993)$	$(-0.079662, 0.193945)$	$(-0.075751, 0.453817)$
z	$(0.198351, 0.795122)$	$(0.058659, 0.619450)$	$(0.075902, 0.569353)$	$(-0.146431, 0.497202)$

We notice that 0 is does not belong to 13 simultaneous confidence intervals in the table above. We then can conclude that there is significant mean VW projective shape change between the two flowers. This difference is also visible with the figure of the boxes of the bootstrap projective shape marginals found in Figure 5. The bootstrap projective shape marginals for landmarks 11 and 12 we can also visually reinforce our choice of rejection of the null hypothesis.

5.2 Results for cross-validation of the mean projective shape of the lily flower in second sample of images

One can show that, as expected, there is no mean VW projective shape change, based on the two samples with sample sizes respectively $n_1 = 5$ and $n_2 = 6$. In the tables below, 0 is contained in all of the simultaneous intervals. Hence, we fail to reject the null hypothesis at level $\alpha = 0.05$.

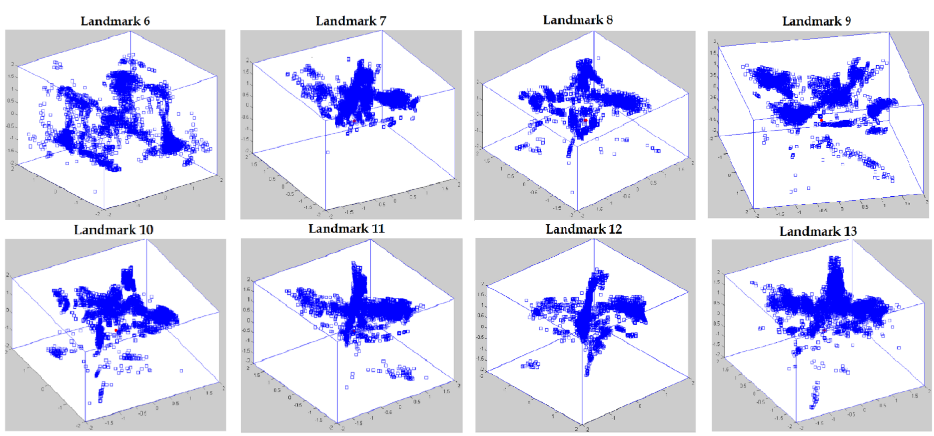

Figure 6: Bootstrap Projective Shape Marginals for Cross Validation of lily Flower

Simultaneous Confidence Intervals for lily's landmarks 6 to 9				
	LM6	LM7	LM8	LM9
x	$(-1.150441, 0.940686)$	$(-1.014147, 1.019635)$	$(-0.960972, 1.142165)$	$(-1.104360, 1.162658)$
y	$(-1.245585, 2.965492)$	$(-1.418121, 1.145503)$	$(-1.250429, 1.300157)$	$(-1.078833, 1.282883)$
z	$(-0.971271, 1.232609)$	$(-1.654594, 1.400703)$	$(-1.464506, 1.318222)$	$(-1.649496, 1.396918)$

Simultaneous Confidence Intervals for lily's landmarks 10 to 13				
	LM10	LM11	LM12	LM13
x	$(-1.078765, 1.039589)$	$(-0.995622, 1.381674)$	$(-0.739663, 1.269416)$	$(-1.015220, 1.132021)$
y	$(-1.126703, 1.140513)$	$(-1.210271, 1.184141)$	$(-1.324111, 1.026571)$	$(-1.650026, 1.593305)$
z	$(-1.092425, 1.795890)$	$(-1.222856, 1.963960)$	$(-1.128044, 1.762559)$	$(-1.035796, 2.227439)$

5.3 Comparing the sample antimean for the two lily flowers

The Veronese-Whitney (VW) antimean is the extrinsic antimean associated with the VW embedding (see Patrangenaru et al. (2010, 2014) for details). The VW antimean changes were considered for comparing the VW antimean shapes of the two flowers. Suppose Q_1 and Q_2 are independent r.o.'s, the hypothesis for their mean change are

$$H_0 : \alpha\mu_{1,j_8}^{-1} \odot_8 \alpha\mu_{2,j_8} = 1_{(\mathbb{R}P^3)^8}$$

Let φ be the affine chart on this product of projective spaces, $\varphi(1_8) = 0_8$, we compute the bootstrap distribution,

$$\alpha D_* = \varphi(\overline{aY}^{*-1}{}_{1,j_8,11} \odot_8 \overline{aY}^{*}{}_{2,j_8,8})$$

and construct the 95% simultaneous nonparametric bootstrap confidence intervals. We will then expect to fail to reject the null, if we have 0 in all of our simultaneous confidence intervals.

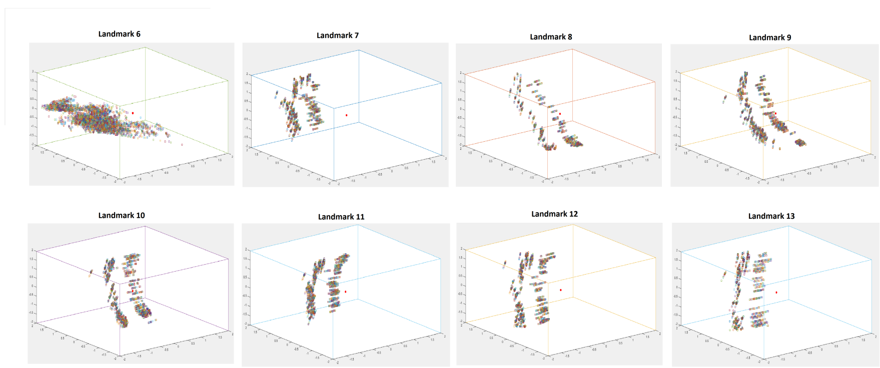

Figure 7: Eight bootstrap projective shape marginals for antimean of lily data

Highlighted in blue are the intervals not containg $0 \in \mathbb{R}$.

simultaneous confidence intervals for lily's landmarks 6 to 9				
	LM6	LM7	LM8	LM9
x	$(-1.02, -0.51)$	$(-1.41, 0.69)$	$(-1.14, 0.40)$	$(-0.87, 0.35)$
y	$(0.82, 2.18)$	$(0.00, 0.96)$	$(-0.15, 0.92)$	$(-0.09, 0.69)$
z	$(-0.75, 0.36)$	$(-6.93, 2.83)$	$(-3.07, 3.23)$	$(-2.45, 2.38)$

Simultaneous confidence intervals for lily's landmarks 10 to 13				
	LM10	LM11	LM12	LM13
x	$(-0.61, 0.32)$	$(-0.87, 0.08)$	$(-0.99, 0.02)$	$(-0.84, -0.04)$
y	$(-0.07, 0.51)$	$(-0.04, 0.59)$	$(0.06, 0.75)$	$(0.18, 0.78)$
z	$(-3.03, 1.91)$	$(-5.42, 1.98)$	$(-7.22, 2.41)$	$(-4.91, 2.62)$

In conclusion there is significant antimean VW projective shape change between the two flowers, showing that the extrinsic antimean is a sensitive parameter for extrinsic analysis.

6 Computational example for VW sample mean and VW sample antimean on a planar Kendall shape space

We use the VW-embedding of the complex projective space (Kendall shape space) to compare VW means and VW antimeans for a configuration of landmarks on midfave in a population of normal children, based on a study on growth measured from X-rays at 8 and 14 years of age(for data sets, see Patrangenaru and Ellingson(2015, Chapter 1)). The figure 8 is from lateral X-ray of a clinically normal skull (top,with landmarks). The figure 9 is the extrinsic sample mean of the coordinates of landmarks. You may find that with only a rotation, figure 8 and figure 9 looks very similar, as the extrinsic mean, is close to the sample observations. Here close is in the sense of small distance relative to the diameter of the object space.

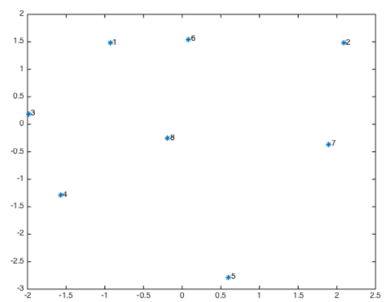

Figure 8: The coordinates of first child's skull image

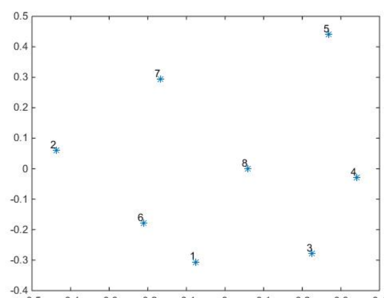

Figure 9: Icon of extrinsic sample mean- coordinates based on children midface skull data

On the other hand, we also have a sample VW- antimean, the representative of which is shown in figure 10. The VW-antimean statistic is far t from the average, since according to the general results presented in this paper, the chord distance between the sample VW antimean and sample mean in the ambient spaces is maximized. The relative location of the landmarks is also different in antimean. The follow result gives the coordinate of representatives (icons) of the VW mean and VW antimean Kendall shapes. Each coordinate of an icon is a complex number.

VW - sample mean $\bar{X}_E = (-0.0766+0.3066i, -0.4368-0.0593i, 0.2254+0.2786i, 0.3401+0.0298i, 0.2685-0.4409i, -0.2110+0.1791i, -0.1676-0.2939i, 0.0580+0.0000i)$.

VW - sample antimean $a\bar{X}_E = (0.0752-0.4103i, 0.0066-0.4731i, -0.1244+0.0031i, 0.1213+0.1102i, -0.1015-0.0422i, -0.0400+0.5639i, -0.2553+0.2485i, 0.3182+0.0000i)$.

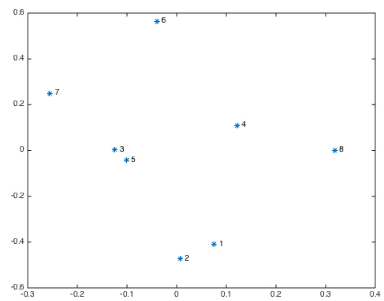

Figure 10: Icon of extrinsic sample antimean - coordinates based on children midface skull data

7 Discussion and thanks

In this paper we introduce a new statistic, the sample extrinsic antimean. Just as with the extrinsic mean, the extrinsic antimean captures important features of a distribution on a compact object space. Certainly the definitions and results extend to the general case of arbitrary Fréchet antimeans, however based on the comparison between intrinsic and extrinsic sample means (see Bhattacharya et al. (2012)), for the purpose of object data analysis (see Patrangenaru and Ellingson(2015)), it is expected that intrinsic sample antimeans take way more time to compute than extrinsic sample means. Therefore future research will parallel research on inference for extrinsic means. This includes results for stickiness of extrinsic means (see Hotz et al.(2013)). The authors would like thank Harry Hendricks and Mingfei Qiu for useful conversations on the subject of the stickiness phenomenon and antimeans and to the referee for useful comments that helped us improve the paper.

References

[1] R. Beran and N.I. Fisher (1998). Nonparametric comparison of mean axes, *Ann. Statist.*, **26**, 472–493.

[2] R. N. Bhattacharya, L. Ellingson, X. Liu and V. Patrangenaru and M. Crane (2012). Extrinsic Analysis on Manifolds is Computationally Faster than Intrinsic Analysis, with Applications to Quality Control by Machine Vision. *Applied Stochastic Models in Business and Industry.* **28**, 222–235.

[3] R.N. Bhattacharya and V. Patrangenaru (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds-Part II, *Ann. Statist.* **33**, 1211– 1245.

[4] R.N. Bhattacharya and V. Patrangenaru (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds-Part I, *Ann. Statist.* **31**, no. 1, 1-29.

[5] L. J. Billera, S. P. Holmes and K. Vogtmann, (2001). Geometry of the space of phylogenetic trees. *Adv. in Appl. Math.* **27** , no. 4, 733767.

[6] M. Crane and V. Patrangenaru. (2011). Random Change on a Lie Group and Mean Glaucomatous Projective Shape Change Detection From Stereo Pair Images. *Journal of Multivariate Analysis*. **102**, 225-237.

[7] I. L. Dryden and K. V. Mardia (1998). *Statistical Shape Analysis*, Wiley, Chichester.

[8] B. Efron (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics, **38**. SIAM, Philadelphia, Pa.

[9] N. I. Fisher, P. Hall, B. Y. Jing and A. T. A. Wood (1996). Properties of principal component methods for functional and longitudinal data analysis. *J. Amer. Statist. Assoc.* **91**, 1062–1070.

[10] M. Fréchet (1948). Les élements aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré* **10**, 215–310.

[11] R. Guo, V. Patrangenaru and D. Lester (2014) Nonparametric Bootstrap test for Equality of Extrinsic Mean Reflection Shapes in Large vs Small Populations of Acrosterigma Magnum Shells, poster, *Geometric Topological and Graphical Model Methods in Statistics Fields Institute, Toronto, Canada, May 22-23, 2014*

[12] S. Helgason (2001) *Differential Geometry and Symmetric Spaces*, AMS Chelsea Publishing, AMS, Providence, Rhode Island.

[13] T. Hotz, S. Huckemann, H. Le, J. S. Marron, J. C. Mattingly, E. Miller, J. Nolen, M. Owen, V. Patrangenaru and S. Skwerer (2013). Sticky Central Limit Theorems on Open Books. *Annals of Applied Probability,* **23**, 2238–2258.

[14] D. G. Kendall (1984). Shape manifolds, Procrustean metrics, and complex projective spaces, *Bull. London Math. Soc.*, **16**, 81–121.

[15] J. T. Kent (1992).New directions in shape analysis, *The Art of Statistical Science, A Tribute to G. S. Watson*, 115–127

[16] Y. Ma and A. Soatto and J. Košecká and S. Sastry (2005). *An Invitation to 3-D Vision: From Images to Geometric Models*, Springer.

[17] K. V. Mardia and V. Patrangenaru (2005). Directions and projective shapes. *The Annals of Statistics*, **33**, 1666-1699.

[18] V. Patrangenaru and L.L. Ellingson (2015). *Nonparametric Statistics on Manifolds and their Applications*. Chapman & Hall/CRC Texts in Statistical Science.

[19] V. Patrangenaru, R. Guo and K. D. Yao (2016). Nonparametric Inference for Location Parameters via Fréchet functions. *Proceedings of Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, Beer Sheva, February 15-18.* To appear.

[20] V. Patrangenaru, M. Qiu and M. Buibas (2014). Two Sample Tests for Mean 3D Projective Shapes from Digital Camera Images. *Methodology and Computing in Applied Probability.* **16**, 485–506.

[21] V. Patrangenaru, X. Liu and S. Sugathadasa (2010). Nonparametric 3D Projective Shape Estimation from Pairs of 2D Images - I, In Memory of W.P. Dayawansa. *Journal of Multivariate Analysis*. **101**, 11-31.

[22] V. Patrangenaru and K. V. Mardia (2003). Affine Shape Analysis and Image Analysis, *Proceedings of the Leeds Annual Statistics Research Workshop*, Leeds University Press, 57–62.

[23] S. M. Sughatadasa (2006). *Affine and Projective Shape Analysis with Applications*, Ph.D. Dissertation, Texas Tech University.

[24] H. Wang and J. S. Marron (2007). Object oriented data analysis: Sets of trees. *The Annals of Statistics*, **35**, 1849-1873.

[25] G. S. Watson (1983). *Statistics on Spheres*, Lecture Notes in the Mathematical Sciences, John Wiley & Sons.

[26] H. Ziezold (1977). On expected figures and a strong law of large numbers for random elements in quasi-metric spaces. *Transactions of Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes A.* 591–602.