On the Meaning of Mean Shape

Stephan Huckemann

Abstract

A stability result for intrinsic means on the quotient due to an isometric and proper Lie group action on a Riemannian manifold is derived, stating that intrinsic means are contained in the highest dimensional manifold stratum assumed with non-zero probability. In consequence, the Central Limit Theorem (CLT) for manifold valued random elements can be extended to non-manifold shape spaces. The relationship to other types of means is disussed: the CLT extends to Ziezold means but not in general to Procrustes means, since the latter may be contained in lower dimensional manifold strata. In contrast, Schoenberg means tend to increase dimension. The finite power of tests may increase when switching from intrinsic tangent space coordinates to residual tangent space coordinates. This and computational speed makes Ziezold means most competitive for multi-sample test. Curiously, intrinsic, Ziezold, and Procrustes means are related in such that in approximation the Ziezold means divides the generalized geodesic segment between intrinsic and Procrustes mean by the ratio 1 : 3. Classical data-sets as well as numerical simulations illustrate the findings.

Key words and phrases: Shape Spaces, Intrinsic Mean, Extrinsic Mean, Mean Location, Stability, General Procrustes Analysis, Full Procrustes Mean, Isometric Group Action, Schoenberg Mean, Power of Tests

Figure *AMS 2000 Subject Classification:* Primary 60D05
Secondary 62H11

1 Introduction

The analysis of shape may be counted among the very early activities of mankind; be it for representation on cultural artefacts, or for morphological, biological and medical application, to name only two of its driving forces. In modern days shape analysis is gaining increased momentum in computer vision, image analysis, biomedicine and many other fields. For a fairly recent overview over key aspects cf. Krim and Yezzi (2006). To begin with, already the concept of an expected, averaged or *mean* shape is surprisingly non trivial and not at all canonical. This is due to the fact that spaces of shapes usually admit several natural structures, and, in many cases none of them is that of a Euclidean space. While some structures are Euclidean (e.g. Mosimann (1970), Hobolth et al. (2002), Hotz et al. (2010)), structures that may be considered benign are that of Riemanian manifolds (e.g. Blum and Nagel (1978), Kendall (1984)). Many shape space, however, admit only the structure of a Riemannian manifold – the *pre-shape space* of configurations – modulo the isometric action of a Lie group, conveying shape equivalence. We call the quotient a *shape space*. It carries a canonical quotient structure of a union of manifold strata of different dimensions, which give in general a Riemannian manifold – the *manifold part* – with singularities at some of which the curvature may tend to infinity. Prominent example are Kendall's three- and higher-dimensional shape spaces, cf. Kendall et al. (1999), and the spaces of closed planar curves, see Zahn and Roskies (1972).

In a Euclidean space, there is a clear and unique concept of a mean in terms of least squares minimization: the arithmetic avarage. Generalizing to manifolds, however, with every embedding in a Euclidean space come specific concepts of *extrinsic means* and *residual means*, and with every manifold structure a specific concept of an *intrinsic mean*. Carrying statistics over to manifolds, strong consistency and central limit theorems (CLTs) for extrinsic, residual and intrinsic means have been derived by Jupp (1988), Hendriks and Landsman (1996, 1998) as well as by Bhattacharya and Patrangenaru (2003, 2005). Under quotienting these means generalize to *Ziezold means* and *Procrustes means*, respectively, and again, to *intrinsic means*. For a CLT to hold, obviously, a manifold structure is inevitable. Due to strong consistency, which has been established by Ziezold (1977) on general quasi-metrical spaces, for a *one-sample test* for a specific mean shape on the manifold part, it may be assumed that sample means eventually lie on the manifold part as well, thus making the above CLT available. Multi-sample tests, however, could not be theoretically justified because it remained unclear whether means of random shapes on the manifold part come to lie on the manifold part again.

As the key result in this paper it is shown that intrinsic and Ziezold means lie on the manifold part whenever the manifold part is assumed with non-zero probability. In consequence the CLT for manifolds can be applied in general, thus making multi-sample tests for these mean shapes possible. We note that specifically for the non-manifold Kendall reflection shape spaces, *Schoenberg means* as have been recently introduced by Bandulasiri and Patrangenaru (2005) as well as by Dryden et al. (2008) also allow for asymptotic inference.

This paper is structured as follows. In motivation in Section 2, Kendall's shape spaces are introduced along with the specific application of the key result. The following Section 3 lays out the various concepts of means in a generic shape space setting. In Section 4 the stability theorem, namely that

intrinsic means preserve dimension

is established. The proof is based on lifting a distribution on the shape space to the pre-shape space and subsequently exploiting the fact that intrinsic means are zeroes of an integral involving the Riemann exponential. The similar argument can be applied to Ziezold means but not to Procrustes means. The population case is more intrigued, however, as the lifting requires the concept of tubular neighborhoods admitting slices. In Section 5, extrinsic means are discussed, in particular that Schoenberg means are due to a non-isometric embedding and that they tend to "increase the dimension". Section 6 tackles local effects of curvature: Section 6.1 motivates why residual tangent space coordinates may yield a higher finite power than intrinsic tangent space coordinates. The closeup on spherical means in Section 6.2 reveals that, if the following means are unique, the generalized geodesic segment between the intrinsic mean and the Procrustes mean (closer to the data) is in approximation divided by the Ziezold mean by the ration $1 : 3$. Section 7 illustrates the practical effects of the theoretical results using classical data-sets as well as simulations.

2 Motivation: Kendall's Shape Spaces

In the statistical analysis of similarity shapes based on landmark configurations, geometrical m-dimensional objects (usually $m = 2, 3$) are studied by placing $k > m$ *landmarks* at specific locations of each object. Each object is then described by a matrix in the space $M(m, k)$ of $m \times k$ matrices, each of the k columns denoting an m-dimensional landmark vector. $\langle x, y\rangle := \operatorname{tr}(xy^T)$ denotes the usual inner product with norm $\|x\| = \sqrt{\langle x, x\rangle}$. For convenience and without loss of generality for the considerations below, only *centered* configurations are considered. Centering can be achieved by multiplying with a sub-Helmert matrix $\mathcal{H} \in M(k, k-1)$ from the right, yielding a configuration $x\mathcal{H}$ in $M(m, k-1)$. For this and other centering methods cf. Dryden and Mardia (1998, Chapter 2). Excluding also all configurations with all landmarks coinciding gives the space of *configurations*

$$F_m^k \quad := \quad M(m, k-1) \setminus \{0\}\,.$$

Since only the similarity shape is of concern, we may assume that all configurations are contained in the unit sphere $S_m^k := \{x \in M(m, k-1) : \|x\| = 1\}$. Then, *Kendall's shape space* is the canonical quotient

$$\Sigma_m^k := S_m^k / SO(m) = \{[x] : x \in S_m^k\} \text{ with the } \textit{fiber } [x] = \{gx : g \in SO(m)\}\,.$$

In some applications reflections are also filtered out giving *Kendall's reflection shape space*

$$R\Sigma_m^k := \Sigma_m^k / \{e, \widetilde{e}\} = S_m^k / O(m)\,.$$

Here, $O(m) = \{g \in M(m, m) : g^T g = e\}$ denotes the orthogonal group with the unit matrix $e = \operatorname{diag}(1, \ldots, 1)$, $\widetilde{e} = \operatorname{diag}(-1, 1, \ldots, 1)$ and $SO(m) = \{g \in O(m) : \det(g) = 1\}$ is the special orthogonal group.

For $1 \le j \le m < k$ consider the (non-unique) embedding

$$S_j^k \quad \hookrightarrow \quad S_m^k \quad : \quad x \quad \mapsto \quad \left(\frac{x}{0}\right) \tag{1}$$

giving rise to a unique and canonical embedding of $\Sigma_j^k \hookrightarrow \Sigma_m^k$ which is isometric w.r.t. the canonical intrinsic distance, the Procrustes distance and the Ziezold distance defined in Section 3, cf. Kendall et al. (1999, pp. 29, 206). In consequence, intrinsic means, Procrustes means and Ziezold means (also defined in Section 3) of random elements taking values in $\Sigma_j^k \subset \Sigma_m^k$, are also contained within Σ_j^k:

Remark 2.1. *Suppose that $[X]$ is a random shape on Σ_m^k assuming shapes of configurations of dimension less than or equal to j $(1 \leq j \leq m)$ with probability one, then every intrinsic, Procrustes and Ziezold mean shape of $[X]$ corresponds to a configuration of dimension less than or equal to j.*

The main result of this paper in form Corollaries 4.4 and 4.9 applied to Kendall's spaces states that intrinsic and Ziezold means are also non decreasing in dimension. In view of Ziezold means, Remark 3.4 provides invariant optimal positioning and (3) on page 8 is valid on the pre-shape sphere. The technical definition of a cut locus on the quotient and a tubular neighborhood admitting a slice can be found in Section 4.1.

Theorem 2.2 (Intrinsic and Ziezold Kendall Means Preserve Dimension)**.** *Suppose that $[X]$ is a random pre-shape on Σ_m^k, $m > k$ with intrinsic or Ziezold mean shape $[\mu] \in \Sigma_m^k$, $\mu \in S_m^k$. Moreover suppose that $[X]$ either assumes discrete values outside the cut locus of $[\mu]$ a.s. or that it is supported by a projection of a tubular neighborhood around μ admitting a slice. If $[X]$ assumes shapes corresponding to non-degenerate configurations up to dimenson j $(1 \leq j \leq m)$ with non-zero probability then μ corresponds to a non-degenerate j-dimensional configuration.*

3 Fréchet ρ-means and General Shape Spaces

In the previous section we introduced Kendall's *shape* and *reflection shape space* based on invariance under similarity transformations; e.g. invariance under congruence transformations only leads to Kendall's *size-and-shape space*. More generally in image analysis, invariance may also be considered under the affine or projective group, cf. Mardia and Patrangenaru (2001, 2005); Munk et al. (2008). A different yet also very popular popular set of shape spaces for two-dimensional configurations modulo the group of similarites has been introduced by Zahn and Roskies (1972). Instead of building on a finite dimensional Euclidean matrix space modeling landmarks, the basic ingredient of these spaces modeling closed planar unit speed curves is the infinite dimensional Hilbert space of Fourier series. In practice for numerical computation, only finitely many Fourier coefficients are considered.

To start with, a shape space is a metric space (Q, d). For this entire paper suppose that $X, X_1, X_2, \ldots$ are i.i.d. random elements mapping from an abstract probability space $(\Omega, \mathcal{A}, \mathcal{P})$ to (Q, d) equipped with its self understood Borel σ-field. Moreover, denote by $\mathbb{E}(Y)$ the classical expected value of a random element Y on a D-dimensional Euclidean space $\mathbb{R}^D$, if existent.

Definition 3.1. *For a continuous function* $\rho : Q \times Q \to [0, \infty)$ *define the* set of population Fréchet ρ-means *by*

$$E^{(\rho)}(X) = \operatorname{argmin}_{\mu \in Q} \mathbb{E}\big(\rho(X, \mu)^2\big) .$$

For $\omega \in \Omega$ *denote by*

$$E_n^{(\rho)}(\omega) = \operatorname{argmin}_{\mu \in Q} \sum_{j=1}^{n} \rho\big(X_j(\omega), \mu\big)^2$$

the set of sample Fréchet ρ-means.

By continuity of ρ, the mean sets are closed random sets. For our purpose here, we rely on the definition of *random closed sets* as introduced and studied by Choquet (1954), Kendall (1974) and Matheron (1975). Since their original definition for $\rho = d$ by Fréchet (1948) such means have found much interest.

Intrinsic means. Independently, for a Riemannian manifold with geodesic distance $d = \rho$, Kobayashi and Nomizu (1969) defined the corresponding means as *centers of gravity* which are nowadays also well known as *intrinsic means* (Bhattacharya and Patrangenaru (2003, 2005)).

Extrinsic means. W.r.t. the chordal or *extrinsic metric* $d = \rho$ due to an embedding of a Riemannian manifold in an ambient Euclidean space, such means have been called *mean locations* by Hendriks (1991) or *extrinsic means* by Bhattacharya and Patrangenaru (2003).

More precisely, let $Q = M \subset \mathbb{R}^D$ be a Riemannian manifold embedded in a Euclidean space $\mathbb{R}^D$ and let $\Phi : R^D \to M$ denote the orthogonal projection, $\Phi(x) = \inf_{p \in M} \|x - p\|$. For any Riemannian manifold an embedding that is even isometric can be found for D sufficiently large, see Nash (1956). Due to an extension of Sard's Theorem by Bhattacharya and Patrangenaru (2003, p.12) for a closed manifold, Φ is uni-valent up to a set of Lebesgue measure zero. Then the set of extrinsic means is given by the set of images $\Phi\big(\mathbb{E}(Y)\big)$ where Y denotes X viewed as taking values in $\mathbb{R}^D$ (cf. Bhattacharya and Patrangenaru (2003)).

Residual means. In this context, setting $\rho(p, q) = \|d\Phi_q(p - q)\|$ $(p, q \in M)$ with the derivative $d\Phi_q$ at q yielding the orthogonal projection to the embedded tangent space $T_q\mathbb{R}^D \to T_q M \subset T_q\mathbb{R}^D$, call the corresponding mean sets $E^{(\rho)}(X)$ and $E_n^{(\rho)}(\omega)$, the sets of *residual population means* and *residual sample means*, respectively. For two-spheres, $\rho(p, q)$ has been studied under the name of *crude residuals* by Jupp (1988). In general, on unit-spheres

$$\rho(p, q) = \|p - \langle p, q\rangle q\| = \sqrt{1 - \langle p, q\rangle^2} = \rho(q, p) \tag{2}$$

is a quasi-metric (symmetric, vanishing on the diagonal $p = q$ and satisfying the triangle inequality).

Due to curvature these concepts of means are essentially non-convex in the following sense.

Remark 3.2. *To a random point on a two-sphere, uniformly distributed along its equator, both north and south pole are intrinsic and extrinsic means.*

Residual means, however, seem less affected.

Theorem 3.3. *If a random point X on a unit sphere is a.s. contained in a unit subsphere S then S contains as well every residual mean of X.*

Proof. Suppose that $x = v + \nu$ is a residual mean of X with $v/\|v\| \in S$ and ν normal to S. Since $1 - \langle X, v+\nu\rangle^2 = 1 - \langle X, v\rangle^2 \geq 1 - \langle X, v\rangle^2/\|v\|^2$ a.s. with equality if and only if $\nu = 0$, the assertion follows at once from (2). $\square$

Let us now incorporate more of the structure common to shape spaces: M is a complete connected D-dimensional Riemannian manifold with geodesic distance d_M on which a Lie group G acts properly and isometrically from the left. Then we call the canonical quotient

$$\pi : M \to Q := M/G = \{[p] : p \in M\} \text{ where } [p] = \{gp : g \in G\},$$

a *shape space*. As a consequence of the isometric action we have that $d_M(gp, q) = d(p, g^{-1}q)$ for all $p, q \in M,\ g \in G$. For $p, q \in M$ we say that p is in *optimal position* to q if $d_M(p, q) = \min_{g\in G} d_M(gp, q)$, the minimum is attained in consequence of the proper action. As is well known (e.g. Bredon (1972, p. 179)) there is an open and dense submanifold M^* of M such that the canonical quotient $Q^* = M^*/G$ restricted to M^* carries a natural manifold structure also being open and dense in Q. Elements in M^* and Q^*, respectively, are called *regular*, the complementary elements are *singular*; Q^* is the *manifold part*.

Intrinsic means on shape spaces. The canonical distance

$$d_Q([p],[q]) := \min_{g\in G} d_M(gp, q) = \min_{g,h\in G} d_M(gp, hq)$$

is called *intrinsic distance* and the corresponding d_Q-Fréchet mean sets are called *intrinsic means*. Note that the intrinsic distance is equal to the canonical geodesic distance on Q^*.

Ziezold and Procrustes means. Now, assume that we have an embedding with orthogonal projection $\Phi : R^D \to M \subset \mathbb{R}^D$ as above. If the action of G is isometric w.r.t. the extrinsic metric, i.e. if $\|gp - gq\| = \|p - q\|$ for all $p, q \in M$ and $g \in G$ then call

$$\begin{aligned} \rho_Q^{(z)}([p],[q]) &:= \min_{g\in G} \|gp - q\| \\ \rho_Q^{(p)}([p],[q]) &:= \min_{g\in G} \|d\Phi_q(gp - q)\| \end{aligned}$$

the *Ziezold distance* and the *Procrustes distance* on Q, respectively. The corresponding Fréchet mean sets are the *Ziezold means* and the *Procrustes means*, respectively. We say that *optimal positioning is invariant* if

$$d_M(g^*p,q) = \min_{g\in G} d_M(gp,q) \Leftrightarrow \|g^*p - q\| = \min_{g\in G} \|gp - q\|$$

for all $p, q \in M$ and $g^* \in G$.

Remark 3.4. *Indeed for $Q = \Sigma_m^k$, optimal positioning is invariant (cf. Kendall et al. (1999, p. 206)), Procrustes means coincide with classical Procrustes means introduced by Gower (1975) and Ziezold means coincide with means as introduced by Ziezold (1994). Moreover for $Q = \Sigma_2^k$, Procrustes means agree with extrinsic means w.r.t. the Veronese-Whitney embedding, cf. Bhattacharya and Patrangenaru (2003) and Section 5.*

4 Stability Results for Intrinsic Means

In this section we derive the stability results underlying Theorem 2.2. To this end we need to recall how a non-manifold shape space is made up from manifold strata of varying dimensions.

4.1 Preliminaries

As before, let M be a complete connected Riemannian manifold called the *pre-shape space* with intrinsic metric d_M on which a Lie group G acts properly and isometrically giving rise to the canonical quotient $Q = M/G$ called the *shape space* with intrinsic metric d_Q. The canonical projection is denoted by $\pi : M \to Q$. T_pM is the tangent space of M at $p \in M$ and $\exp_p$ denotes the *Riemannian exponential* at p. On a complete and connected manifold M we have for every $p' \in M$ that there is $v' \in T_pM$ such that $p' = \exp_p v'$. $v' = \exp_p^{-1} p'$ of minimal modulus is uniquely determined as long as $p \in M\setminus C(p)$ with the *cut locus* $C(p)$ of p which is a subset of measure zero in M. E.g. on a sphere, the cut locus of any point is its antipode. For $q \in Q$ let

$$C(q) := \{[p'] : p' \in C(p) \text{ is in optimal position to some } p \in q\}$$

be the *cut locus on the quotient* of $q \in Q$.

Next we recollect consequences of the Lie group action, see Bredon (1972). The stability result in the sample case (below in Section 4.2) rests on (A) and (B) alone.

(A) With the *isotropy group* $I_p = \{g \in G : gp = p\}$ for $p \in M$, every orbit carries the natural structure of a coset space $[p] \cong G/I_p$. Moreover, $p' \in M$ is of *orbit type* (G/I_p) if $I_{p'} = gI_pg^{-1} = I_{gp}$ for a suitable $g \in G$. If $I_p \subset I_{gp'}$ for suitable $g \in G$ then p' is of *lower orbit type* than p.

(B) The pre-shapes of equal orbit type

$$M^{(I_p)} := \{p' \in M : \ p' \text{ is of orbit type } (G/I_p)\}$$

and the corresponding shapes $Q^{(I_p)} := \{[p'] : p' \in M^{(I_p)}\}$ are manifolds in M and Q, respectively.

(C) The orthogonal complement H_pM in T_pM of the tangent space $T_p[p]$ along the orbit is called the *horizontal space*: $T_pM = T_p[p] \oplus H_pM$.

(D) The *Slice Theorem* (cf. Palais (1960)) states that every $p \in M$ has a *tubular neighborhood* $[p] \subset U \subset M$ such that with a suitable open subset $D \subset H_pM$ the *twisted product* $\exp_p D \times_{I_p} [p]$ is diffeomorphic with U. Here, the twisted product is the natural topological quotient of the product space $\exp_p D \times [p]$ modulo the equivalence

$$(\exp_p v, gp) \sim_{I_p} (\exp_p v', g'p) \ \Leftrightarrow \exists h \in I_p \text{ such that } v' = dhv,\ g' = gh^{-1}.$$

For short we say that the tubular neighborhood U *admits a slice* $\exp_p D$ *through* p.

(E) If a tubular neighborhood U of p admits a slice $\exp_p D$ then every $p' \in \exp_p D$ is in optimal position to p. Moreover, all points $p' \in U$ are of orbit type larger than or equal to the orbit type of p and only finitely many orbit types occur in U. If p is regular, i.e. of maximal orbit type, then the quotient is trivial: $\exp_p D \times_{I_p} [p] = \exp_p D \times [p]$.

Finally, we link intrinsic to Ziezold means. The gradient of the mapping $f_{int} : T_pM \to [0, \infty)$ defined by $f_{int}^{(p)}(v) = d_M(p, \exp_p v)^2$ is given by $\operatorname{grad} f_{int}^{(p)}(v) = 2v$. Hence, we have for $p_1, p_2 \in M \setminus C(p)$ that

$$\operatorname{grad} f_{int}^{(p)}(\exp_p^{-1} p_1) = \operatorname{grad}(f_{int}^{(p)})(\exp_p^{-1} p_2) \ \Leftrightarrow \ p_1 = p_2$$

In view of Ziezold means let $f_{ext}^{(p)} : T_pM \to [0, \infty)$ be defined by $f_{ext}^{(p)}(v) = \|p - \exp_p v\|^2/2$. Mimicking the above property introduce the following condition

$$\operatorname{grad}(f_{ext}^{(p)})(\exp_p^{-1} p_1) = \operatorname{grad}(f_{ext}^{(p)})(\exp_p^{-1} p_2) \quad \Leftrightarrow \quad p_1 = p_2 \tag{3}$$

for $p_1, p_2 \in M \setminus C(p)$. E.g. on spheres, (3) is valid. Using the residual distance instead, the analog to (3) is only true on half-spheres.

4.2 Stability of Sample Means

For short, for sampled points $p_1, \ldots, p_n \in M$ denote the set of intrinsic or extrinsic sample means, respectively, by $E(p_1, \ldots, p_n)$; similarly for $q_1, \ldots, q_n \in Q$ denote the set of intrinsic or Ziezold sample means, respectively, by $E(q_1, \ldots, q_n)$. Obviously, both sets are non-void.

Lemma 4.1. *In case of extrinsic and Ziezold means assume that optimal positioning is invariant. Let $q_1 \dots, q_n \in Q$ and $\overline{q} \in E(q_1, \dots, q_n)$. If $\overline{p} \in \overline{q}$ and $p_1 \in q_1, \dots, p_n \in q_n$ are in optimal position to $\overline{p}$ then $\overline{p} \in E(p_1, \dots, p_n)$.*

Proof. The proof in case of extrinsic and Ziezold means is completely analogous to the following case of intrinsic means. If the assertion would not be true, we would have with $p \in E(p_1, \dots, p_n)$ that

$$\begin{aligned} \sum_{j=1}^n d_Q([p], q_j)^2 &= \sum_{j=1}^n d_M(p, g_j p_j)^2 \\ &< \sum_{j=1}^n d_M(\overline{p}, p_j)^2 = \sum_{j=1}^n d_Q(\overline{q}, q_j)^2 \end{aligned}$$

for $g_1 p_1, \dots, g_n p_n$ in optimal position to p, a contradiction to the hypothesis $\overline{q} \in E(q_1, \dots, q_n)$ □

The following necessary condition is taken from (Kobayashi and Nomizu, 1969, p. 110).

Lemma 4.2. *In case of intrinsic means let $\overline{p} \in M$ and $p_1, \dots, p_n \in M \setminus C(\overline{p})$ with $\overline{p} \in E(p_1, \dots, p_n)$ for a complete and connected manifold M. Then*

$$\sum_{j=1}^n \exp_{\overline{p}}^{-1} p_j = 0 \,.$$

For extrinsic means the analog condition is

$$\sum_{j=1}^n f_{ext}^{(\overline{p})}(\overline{p}, \exp_{\overline{p}}^{-1} p_j) = 0 \,. \tag{4}$$

Theorem 4.3. *Suppose that G is a Lie group acting properly and isometrically on a complete connected Riemannian manifold M giving rise to the natural quotient $Q = M/G$. If $\overline{q} \in E(q_1, \dots, q_n)$ is an intrinsic sample mean of $q_1, \dots, q_n \in Q \setminus C(\overline{q})$ then*

$$I_{\overline{p}} \subset \cap_{j=1}^n I_{p_j}$$

with arbitrary $\overline{p} \in \overline{q}$ and $p_1 \in q_1, \dots, p_n \in q_n$ in optimal position to $\overline{p}$. In case of invariant optimal positioning and under condition (3) the same assertion is true if $\overline{q}$ is a Ziezold sample mean.

Proof. The proof in case of Ziezold means is completely analogous to the following case of intrinsic means, using (4) instead of Lemma 4.2. If the assertion would be false, we may assume that $I_{\overline{p}} \not\subset I_{p_1}$, i.e. that there is $g \in G$ with $g\overline{p} = \overline{p}$ but $gp_1 \neq p_1$. In consequence of Lemma 4.1, $\overline{p}$ is an intrinsic mean of both $p_1, \dots, p_n$ and $gp_1, p_2, \dots, p_n$, both being in optimal position to $\overline{p}$. Since by hypothesis, $\exp_{\overline{p}}^{-1}(gp_1) \neq \exp_{\overline{p}}^{-1}(p_1)$ is well defined, we have in consequence of Lemma 4.2 the contradiction

$$\begin{aligned} 0 &= \exp_{\overline{p}}^{-1} gp_1 + \sum_{j=2}^{n} \exp_{\overline{p}} p_j \\ &= \exp_{\overline{p}}^{-1} gp_1 - \exp_{\overline{p}}^{-1} p_1 + \sum_{j=1}^{n} \exp_{\overline{p}} p_j \neq 0 \,. \end{aligned}$$

□

Corollary 4.4. *Suppose that G is a Lie group acting properly and isometrically on a complete connected Riemannian manifold M giving rise to the natural quotient $Q = M/G$. Then every intrinsic mean $\overline{q} \in Q$ of a sample $q_1, \dots, q_n \in Q \setminus C(\overline{q})$ is regular if the sample contains at least one regular point. In case of invariant optimal positioning and under condition (3) the same assertion is true for Ziezold sample means.*

Proof. The assertion is a consequence of Theorem 4.3 and the well known fact that regular points $p^* \in M^*$ are characterized by the property that for every $p \in M$ there is a $g = g_p \in G$ such that $gI_{p^*}g^{-1} \subset I_p$: if additionally p is in optimal position to p^* then $g \in I_p$. □

4.3 Stability of Population Means

Theorem 4.5. *Assume that X is a random element on the canonical quotient $Q = M/G$ due to the isometric and proper action of a Lie group G on a complete Riemannian manifold M. Let $p \in M$ and suppose that X is supported on $\pi(U)$ with a tubular neigborhood $U \in M$ of $[p]$ that admits a slice $\exp_p D \times_{I_p} [p] \cong U$, $D \in H_pM$. If $\mathcal{P}\{X \in Q^{(I_{p'})}\} \neq 0$ for some $p' \in U$ and if either $[p]$*

(i) is an intrinsic mean of X or

(ii) is a Ziezold mean of X while optimal positioning is invariant and (3) is valid,

then p' is of lower orbit type than p.

The proof of Theorem 4.5 further below relies on the two following lemmas for which we first develop an additional concept.

In the following, *measurable* will refer to the corresponding Borel σ-algebras, respectively.

Definition 4.6. *Call a measurable subset $L \subset M$ a* measurable horizontal lift *of a subset R of M/G* in optimal position to *$p \in M$ if*

1. the canonical projection $L \to R \subset M/G$ surjective,

2. every $p' \in L$ is in optimal position to p, *3. every orbit $[p']$ of $p' \in L$ meets L once.*

Lemma 4.7. *Let $U \subset M$ be a tubular neighborhood about $p \in M$ that admits a slice $\exp_p D \times_{I_p} [p] \cong U$. Then, there is a measurable horizontal lift $L \subset \exp_p D$ of $\pi(U)$.*

Proof. If p is regular, then $L = \exp_p D$ has the desired properties, cf. (E) above. Now assume that p is not of maximal orbit type. W.l.o.g. assume that D contains the closed ball B of radius $r > 0$ with bounding sphere $S = \partial B$ and that there are $p_1, \ldots, p_J \in \exp_p(S)$ having the distinct orbit types orccuring in S. S^j denotes all points on S of orbit type (G/I_{p_j}), $j = 1, \ldots, J$, respectively. Observe that each S^j is a manifold on which I_p acts properly and isometrically. Hence for every $1 \leq j \leq J$, there is a finite ($K_j < \infty$) or countable ($K_j = \infty$) sequence of tubular neighborhoods $U^j_k \subset S^j$ covering S^j, admitting trivial slices

$$\exp^{S^j}_{p^j_k} D^j_k \times \{gp^j_k : g \in I_{p^j_k}\} \cong U^j_k, \quad 1 \leq k \leq K_j\,.$$

Here, $\exp^{S^j}_{p^j_k}$ denotes the Riemann exponential of S^j. Defining a disjoint sequence

$$\begin{aligned} \widetilde{U}^j_1 &:= U^j_1 \\ \widetilde{U}^j_{k+1} &:= U^j_{k+1} \setminus \widetilde{U}^j_k \text{ for } 1 \leq k \leq K_j - 1 \end{aligned}$$

exhausting S^j we obtain a corresponding sequence of disjoint measurable sets $\widetilde{D}^j_k$ with

$$\exp^{S^j}_{p^j_k} \widetilde{D}^j_k \times \{gp^j_k : g \in I_{p^j_k}\} \cong \widetilde{U}^j_k, \quad 1 \leq k \leq K_j\,.$$

Setting

$$L^j_k := \exp^{S^j}_{p^j_k} \widetilde{D}^j_k \text{ and } L^j := \bigcup_{k=1}^{K_j} L^j_k$$

observe that every $p' \in S^j$ has a unique lift in L^j which is contained in a unique L^j_k. This lift is by construction in optimal position to p. If $gp' \in L^j_{k'}$ for some $g \in G$ and $1 \leq k' \leq K_j$ we have by the disjoint construction of $\widetilde{U}^j_k$ and $\widetilde{U}^j_{k'}$ that $k = k'$, hence the isotropy groups of gp' and p' agree, yielding $gp' = p'$. In consequence, L^j is a measurable horizontal lift of S^j in optimal position to p. Since every horizontal geodesic segment $t \mapsto \exp_p(tv)$, $v \in H_pM$ contained in $\exp_p D$ features a constant isotropy group, except possibly for the initial point we obtain with the definition of

$$L := \{\exp_p(tv) \in \exp_p D : v \in \exp_p^{-1}(L^j) \text{ for some } 1 \leq j \leq J, t \geq 0\}$$

a measurable horizontal lift of $\pi(U)$ in optimal position to p. $\square$

The following Lemma 4.8 is the analog in the population case of Lemmas 4.1 and 4.2. The proof goes similar.

Lemma 4.8. *Assume that X is a random element on the canonical quotient $Q = M/G$ due to the isometric and proper action of a Lie group G on a complete Riemannian manifold M. Moreover, assume that there is $p \in M$ with tubular neigborhood $U \in M$ of $[p]$ that admits a slice $\exp_p D \times_{I_p} [p] \cong U$ such that X is supported by $\pi(U)$. With a measurable horizontal lift L of $\pi(U)$ define the random element Y on $L \subset M$ by $\pi \circ Y = X$. If $[p]$ is an intrinsic mean of X on Q, then p is an intrinsic mean of Y on M and*

$$\mathbb{E}(\exp_p^{-1} Y) = 0\,.$$

If $[p]$ is a Ziezold mean of X on Q and optimal positioning is invariant, then p is an extrinsic mean of Y on M and

$$\mathbb{E}\big(\operatorname{grad}(f_{ext})(\exp_p^{-1} Y)\big) = 0\,.$$

Proof of Theorem 4.5. The proof in case of Ziezold means is completely analogous to the following case of intrinsic means. With the hypotheses of Theorem 4.5, suppose that L is a horizontal measurable lift of $\pi(U)$ guaranteed by Lemma 4.7 through an intrinsic mean $p \in M$ of the random element Y on M defined as in Lemma 4.8 with $[p] \in E^{d_Q}(X)$. With the notation of Lemma 4.7 and its proof, if the assertion of the Theorem would be false, w.l.o.g. there would be a $q_j \in S^j$ with $gp_j \neq p_j$ for some $g \in I_p$ and $P\{Y \in M^j\} > 0$ for some $1 \leq j \leq J$ with

$$M^j := \{\exp_p(tv) \in \exp_p(D) : v \in \exp_p^{-1}(L^j), t \geq 0\}$$

In particular, w.l.o.g. in the proof Lemma 4.7, we may choose an arbitrarily small L_j^k around p_j such that

$$\int_{M_k^j} \big(\exp_p^{-1} Y - \exp_p^{-1}(gY)\big)\, dP_Y \ \neq\ 0\,.$$

with $M_k^j := \{\exp_p(tv) \in \exp_p(D) : v \in \exp_p^{-1}(L_k^j), t \geq 0\}$. Suppose that L is obtained as in the proof of Lemma 4.7 by using L_k^j and suppose that L' is obtained from L by replacing L_k^j with gL_k^j. Then L' is also a measurable horizontal lift in optimal position to p. Moreover with the lifts Y' of X to L' and Y of X to L guaranteed by Lemma 4.8, we have that

$$\begin{aligned} 0 &= \int_L \exp_p^{-1} Y\, dP_Y - \int_{L'} \exp_p^{-1} Y'\, dP'_Y \\ &= \int_{M_k^j} \big(\exp_p^{-1} Y - \exp_p^{-1}(dgY)\big)\, dP_Y \ \neq\ 0\,, \end{aligned}$$

a contradiction to the above, completing the proof. ☐

We have at once the population analog to Corollary 4.4.

Corollary 4.9. *Suppose that a random pre-shape Y is distributed on M absolutely continuous w.r.t. Riemmanian measure and supported by a tubular neighborhood that admits a slice through $p \in M$. Let $X = \pi \circ Y$ be the corresponding random shape. Then p is regular if $[p]$ is an intrinsic mean of X, or if $[p]$ is a Ziezold mean, optimal positioning is invariant and (3) is valid.*

5 Extrinsic Means for Kendall's Shape Spaces

Let us recall the well known *Veronese-Whitney* embedding for Kendall's planar shape spaces Σ_2^k. Identify F_2^k with $\mathbb{C}^{k-1} \setminus \{0\}$ such that every landmark column corresponds to a complex number. This means in particular that $z \in \mathbb{C}^{k-1}$ is a complex row(!)-vector. With the Hermitian conjugate $a^* = (\overline{a_{kj}})$ of a complex matrix $a = (a_{jk})$ the pre-shape sphere S_2^k is identified with $\{z \in \mathbb{C}^{k-1} : zz^* = 1\}$ on which $SO(2)$ identified with $S^1 = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ acts by complex scalar multiplication. Then the well known Hopf-Fibration mapping to complex projective space gives $\Sigma_2^k = S_2^k/S^1 = \mathbb{C}P^{k-2}$. Moreover, denoting with $M(k-1,k-1,\mathbb{C})$ all complex $(k-1)\times(k-1)$ matrices, the Veronese-Whitney embedding is given by

$$\begin{array}{rcl} S_2^k/S^1 & \to & \{a \in M(k-1,k-1,\mathbb{C}) : a^* = a\} \\ [z] & \mapsto & z^*z\,. \end{array}$$

Remark 5.1. *The Procrusets metric of Σ_2^k is isometric with the Euclidean metric of $M(k-1,k-1,\mathbb{C})$ since we have $\langle z,w\rangle = \mathrm{Re}(zw^*)$ for $z,w \in S_2^k$ and hence, $d^{(p)}([z],[w]) = \sqrt{1-wz^*zw^*} = \|w^*w - z^*z\|/\sqrt{2}$.*

The idea of the Veronse-Whitney embedding can be transferred to the general case of shapes of arbitrary dimension $m \geq 2$. Even though the embedding given below is apt only for reflection shape space it can be applied to practical situations in similarity shape analysis whenever the geometrical objects considered have a common orientation. As above, the number k of landmarks is essential and will be considered fixed throughout this section; the dimension $1 \leq m < k$, however, is lost in the embedding and needs to be retrieved by projection. To this end recall the (non-unique) embedding of S_j^k in S_m^k $(1 \leq j \leq m)$ in (1) which gives rise to a unique and canonical embedding of $R\Sigma_j^m$ in $R\Sigma_m^k$. Moreover, consider the strata

$$(R\Sigma_m^k)^j := \{[x] \in R\Sigma_m^k : \mathrm{rank}(x) = j\}, \quad (\Sigma_m^k)^j := \{[x] \in \Sigma_m^k : \mathrm{rank}(x) = j\}$$

for $j = 1,\dots,m$, each of which carries a canonical manifold structure; due to the above embedding, $(R\Sigma_m^k)^j$ can and will be identified with $(R\Sigma_j^k)^j$ such that

$$R\Sigma_m^k \;=\; \bigcup_{j=1}^m (R\Sigma_j^k)^j\,.$$

At this point we note that $SO(m)$ is connected, while $O(m)$ is not; and the consequences for the respective manifold parts, i.e. points of maximal orbit type:

$$\begin{array}{rcl} (\Sigma_m^k)^* & = & (\Sigma_m^k)^{m-1} \cup (\Sigma_m^k)^m\,, \\ (R\Sigma_m^k)^* & = & (R\Sigma_m^k)^m\,. \end{array}$$

Similarly, we have a stratifiction

$$\mathcal{P} := \left\{a \in M(k-1,k-1) : a = a^T \geq 0, \mathrm{tr}(a) = 1\right\} \quad = \quad \bigcup_{j=1}^{k-1} \mathcal{P}^j$$

of a bounded flat convex manifold $\mathcal{P}$ with non-flat unbounded manifolds

$$\mathcal{P}^j := \{a \in \mathcal{P} : \mathrm{rank}(a) = j\} \ (j = 1, \ldots, k-1)\,,$$

all embedded in $M(k-1,k-1)$. The *Schoenberg map* $\mathfrak{s} : R\Sigma_m^k \to \mathcal{P}$ is then defined on each stratum by

$$\begin{array}{rcl} \mathfrak{s}|_{(R\Sigma_m^k)^j} =: \mathfrak{s}^j : (R\Sigma_m^k)^j & \to & \mathcal{P}^j \\ {[x]} & \mapsto & x^T x \end{array} .$$

For $x \in S_j^k$ recall the tangent space decomposition $T_x S_j^k = T_x[x] \oplus H_x S_j^k$ into the *vertical tangent space* along the fiber $[x]$ and its orthogonal complement the *horizontal tangent space*. For $x \in (S_j^k)^j$ identify canonically (cf. Kendall et al. (1999, p. 109)):

$$T_{[x]}(R\Sigma_j^k)^j \cong H_x S_j^k = \{w \in M(j,k-1) : \mathrm{tr}(wx^T) = 0, wx^T = xw^T\}\,.$$

Then the assertion of the following Theorem condenses results of Bandulasiri and Patrangenaru (2005), cf. also Dryden et al. (2008).

Theorem 5.2. *Each* $\mathfrak{s}^j$ *is a diffeomorphism with inverse* $(\mathfrak{s}^j)^{-1}(a) = [(\sqrt{\lambda}u^T)_1^j]$ *where* $a = u\lambda u^T$ *with* $u \in O(k-1)$, $\lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, *and* $0 = \lambda_{j+1} = \ldots = \lambda_{k-1}$ *in case of* $j < k-1$. *Here,* $(a)_1^j$ *denotes the matrix obtained from taking only the first* j *rows from* a. *For* $x \in S_j^k$ *and* $w \in H_x S_j^k \cong T_{[x]}(R\Sigma_j^k)^j$ *the derivative is given by*

$$d(\mathfrak{s}^j)_{[x]} w = x^T w + w^T x\,.$$

Remark 5.3. *In contrast to the Vernose-Whitney embedding, the Schoenberg embedding is not isometric as the example of*

$$x = \begin{pmatrix} \cos\phi & 0 \\ 0 & \sin\phi \end{pmatrix}, \quad w_1 = \begin{pmatrix} \sin\phi & 0 \\ 0 & -\cos\phi \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 & \cos\phi \\ \sin\phi & 0 \end{pmatrix},$$

teaches:

$$\|x^T w_1 + w_1^T x\| = \sqrt{2}\, 2|\cos\phi \sin\phi|, \quad \|x^T w_2 + w_2^T x\| = \sqrt{2}\,.$$

Since $\mathcal{P}$ is bounded, convex and Euclidean, the classical expectation $\mathbb{E}(X^T X) \in \mathcal{P}^j$ for some $1 \leq j \leq k-1$ of the Schoenberg image $X^T X$ of an arbitrary random reflection shape $[X] \in R\Sigma_m^k$ is well defined. Then we have at once the following relation between rank of Euclidean mean and increasing sample size.

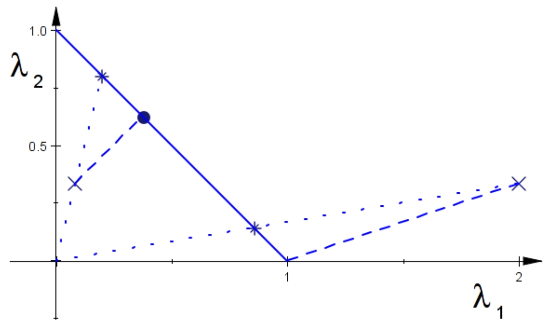

Figure 1: *Projections (if existent) of two points (crosses) in the λ-plane to the open line segment $\Lambda = \{(\lambda_1, \lambda_2) : \lambda_1 + \lambda_2 = 1, \lambda_1, \lambda_2 > 0\}$. The dotted line gives the central projections (denoted by stars) suggested by Dryden et al. (2008) which is well defined for all symmetric, positive definite matrices (corresponding to the first open quadrant), the dashed line gives the orthogonal projection (circle) which is well defined in the triangle below Λ (corresponding to $\mathcal{P}$) and above Λ in an open strip. In particular it exists not for the right point.*

Theorem 5.4. *Suppose that a random reflection shape $[X] \in R\Sigma_m^k$ is distribed absolutely continous w.r.t. the projection of the spherical measure on S_m^k. Then*

$$\mathbb{E}(X^T X) \in \mathcal{P}^{k-1} \text{ and } \frac{1}{n}\sum_{i=1}^n X_i^T X_i \in \mathcal{P}^j \text{ a.s.}$$

for every i.i.d. sample $X_1, \ldots, X_n \sim X$ if $j \leq nm < j+1$ for $nm < k-1$ and $j = k-1$ for $nm \geq k-1$.

Hence, in stastical settings involving a higher number of landmarks, a sufficiently well behaved projection of a high rank Euclidean mean onto lower rank $\mathcal{P}^r \cong (\Sigma_r^k)^r$, usually $m = r$, is to be employed, giving at once a mean shape satisfying strong consistency and a CLT. Here, unlike to intrinsic or Procrustes analysis, the dimension r chosen is crucial for the dimensionality of the mean obtained.

The orthogonal projection

$$\begin{array}{rcl} \phi^r : \bigcup_{i=r}^{k-1} \mathcal{P}^i & \to & \mathcal{P}^r \\ a & \mapsto & \operatorname{argmin}_{b \in \mathcal{P}^r} \operatorname{tr}\left((a-b)^2\right) \end{array}$$

giving the set of *extrinsic Schoenberg means* has been computed by Bhattacharya (2008):

Theorem 5.5. *For $1 \leq r \leq k-1$, $a = u\lambda u^T \in \mathcal{P}$ with $u \in O(k-1)$, $\lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$, $\lambda_1 \geq \ldots \geq \lambda_{k-1}$ and $\lambda_r > 0$ the orthogonal projection onto $\mathcal{P}^r$ is given by*

$$\phi^r(a) = u\mu u^T$$

with $\mu = \operatorname{diag}(\mu_1, \ldots, \mu_r, 0, \ldots, 0)$,

$$\mu_i = \lambda_i + \frac{1}{r} - \overline{\lambda}_r \; (i = 1, \ldots, r)$$

and $\overline{\lambda}_r = \frac{1}{r}\sum_{i=1}^r \lambda_i \leq \frac{1}{r}$ *which is uniquely determined if and only if* $\lambda_r > \lambda_{r+1}$*.*

With the notation of Theorem 5.5, a non-orthogonal *central projection* $\psi^r(a) = u\nu u^T$ equally well and uniquely determined has been proposed by Dryden et al. (2008) with $\nu = \mathrm{diag}(\nu_1, \ldots, \nu_r, 0, \ldots, 0)$,

$$\nu_i = \frac{\lambda_i}{r\overline{\lambda}_r} \quad (i = 1, \ldots, r)\,.$$

Orthogonal and central projection are depicted in Figure 5.

6 Local Effects of Curvature

In this section we assume that a manifold stratum M supporting a random element X is isometrically embedded in a Euclidean space $\mathbb{R}^D$ of dimension $D > 0$. With the orthogonal projection $\Phi : \mathbb{R}^D \to M$ from Section 3 and the Riemann exponential $\exp_p$ of M at p we have the

intrinsic tangent space coordinate	$\exp_p^{-1} X$ and the
residual tangent space coordinate	$d\Phi_p(X - p)\,,$

respectively, of X at p, if existent.

6.1 Finite Power of Tests and Tangent Space Coordinates

With the above setup, assume that $\mu \in M$ is a unique mean of X. If $c \in \mathbb{R}^D$ is the center of the osculatory circle touching the geodesic segment in M from X to μ at μ with radius r, and if X_r is the orthogonal projection of X to that circle, then $X = X_r + O(\|X - \mu\|^3)$. Moreover, with

$$\begin{aligned}\cos\alpha &= \left\langle \frac{X - c}{\|X - c\|}, \frac{\mu - c}{r} \right\rangle \\ &= \frac{1}{r^2}\langle X_r - c, \mu - c\rangle + O(\|X - \mu\|^3)\end{aligned}$$

we have the residual tangent space coordinate

$$v = X - c - \frac{\mu - c}{r}\|X - c\|\,\cos\alpha = X_r - c - (\mu - c)\cos\alpha + O(\|X - \mu\|^3)$$

having squared length $\|v\|^2 = r^2 \sin\alpha^2 + O(\|X - \mu\|^3)$. By isometry of the embedding, the intrinsic tangent space coordinate is given by

$$\exp_\mu X = \frac{r\alpha}{\|v\|}\,v + O(\|X - \mu\|^3)\,.$$

With the component

$$\nu = \mu - c - \|X - c\|\frac{\mu - c}{r}\cos\alpha = (\mu - c)(1 - \cos\alpha) + O(\|X - \mu\|^3)$$

of X normal to the mentioned geodesic segment of squared length $\|\nu\|^2 = r^2(1-\cos\alpha)^2 + O(\|X-\mu\|^3)$ we obtain

$$\|\exp_\mu X\|^2 \quad = \quad \|v\|^2 + \|\nu\|^2 + O(\|X-\mu\|^3)\,,$$

since

$$(1-\cos\alpha)^2 + \sin^2\alpha \quad = \quad 2(1-\cos\alpha) \;=\; \alpha^2 + 2\frac{\alpha^4}{4!} + \cdots$$

and $\alpha = O(\|X-\mu\|)$. In consequence we have

Remark 6.1. *In approximation, the variation of intrinsic tangent space coordinates is the sum of the variation $\|v\|^2$ of residual tangent space coordinates and the variation normal to it. In particular, for spheres*

$$\|\exp_\mu X\|^2 \;\geq\; \|v\|^2 + \|\nu\|^2\,.$$

Since the variation in normal space is irrelevant for a two-sample test for equality of means, say, a higher power for tests based on intrinsic means can be expected when solely residual tangent space coordinates obtained from an isometric embedding are used rather than intrinsic tangent space coordinates.

Note that the natural tangent space coordinates for Ziezold means are residual.

A simulated classification example in Section 7 illustrates this slight effect.

6.2 The $1:3$ - Property for Spherical and Kendall Shape Means

In this section $M = S^{D-1} \subset \mathbb{R}^D$ is the $(D-1)$-dimensional unit-hypersphere embedded isometrically in Euclidean D-dimensional space. The orthogonal projection $\Phi : \mathbb{R}^D \to S^{D-1} : p \to \frac{p}{\|p\|}$ is well defined except for the origin $p = 0$, and the normal space at $p \in S^{D-1}$ is spanned by p itself. In consequence, a random point X on S^{D-1} has

$$d\Phi_p(X-p) = X - p\cos\alpha, \quad \exp_p(X) = \frac{\alpha}{\sin\alpha}\, d\Phi_p(X-p)$$

as residual and intrinsic, resp., tangent space coordinate at $-X \neq p \in S^{D-1}$ where $\cos\alpha = \langle X, p\rangle$.

Remark 6.2. *An adaption (i.e. multiplying by the respective $\frac{\sin\alpha}{\alpha}$ in every iteration step) of the algorithm for intrinsic means in Huckemann and Ziezold (2006, Section 5.4) gives at once an algorithm converging usually quickly to one of the two sample residual means.*

Theorem 6.3. *If X is contained in a closed half sphere with non-zero probability in its interior, it has a unique intrinsic mean which is assumed in the interior of that half sphere.*

Proof. Below, we show that every intrinsic mean necessarily lies within the interior of the half sphere. Then, Kendall (1990, Theorem 7.3) yields uniqueness. W.l.o.g. let $X = (\sin\phi, x_2, \dots, x_n)$ such that $\mathcal{P}\{\sin\phi < 0\} = 0 < \mathcal{P}\{\sin\phi > 0\}$ and assume that $p = (\sin\psi, p_2, \dots, p_n) \in S^{D-1}$ is an intrinsic mean. Moreover let $p' = (\sin(|\psi|), p_2, \dots, p_n)$. Since

$$\begin{aligned}
\mathbb{E}\left(\|\exp_p(X)\|^2\right) &= \mathbb{E}\left(\arccos^2\langle p, X\rangle\right)\\
&= \mathbb{E}\left(\arccos^2\left(\sin\psi\sin\phi + \sum_{j=2}^n p_j x_j\right)\right)\\
&\leq \mathbb{E}\left(\|\exp_{p'}(X)\|^2\right)
\end{aligned}$$

with equality if and only if $|\psi| = \psi$, this can only happen for $\psi \geq 0$. Now, suppose that $p = (0, p_2, \dots, p_n)$ is an intrinsic mean. For deterministic $\psi \geq 0$ consider $p(\psi) = (\sin\psi, p_1\cos\psi, \dots, p_n\cos\psi)$. Then

$$\begin{aligned}
\mathbb{E}\left(\|\exp_{p(\psi)}(X)\|^2\right) &= \mathbb{E}\left(\arccos^2\left(\sin\psi\sin\phi + \cos\phi\sum_{j=2}^n p_j x_j\right)\right)\\
&= \mathbb{E}\left(\|\exp_p(X)\|^2\right) - C_1\psi + O(\psi^2)
\end{aligned}$$

with $C_1 > 0$ since $\mathcal{P}\{\sin\phi > 0\} > 0$. In consequence, p cannot be an intrinsic mean. Hence, we have shown that every intrinsic mean is contained in the interior of the half sphere. □

Remark 6.4. *For the special case of spheres, this extends a result of H. Le, asserting a unique intrinsic mean on general Riemannian manifolds, if among others, the support of X is contained in a geodesic quarter-ball, see Le (2001, 2004). H. Le's result is an extension of Karcher (1977), asserting uniqueness of intrinsic means in geodesic balls if among others, X is supported in a geodesic quarter ball. Intrinsic means restricted to geodesic balls are often called* Karcher means. *As mentioned in the above proof, Kendall (1990, Theorem 7.3) extended Karcher's result to obtain uniquenes of Karcher means, if among others, the support of X is contained in a half ball. In view of Theorem 6.3, one might expect to extend Le's result on uniquenes of intrinsic means on the entire manifold accordingly, to random elements contained in a geodesic half ball.*

The following Theorem characterizes the three spherical means.

Theorem 6.5. *Let X be a random point on S^{D-1}. Then $x^{(e)} \in S^{D-1}$ is the unique extrinsic mean if and only if the Euclidean mean $\mathbb{E}(X) = \int_{S^{D-1}} X\,dP_X$ is non-zero. If the following right hand sides are non-zero, then there are suitable $\lambda^{(e)} = \|\mathbb{E}(X)\| > 0$, $\lambda^{(r)} > 0$ and $\lambda^{(i)} > 0$ such that*

$$\lambda^{(e)} x^{(e)} = \mathbb{E}(X)\,,$$

every residual mean $x^{(r)} \in S^{D-1}$ satisfies

$$\lambda^{(r)} x^{(r)} = \mathbb{E}\big(\langle X, x^{(r)}\rangle\, X\big)\,,$$

and every intrinsic mean $x^{(i)} \in S^{D-1}$ satisfies

$$\lambda^{(i)} x^{(i)} = \mathbb{E}\left(\frac{\arccos\langle X, x^{(i)}\rangle}{\sqrt{1-\langle X, x^{(i)}\rangle^2}}\, X\right) .$$

In the last case we additionally require that $\mathbb{E}\left(\frac{\arccos\langle X,x^{(i)}\rangle}{\sqrt{1-\langle X,x^{(i)}\rangle^2}}\,\langle X, x^{(i)}\rangle\right) > 0$ *which is in particular the case if X is contained in a closed half sphere with non-zero probability in the interior.*

Proof. The assertions for the extrinsic mean are well known from Hendriks et al. (1996). The second assertion for residual means follows from minimization of

$$\int_{S^{D-1}} \|p - \langle p, x\rangle x\|^2 \, dP_X(p) = 1 - \int_{S^{D-1}} \langle p, x\rangle^2 \, dP_X(p)$$

with respect to $x \in \mathbb{R}^D$ under the constraining condition $\|x\| = 1$. Using a Lagrange ansatz this leads to the necessary condition

$$\int_{S^{D-1}} \langle p, x\rangle \, p \, dP_X(p) = \lambda x$$

with a Lagrange multiplier λ of value $\mathbb{E}(\langle X, x\rangle^2)$ which is positive unless X is supported by the hypersphere orthogonal to x. In that case, by Theorem 3.3, x cannot be a residual mean of X, as every residual mean is as well contained in that hypersphere. Hence, we have $\lambda > 0$. The same method applied to

$$\int_{S^{D-1}} \|\exp_x^{-1}(p)\|^2 \, dP_X(p) = \int_{S^{D-1}} \arccos^2(\langle p, x\rangle) \, dP_X(p)$$

taking into account Theorem 6.3 yields the third assertion on the intrisic mean. $\square$

In particular, residual means are eigenvectors to the largest eigenvalue of the matrix $\mathbb{E}(XX^T)$. As such, they rather reflect the mode then the classical mean of a distribution:

Example 6.6. *Consider $\gamma \in (0, \pi)$ and a random variable X on the unit circle $\{e^{i\theta} : \theta \in [0, 2\pi)\}$ which takes the value 1 with probability $2/3$ and $e^{i\gamma}$ with probability $1/3$. Then, explicit computation gives the unique intrinsic and extrinsic mean as well as the two residual means*

$$x^{(i)} = e^{i\frac{\gamma}{3}}, \quad x^{(e)} = e^{i \arctan \frac{\sin\gamma}{2+\cos\gamma}}, \quad x^{(r)} = \pm\, e^{i\frac{1}{2}\, \arctan \frac{\sin(2\gamma)}{2+\cos(2\gamma)}} .$$

Figure 2 shows the case $\gamma = \frac{\pi}{2}$.

In contrast to Figure 2, one may assume in many practical applications that the mutual proximities of the unique intrinsic mean $x^{(i)}$, the unique extrinsic mean $x^{(e)}$ and the unique residual mean $x^{(r_0)}$ closer to $x^{(e)}$ are rather small,

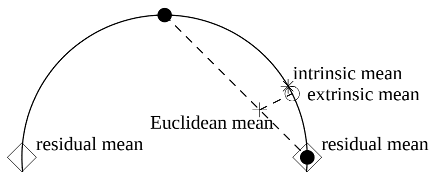

Figure 2: *Means on a circle of a distribution taking the upper dotted value with probability $1/3$ and the lower right dotted value with probability $2/3$. The latter happens to be one of the two residual means.*

namely of the same order as the squared proximity of the modulus $\|\mathbb{E}(X)\|$ of the Euclidean mean to 1, cf. Table 1. We will use the following condition

$$\|x^{(e)} - x^{(r_0)}\|\,,\quad \|x^{(e)} - x^{(i)}\| \quad = \quad O\big((1 - \|\,\mathbb{E}(X)\|)^2\big) \tag{5}$$

with the *concentration parameter* $1 - \|\,\mathbb{E}(X)\|$.

Corollary 6.7. *Under condition (5), if all three means are unique, then the great circular segment between the residual $x^{(r_0)}$ mean closer to the extrinsic mean $x^{(e)}$ and the intrinsic mean $x^{(i)}$ is divided by the extrinsic mean in approximation by the ratio $1:3$:*

$$\begin{aligned} x^{(r_0)} &= \frac{\|\,\mathbb{E}(X)\|}{\lambda^{(r_0)}}\left(x^{(e)} - \frac{\mathbb{E}\big(\langle X - x^{(e)}, X\rangle\, X\big)}{\|\,\mathbb{E}(X)\|} + O\big((1 - \|\,\mathbb{E}(X)\|)^2\big)\right) \\ x^{(i)} &= \frac{\|\,\mathbb{E}(X)\|}{\lambda^{(i)}}\left(x^{(e)} + \frac{1}{3}\,\frac{\mathbb{E}\big(\langle X - x^{(e)}, X\rangle\, X\big)}{\|\,\mathbb{E}(X)\|} + O\big((1 - \|\,\mathbb{E}(X)\|)^2\big)\right) \end{aligned}$$

with $\lambda^{(i)}$ and $\lambda^{(r_0)}$ from Theorem 6.5.

Proof. For any $x, p \in S^{D-1}$ decompose $p - x = p - \langle x, p\rangle\, x - z(x,p)x$ with $z(x,p) = 1 - \langle x, p\rangle$, the length of the part of p normal to the tangent space at x. Note that $\mathbb{E}\left(z(x^{(e)}, X\right) = 1 - \|\,\mathbb{E}(X)\|$. Now, under condition (5), verify the first assertion using Theorem 6.5:

$$x^{(r0)} \quad = \quad \frac{1}{\lambda^{(r_0)}}\Big(\mathbb{E}(X) - \mathbb{E}\left(z(x^{(r_0)}, X)X\right)\Big)\,.$$

On the other hand since

$$\frac{\arccos(1-z)}{\sqrt{1-(1-z)^2}} \quad = \quad 1 + \frac{1}{3}\,z + \frac{2}{15}\,z^2 + \dots$$

we obtain with the same argument the second assertion

$$\begin{aligned}
x^{(i)} &= \frac{1}{\lambda^{(i)}}\,\mathbb{E}\left(\frac{\arccos\langle X, x^{(i)}\rangle}{\sqrt{1-\langle X, x^{(i)}\rangle^2}}\,X\right)\\
&= \frac{1}{\lambda^{(i)}}\left(\mathbb{E}(X) + \frac{1}{3}\,\mathbb{E}\left(z(x^{(i)},X)\,X\right) + \frac{2}{15}\,\mathbb{E}\left(z(x^{(i)},X)^2\,X\right) + \dots\right)\\
&= \frac{\|\,\mathbb{E}(X)\|}{\lambda^{(i)}}\left(x^{(e)} + \frac{1}{3\|\,\mathbb{E}(X)\|}\,\mathbb{E}\left(z(x^{(e)},X)\,X\right) + O\big((1-\|\,\mathbb{E}(X)\|)^2\big)\right)\,.
\end{aligned}$$

$\square$

Remark 6.8. *The tangent vector defining the great circle approximately connecting the three means is obtained from correcting with the expected normal component of any of the means. As numerical experiments show, this great circle is different from the first* principal component geodesic *as defined in Huckemann and Ziezold (2006).*

Recall the following connection between intrinsic and extrinsic spherical versus intrinsic and Ziezold shape means, respectively, cf. Lemmas 4.1 and 4.8. It is easy to see the corresponding connection for residual versus Procrustes means.

Remark 6.9. *Suppose that a random shape X on Σ_m^k is supported by $R \subset \Sigma_m^k$ admitting a horizontal measurable lift $L \subset S_m^k$ in optimal position to $p \in S_m^k$. Define the random variable Y on S_m^k by $\pi \circ Y = X$. Then we have that*

if $[p]$ is an intrinsic mean of X then p is an intrinsic mean of Y,

if $[p]$ is a Procrustes mean of X then p is a residual mean of Y,

if $[p]$ is a Ziezold mean of X then p is an extrinsic mean of Y

In consequence, Corollary 6.7 extends at once to Kendall's shape spaces.

Corollary 6.10. *Suppose that a random shape X on Σ_m^k with unique intrinsic mean $\mu^{(i)}$, unique Ziezold mean $\mu^{(z)}$ and unique Procrustes mean $\mu^{(p_0)}$ closer to $\mu^{(i)}$ is supported by $R \subset \Sigma_m^k$ admitting a horizontal measurable lift $L \subset S_m^k$ in optimal position to $\mu^{(z)} \in S_m^k$. If the means are sufficiently close to each other in the sense of*

$$d_{\Sigma_m^k}(\mu^{(z)} - \mu^{(p_0)})\,,\quad d_{\Sigma_m^k}(\mu^{(z)} - \mu^{(i)}) \;=\; O\big((1-\|\,\mathbb{E}(Y)\|)^2\big)$$

with the random pre-shape Y on L defined by $X = \pi \circ Y$, then the generalized geodesic segment between $\mu^{(i)}$ and $\mu^{(p_0)}$ is approximately divided by $\mu^{(z)}$ by the ratio $1:3$ with an error of order $O\big((1-\|\,\mathbb{E}(Y)\|)^2\big)$.

7 Examples: Exemplary Datasets and Simulations

7.1 The $1:3$ property

In the first example we illustrate Corollary 6.10 on the basis of four data sets:

poplar leaves: contains 104 quadrangular planar shapes extracted from poplar leaves in a joint collaboration with Institute for Forest Biometry and Informatics at the University of Göttingen, cf. Huckemann et al. (2009).

digits '3': contains 30 planar shapes with 13 landmarks each, extracted from handwritten digits '3', cf. Dryden and Mardia (1998, p. 318).

macaque skulls: contains three-dimensional shapes with 7 landmarks each, of 18 macaque skulls, cf. Dryden and Mardia (1998, p. 16).

iron age brooches: contains 28 three-dimensional tetrahedral shapes of iron age brooches, cf. Small (1996, Section 3.5).

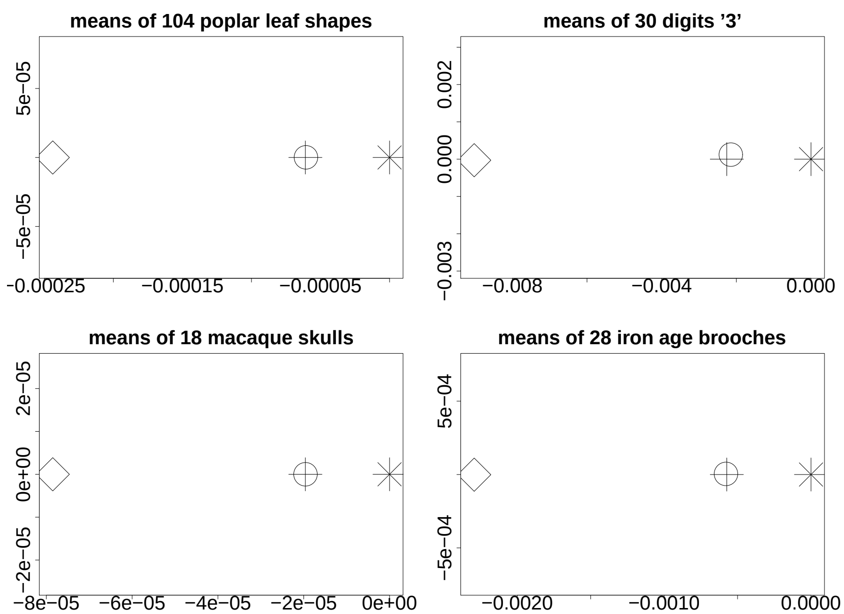

Figure 3: *Depicting shape means for four typical data sets: intrinsic (star), Ziezold (circle) and Procrustes (diamond) projected to the tangent space at the intrinsic mean. The cross divides the generalized geodesic segment joining the intrinsic with the Procrustes mean by the ratio* $1:3$.

data set	$d_{\Sigma_m^k}(\mu^{(i)}, \mu^{(z)})$	$d_{\Sigma_m^k}(\mu^{(p_0)}, \mu^{(z)})$	$d_{\Sigma_m^k}(\mu^{(p_0)}, \mu^{(i)})$	$(1 - \lvert\mathbb{E}(Y)\rvert)^2$
poplar leaves	$6.05e-05$	$1.83e-04$	$2.44e-04$	$5.24e-05$
digits '3'	0.00154	0.00452	0.00605	0.00155
macaque skulls	$1.96e-05$	$5.89e-05$	$7.85e-05$	$7.59e-06$
iron age brooches	0.000578	0.001713	0.002291	0.000217

Table 1: *Mutual shape distances between intrinsic mean $\mu^{(i)}$, Ziezold mean $\mu^{(z)}$ and Procrustes mean $\mu^{(p_0)}$ closer to the Ziezold mean for various data sets. Right column: the squared distance between modulus of Euclidean mean and* 1*, cf. Corollary 6.10.*

As clearly visible from Figure 3 and Table 1, the approximation of Corollary 6.10 for two- and three-dimensional shapes is highly accurate for data of little dispersion (the macaque skull data) and still fairly accurate for highly dispersed data (the digits '3' data).

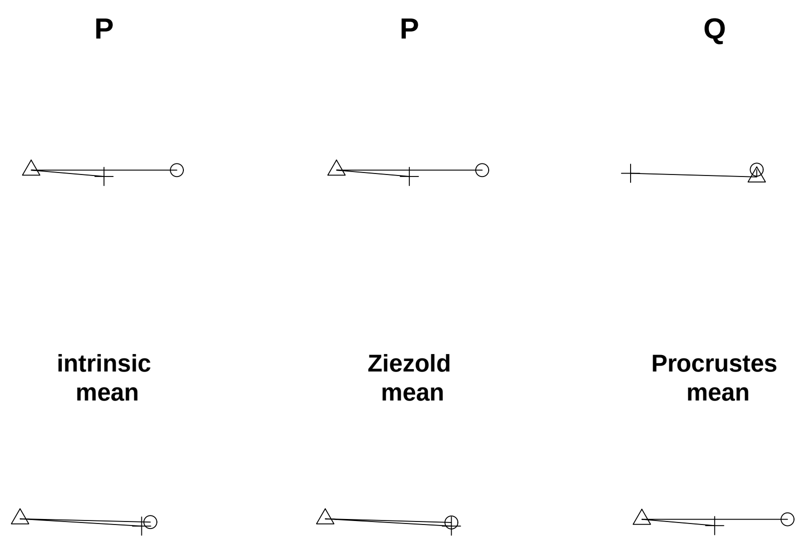

Figure 4: *A data set of three planar triangles (top row) with its corresponding intrinsic mean (bottom left), Ziezold mean (bottom center) and Procrustes mean (bottom right).*

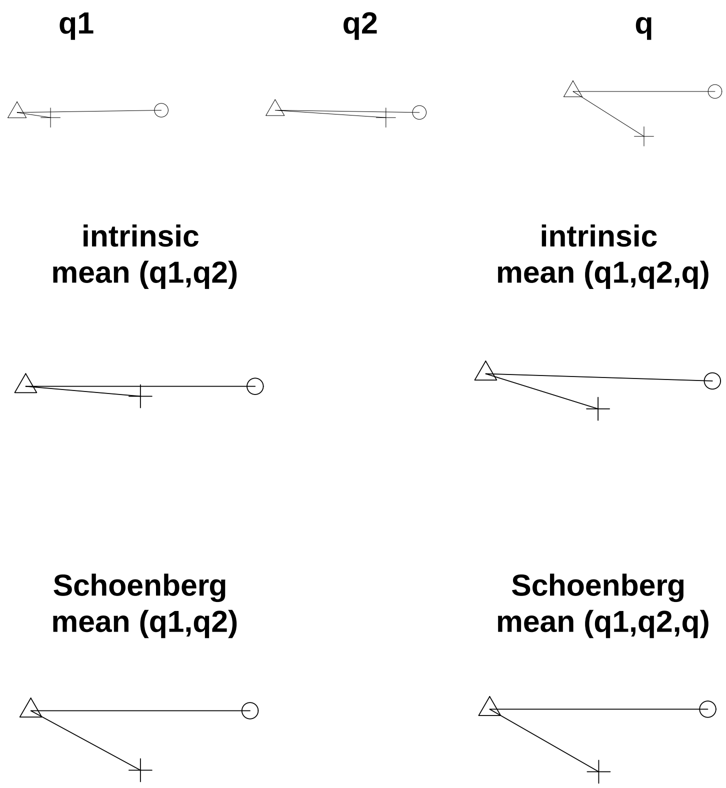

Figure 5: *Planar triangles* $q_1 = x\cos\beta - w_2 \sin\beta$, $q_2 = x\cos\beta - w_2\sin\beta$ *and* $q = x\cos\beta' + w_1 \sin\beta'$ *with* x, w_1, w_2 *from Remark 5.3,* $\phi = 0.05$, $\beta = 0.3 = \beta'$ *(top row). Intrinsic means (middle row) of sample* (q_1, q_2) *(left) and* (q_1, q_2, q) *(right). Schoenberg means (bottom row) of sample* (q_1, q_2) *(left) and* (q_1, q_2, q) *(right).*

7.2 "Blindness" of Procrustes and Schoenberg Means

In the second example we illustrate an effect of "blindness to data" of Procrustes means and Schoenberg means. The former blindness is due to the affinity of the Procrustes mean to the mode in conjunction with curvature, the latter is due to non-isometry of the Schoenberg embedding. While the former effect occurs only for some highly dispersed data when the analog of condition (5) is violated, the latter effect is local in nature and may occur for concentrated data as well.

Example 7.1. *Reenacting the situation of Example 6.6 and Figure 2, the shapes of the triangles* P *and* Q *in Figure 4 are almost maximally remote. Since the* *mode P is assumed twice and Q only once, the Procrustes mean is nearly blind to Q. In case of*

$$P = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 1 & -2 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & -\frac{3}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

the Procrustes mean shape of the sample P, P, Q is the one-dimensional P because of blindness to the two-dimensional Q.

Let us record this fact.

Remark 7.2. *The Procrustes mean of a random shape assuming both regular and degenerate shapes with non-zero probability may be degenerate.*

Even though Schoenberg means have been introduced to tackle 3D shapes, the effect of "blindness" can be well illustrated already for 2D. To this end consider $x = x(\phi)$, $w_1 = w_1(\phi)$ and $w_2 = w_2(\phi)$ as introduced in Remark 5.3. Along the horizontal geodesic through x with initial velocity w_2 we pick two points $q_1 = x\cos\beta + w_2 \sin\beta$ and $q_2 = x\cos\beta - w_2\sin\beta$. On the orthogonal horizontal geodesic through x with initial velocity w_1 pick $q = x\cos\beta' + w_1\sin\beta'$. Recall from Remark 5.3, that along that geodesic the derivative of the Schoenberg embedding can be made arbitrarily small for ϕ near 0. Indeed, Figure 5 illustrates that for small ϕ, the Schoenberg mean is nearly unchanged if the triangle q is added to the sample q_1, q_2. The corresponding intrinsic means, however, are sensitive to the sample's expansions.

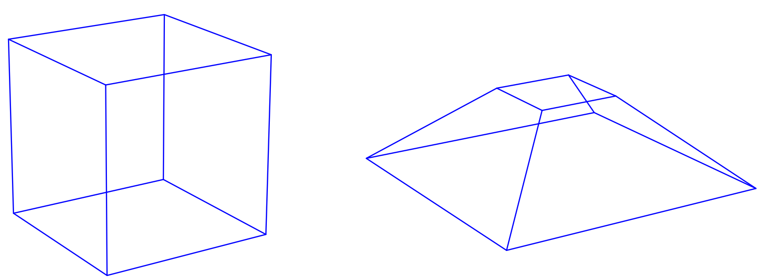

Figure 6: *cube (left) and pyramid (right) for classification.*

7.3 Classification Power of a Test

In the ultimate example we illustrate the consequences of the choice of tangent space coordinates and the effect of the tendency of the Schoenberg mean to increase dimension by a classification simulation. To this end we apply a Hotelling T^2-test to discriminate the shapes of 10 noisy samples of regular unit cubes from the shapes of 10 noisy samples of pyramids with top section chopped off, each with 8 landmarks, given by the following configuration matrix

$$\begin{pmatrix} 0 & 1 & \frac{1+\epsilon}{2} & \frac{1-\epsilon}{2} & 0 & 1 & \frac{1+\epsilon}{2} & \frac{1-\epsilon}{2} \\ 0 & 0 & \frac{1-\epsilon}{2} & \frac{1-\epsilon}{2} & 1 & 1 & \frac{1+\epsilon}{2} & \frac{1+\epsilon}{2} \\ 0 & 0 & \epsilon & \epsilon & 0 & 0 & \epsilon & \epsilon \end{pmatrix}$$

(cf. Figure 6) determined by $\epsilon > 0$. In the simulation, indenpendent Gaussian noise with variance $\sigma^2 > 0$ is added to each landmark measurement. Table 2 gives the number of correct classifications for 1,000 simulations for given α, ϵ and significance level.

σ	ϵ	level	intrinsic mean with residual tangent space coordinates	intrinsic mean with intrinsic tangent space coordinates	Ziezold mean	Procrustes mean	Schoenberg mean
0.35	0.05	0.1	522	542	538	531	442
0.3	0.1	0.1	558	583	581	575	483
0.2	0.2	0.05	556	575	574	575	512
0.1	0.3	0.01	359	369	369	368	366
0.05	0.35	0.01	764	780	780	780	809

Table 2: *Number of correct classifications within 1,000 simulations each of* 10 *unit-cubes and pyramids determined by* ϵ *(which gives the height), where each landmark is independently corrupted by Gaussian noise of variance* σ^2*, via a Hotelling* T^2*-test for equality of means to the given significance level.*

As visible from Table 2, discriminating nearly flat pyramids ($\epsilon = 0.05$) from cubes is achieved much better by employing intrinsic, Ziezold or Procrustes means rather than Schoenberg means. This finding is in concord with Theorem 5.4: samples of size 10 of three-dimensional configurations yield Euclidean means a.s. in $\mathcal{P}^7$ which are projected to $\mathcal{P}^3$ to obtain Schoenberg means in Σ_3^8. In consequence, Schoenberg means of nearly two-dimensional pyramids are essentially three dimensional. With increased height of the pyramid ($\epsilon = 0.35$), i.e. for more pronounced third dimension and increased proximity to the unit cube), all means perform almost equally well, with a tendency of the Schoenberg mean to eventually outperform the others. Moreover in any case, intrinsic means with intrinsic tangent space coordinates behave much poorer in view of shape discrimination than intrinsic means with residual tangent space coordinates, cf. Remark 6.1. The latter (intrinsic means with residual tangent space coordinates) are better or equally good as Ziezold and Procrustes means (which naturally use residual tangent space coordinates).

intrinsic mean	Ziezold mean	Procrustes mean	Schoenberg mean
235.167	182.335	225.982	35.595

Table 3: *Average time in seconds for computation of* $1,000$ *means of sample size* 20.

In conclusion we record the time for mean computation in Table 3. While

Ziezold means compute in approximately 3/4 of the computational time for intrinsic means, Schoenberg means are obtained approximately 5 times faster.

Note that we have included inference based on Procrustes means for illustration, even though no stability result is available and hence at this point, it cannot be theoretically justified that the CLT holds for Procrustes means as well.

8 Discussion

By establishing stability results for intrinsic and Ziezold means on the manifold part of a shape space, a gap in asymptotic theory for general non-manifold shape spaces could be closed, now allowing for multi-sample tests of equality of intrinsic means and Ziezold means. A similar stability assertion in general is false for Procrustes means. Note that the argument applied to intrinsic and Ziezold means fails for Procrustes means, since in contrast to (3) the sum of Procrustes residuals is in general non-zero. The findings on dimensionality condense to

Procrustes means may decrease dimension,
intrinsic and Ziezold means preserve dimension,
Schoenberg means tend to increase dimension.

Due to the proximity of Ziezold and intrinsic means on Kendall's shape spaces in most practical applications, taking into account that the former are computationally easier accessable (optimally positioning and Euclidean averaging in every iteration step) than intrinsic means (optimally positioning and weighted averaging in every iteration step), in practical applications Ziezold means, may be preferred over intrinsic means. They may be even more preferred over intrinsic means, since Ziezold means naturally come with residual tangent space coordinates which may allow in case of intrinsic means for a slightly higher finite power of tests than intrinsic tangent space coordinates.

Computationally much faster (not relying on iteration at all) are Schoenberg means which are available for Kendall's reflection shape spaces. As a drawback, however, Schoneberg means, seem less sensitive for dimensionality of configurations considered than intrinsic or Ziezold means. In consequence, multi-sample test based on Schoenberg means may have a considerably lower power than when based on intrinsic or Ziezold means.

We note that Ziezold means may be defined for the shape spaces of planar curves introduced by Zahn and Roskies (1972), which are currently very popular e.g. Klassen et al. (2004) or Schmidt et al. (2006). Employing Ziezold means there, a computational advantage greater than found here can be expected since computation of iterates of intrinsic means involve computations of geodesics which themselves can only be found iteratively.

Acknowledgment

The author would like to thank Alexander Lytchak for helpful advice on differential geometic issues. Also, the author gratefully acknowledges support by DFG grant MU 1230/10-1 and GRK 1023.

References

Bandulasiri, A., Patrangenaru, V., 2005. Algorithms for nonparametric inference on shape manifolds.

Bhattacharya, A., 2008. Statistical analysis on manifolds: A nonparametric approach for inference on shape spaces. , to appear.

Bhattacharya, R. N., Patrangenaru, V., 2003. Large sample theory of intrinsic and extrinsic sample means on manifolds I. Ann. Statist. 31 (1), 1–29.

Bhattacharya, R. N., Patrangenaru, V., 2005. Large sample theory of intrinsic and extrinsic sample means on manifolds II. Ann. Statist. 33 (3), 1225–1259.

Blum, H., Nagel, R. N., 1978. Shape description using weighted symmetric axis features. Pattern Recognition 10 (3), 167–180.

Bredon, G. E., 1972. Introduction to Compact Transformation Groups. Vol. 46 of Pure and Applied Mathematics. Academic Press.

Choquet, G., 1954. Theory of capacities. Ann. Inst. Fourier 5, 131–295.

Dryden, I. L., Kume, A., Le, H., Wood, A. T. A., 2008. A multidimensional scaling approach to shape analysis. To appear.

Dryden, I. L., Mardia, K. V., 1998. Statistical Shape Analysis. Wiley, Chichester.

Fréchet, M., 1948. Les éléments aléatoires de nature quelconque dans un espace distancié. Ann. Inst. H. Poincaré 10 (4), 215–310.

Gower, J. C., 1975. Generalized Procrustes analysis. Psychometrika 40, 33–51.

Hendriks, H., 1991. A Cramér-Rao type lower bound for estimators with values in a manifold. Journal of Multivariate Analysis 38, 245–261.

Hendriks, H., Landsman, Z., 1996. Asymptotic behaviour of sample mean location for manifolds. Statistics & Probability Letters 26, 169–178.

Hendriks, H., Landsman, Z., 1998. Mean location and sample mean location on manifolds: asymptotics, tests, confidence regions. Journal of Multivariate Analysis 67, 227–243.

Hendriks, H., Landsman, Z., Ruymgaart, F., 1996. Asymptotic behaviour of sample mean direction for spheres. Journal of Multivariate Analysis 59, 141–152.

Hobolth, A., Kent, J., Dryden, I., 2002. On the relation between edge and vertex modelling in shape analysis. Scandinavian Journal of Statistics 29 (3), 355–374.

Hotz, T., Huckemann, S., Gaffrey, D., Munk, A., Sloboda, B., 2010. Shape spaces for pre-alingend star-shaped objects in studying the growth of plants. Journal of the Royal Statistical Society, Series C 59 (1), 127–143.

Huckemann, S., Hotz, T., Munk, A., 2009. Intrinsic MANOVA for Riemannian manifolds with an application to Kendall's space of planar shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence. To appear.

Huckemann, S., Ziezold, H., 2006. Principal component analysis for Riemannian manifolds with an application to triangular shape spaces. Adv. Appl. Prob. (SGSA) 38 (2), 299–319.

Jupp, P. E., 1988. Residuals for directional data. J. Appl. Statist. 15 (2), 137–147.

Karcher, H., 1977. Riemannian center of mass and mollifier smoothing. Communications on Pure and Applied Mathematics XXX, 509–541.

Kendall, D., 1974. Foundations of a theory of random sets. Stochastic Geom., Tribute Memory Rollo Davidson, 322-376 (1974).

Kendall, D. G., 1984. Shape manifolds, Procrustean metrics and complex projective spaces. Bull. Lond. Math. Soc. 16 (2), 81–121.

Kendall, D. G., Barden, D., Carne, T. K., Le, H., 1999. Shape and Shape Theory. Wiley, Chichester.

Kendall, W. S., 1990. Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence. Proc. London Math. Soc. 61, 371–406.

Klassen, E., Srivastava, A., Mio, W., Joshi, S., Mar. 2004. Analysis on planar shapes using geodesic paths on shape spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (3), 372–383.

Kobayashi, S., Nomizu, K., 1969. Foundations of Differential Geometry. Vol. II. Wiley, Chichester.

Krim, H., Yezzi, A. J. J. E., 2006. Statistics and Analysis of Shapes. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser, Boston.

Le, H., Jun. 2001. Locating Fréchet means with an application to shape spaces. Adv. Appl. Prob. (SGSA) 33 (2), 324–338.

Le, H., 2004. Estimation of Riemannian barycenters. LMS Journal of Computation and Mathematics 7, 193–200.

Mardia, K., Patrangenaru, V., 2001. On affine and projective shape data analysis. Functional and Spatial Data Analysis, Proceedings of the 20th LASR Workshop (Eds: K.V. Mardia and R.G. Aykroyd), 39–45.

Mardia, K., Patrangenaru, V., 2005. Directions and projective shapes. Ann. Statist. 33, 1666–1699.

Matheron, G., 1975. Random sets and integral geometry. Wiley Series in Probability and Mathematical Statistics. New York.

Mosimann, J. E., 1970. Size allometry: Size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association 65 (330), 930–945.

Munk, A., Paige, R., Pang, J., Patrangenaru, V., Ruymgaart, F., 2008. The one- and multi-sample problem for functional data with application to projective shape analysis. Journal of Multivariate Analysis (99), 815–833.

Nash, J., 1956. The imbedding problem for Riemannian manifolds. Ann. of Math. 63, 20–63.

Palais, R. S., 1960. Slices and equivariant embeddings. In: Borel, A. (Ed.), Seminar on Transformation Groups. No. 46 in Annals of Mathematics Studies. Princeton Univ. Press, Princton NJ, Ch. 8, pp. 101–115.

Schmidt, F. R., Clausen, M., Cremers, D., 2006. Shape matching by variational computation of geodesics on a manifold. In: Pattern Recognition (Proc. DAGM). Vol. 4174 of LNCS. Springer, Berlin, Germany, pp. 142–151.

Small, C. G., 1996. The Statistical Theory of Shape. Springer-Verlag, New York.

Zahn, C., Roskies, R., 1972. Fourier descriptors for plane closed curves. IEEE Trans. Computers C-21, 269–281.

Ziezold, H., 1977. Expected figures and a strong law of large numbers for random elements in quasi-metric spaces. Trans. 7th Prague Conf. Inf. Theory, Stat. Dec. Func., Random Processes A, 591–602.

Ziezold, H., 1994. Mean figures and mean shapes applied to biological figure and shape distributions in the plane. Biom. J. (36), 491–510.