2016

Statistical Analysis on Object Spaces with Applications

Kouadio David Yao

FLORIDA STATE UNIVERSITY

COLLEGE OF ART AND SCIENCES

STATISTICAL ANALYSIS ON OBJECT SPACES WITH APPLICATIONS

By

KOUADIO DAVID YAO

A Dissertation submitted to the
Department of Mathematics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2016

Kouadio David Yao defended this dissertation on November 21, 2016.
The members of the supervisory committee were:

Vic Patrangenaru
Professor Co-Directing Dissertation

Alec Kercheval
Professor Co-Directing Dissertation

Xiuwen Liu
University Representative

Washington Mio
Committee Member

Xiaoming Wang
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

ACKNOWLEDGMENTS

I am very grateful to my major professor, Dr. Victor Patrangenaru for giving me the opportunity to work in the vibrant area of extrinsic analysis. With professor Patrangenaru's guidance and constant support, I was able to conduct some interesting and meaningful work. I am also thankful for the financial support provided by the many research assistantships I benefited from while under his supervision. I would like to thank my Co-Advisor, Dr. Alec Kercheval for his constant guidance and advices. He certainly made this interdisciplinary endeavor much smoother by always providing clear directions and helping a lot with organizing and improving my dissertation. I would also like to thank the other distinguished members of my committee, Professors Xiuwen Liu, Washington Mio and Xiaoming Wang for their active participation, their patience and encouragements during this process and also for their many great comments that helped significantly improving the quality of the presentation in my dissertation.

I am thankful to Mr. Ruite Guo, Mr. David Lester and Dr. Mingfei Qiu for their assistance and great contributions to our collaborative work. Last but not least, I would like to thank National Science Foundation and the Mathematical Sciences Program of the National Security Agency for financial support, and the Statistical and Applied Mathematical Sciences Institute (SAMSI) for supporting my travel expenses to workshops in the Research Triangle.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Most of the data encountered is bounded nonlinear data. The Universe is bounded, planets are sphere like shaped objects, and life growing on Earth comes in various shapes and colors that can hardly be represented as points on a linear space, and even if the object space they sit on is embedded in a Euclidean space, their mean vector can not be represented as a point on that object space, except for the case when such space is convex. To address this misgiving, since the mean vector is the minimizer of the expected square distance, following Fréchet (1948)[11], on a compact metric space, one may consider both minimizers and maximizers of the expected square distance to a given point on the object space as mean, respectively **anti-mean** of a given random point. Of all distances on a object space, one considers here the chord distance associated with an embedding of the object space, since for such distances one can give a necessary and sufficient condition for the existence of a unique Fréchet mean (respectively Fréchet anti-mean). For such distributions these location parameters are called extrinsic mean (respectively extrinsic anti-mean), and the corresponding sample statistics are consistent estimators of their population counterparts. Moreover one derives the limit distribution of such estimators around an anti-mean located at a smooth point. Extrinsic analysis is thus a general framework that allows one to run object data analysis on nonlinear object spaces that can be embedded in a numerical space. New sample tests for extrinsic means, and a test statistic for extrinsic MANOVA on manifolds are also developed here. In particular one focuses on Veronese-Whitney (VW) means and anti-means of 3D projective shapes of configurations extracted from digital camera images. The 3D data extraction is greatly simplified by an RGB based 3D surface reconstruction algorithm using the Faugeras-Hartley-Gupta-Chang 3D reconstruction method (see [10],[12]), that is used to collect 3D image data. In particular one derives two sample tests for face analysis based on projective shapes, and more generally a MANOVA on manifolds method to be used in 3D projective shape analysis. The manifold based approach is also applicable to financial data analysis for exchange rates.

CHAPTER 1

OVERVIEW

Due to technological advances in digital imaging, we are now able to collect and quantify a wide variety of data sets, including 3D surface data from RGB regular digital camera images. Indeed if color pictures of the same scene are collected under fairly uniform lighting conditions, a correlation based algorithm coupled with a 3D reconstruction algorithm may help retrieve surfaces of a 3D scene, including texture. One of the task of this dissertation was to collect such 3D data, and in particular face data including the mid-face of individuals that accepted to have their pictures taken, and volunteered, without being compensated for offering their time. Some of the digital camera data collected this way is posted at stat.fsu.edu/$\sim$vic/Kouadio/collected-by-Davids. The face surfaces, regarded as 2D manifolds in 3D could be partially retrieved using the technique mentioned above and are presented in the data analysis for Chapters 3 and 6. Such surface data is infinite dimensional, thus a drastic data reduction method consisting in landmark coordinate selection post 3D reconstruction was key to speed up the analysis. Moreover, since the camera internal parameters are unknown, for the landmark configurations considered, one may retrieve only the projective shapes (see Patrangenaru et. al.(2010))[23]. Therefore, the object spaces we have to consider are projective shape spaces (see Mardia and Patrangenaru(2005)[20]), which are direct products of real projective spaces, thus having in fact a nonlinear structure of compact smooth manifolds. There are many other examples of object spaces with a manifold structure, arising from morphometric data, protein and DNA structures, aerial or satellite imaging, medical imaging outputs (angiography, CT scans, MRI) beside digital camera imaging considered here (see Patrangenaru and Ellingson (2015)[21]). Fréchet (1948)[11] 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*. A natural way of addressing the problem of analyzing data on such a nonlinear object space, consists of regarding a *random object* X as a random point on a complete metric space $(\mathcal{M}, \rho)$ that often times has a smooth manifold structure (see Patrangenaru and Ellingson (2015)[21]). The numerical space $\mathbb{R}^m$ is the most elementary example of a manifold arising as an object space in Statistics. Therefore, multivariate data analysis is the key basic example of data analysis on a manifold.

Given a random object (r.o.) X on a complete separable metric space $(\mathcal{M}, \rho)$, the expected square distance from X to an arbitrary point $p \in \mathcal{M}$ 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.*[5]. **Unless otherwise specified, throughout this dissertation we will assume that the object space $\mathcal{M}$ can be regarded as a subset of a numerical space via a one to one map $j : \mathcal{M} \to \mathbb{R}^N$, and the distance on $\mathcal{M}$ is ρ_j, the *chord distance* given by**

$$\rho_j(p_1, p_2) = \|j(p_1) - j(p_2)\|. \tag{1.2}$$

If, in addition $\mathcal{M}$ has a smooth manifold structure (see Lee[18] for a definition), we will assume that j is an embedding, that is to say that at each point $p \in \mathcal{M}$, the differential map d_p is a one to one map from the tangent space $T_p\mathcal{M}$ to the tangent space $T_p\mathbb{R}^N$.

In our case, the Fréchet function becomes

$$\mathcal{F}(p) = \int_{\mathcal{M}} \|j(x) - j(p)\|^2 Q(dx), \tag{1.3}$$

where $Q = P_X$ is the probability measure on $\mathcal{M}$, associated with X, and the Fréchet mean set is called *extrinsic mean set*. The complete separable metric space $(\mathcal{M}, \rho_j)$ with chord distance ρ_j and with an additional smooth manifold structure, is isometric to $(j(\mathcal{M}), \rho_0)$ where ρ_0 is the Euclidean distance. **This is by definition an isometric embedding (distance preserving between two points and their images in the ambient space), if we consider the chord distance**.

In general inference for extrinsic mean sets was never considered yet in literature, none the less, in case the extrinsic mean set has a unique point, called the extrinsic mean, there is a large body of literature on this subject (see Patrangenaru and Ellingson (2015)[21], and the related reference therein); this is due to a a simple condition for the existence and uniqueness of the extrinsic mean (see Bhattacharya and Patrangenaru (2003)[5]), saying the extrinsic mean exists if and only if the probability measure Q is j-nonfocal. I will detail this condition in Chapter 2.

1.1 Short summary of results in chapters 3 through 7

In Chapter 3, I use two sample hypothesis testing methods for means of r.o.'s on a Lie group, as developed by Crane and Patrangenaru(2011)[7], that are applied in the context of 3D projective shape analysis to differentiate between faces. I conduct a landmark based analysis on the space of 3D projective shapes of k-ads (labeled points). The object spaces of interest are often nonlinear spaces, and this poses some challenges when attemping a two sample testing problem for mean change for random samples of different sizes. For my statistical testing problems I consider Lie groups, which are smooth manifolds with an additional group structure (in the algebraic sense) where the mulitplicative operation $\otimes$ and the inverse operation are both smooth. With such object spaces I can conduct a two sample hypothesis testing problem for mean change (see Crane and Patrangenaru (2011) [7].) The 3D projective shape spaces of k-ads containing a projective frame at five fixed landmark indices, denoted ΣP_3^k can be identified with $\mathcal{M} = (\mathbb{R}P^3)^q,\ q = k - 5$ which is a Lie group with multiplicative operation denoted $\odot_q$. For $a = 1, 2$, let $Y_{a,1}, \cdots, Y_{a,n_a}$ identically independent distributed random objects (i.i.d.r.o.'s) from the independent j_k-nonfocal probability measures Q_a on $(\mathbb{R}P^3)^q$, where j_k-nonfocal refers to a probability measure for which there is an extrinsc mean. We consider the following hypothesis testing problem,

$$H_0 : \mu_{2,E}^{-1} \odot_q \mu_{1,E} = 1_{(\mathbb{R}P^3)^q} \quad \text{vs.} \quad H_1 : \mu_{2,E}^{-1} \odot_q \mu_{1,E} \neq 1_{(\mathbb{R}P^3)^q} \tag{1.4}$$

were $\mu_{1,E},\ \mu_{2,E}$ are the Veronese-Whitney means on $(\mathbb{R}P^3)^q$. We are able to construct an asymptotic p-value for large samples and $100(1-\alpha)\%$ bootstrap confidence region as well for small sample size at the α level. These results were made possible by knowing the asymptotic convergence of the sequence of random vectors $n^{1/2}\left(\varphi_q(\bar{Y}_{2,E}^{-1} \odot_q \bar{Y}_{1,E})\right)$ where $\bar{Y}_{a,E}$ are the corresponding VW (Veronese-Whitney) sample means and φ_q is an affine chart (i.e. a smooth one-to-one and onto function from $(\mathbb{R}P^3)^q$ to $\mathbb{R}^{3q}$). The data analysis was conducted on three human faces. I placed all ten landmarks on all three subjects using Matlab for all 29 pairs of noncalibrated digital camera images. The reconstruction of the corresponding 3D coordinates was also done in Matlab. I was then able to use the first five reconstructed coordinates to construct the resulting 5-tuples of projective coordinates represent the 3D projective shapes and are the elements that make up the random samples. After conducting the analysis I was able to effectively use hypothesis testing for 3D projective shape mean change to differentiate between faces and also to identify the same face in cross-validation analysis. The analysis I ran, along with the various results, can be found in a couple of publications [24] and [26]. Using the Agisoft software I was able to build a couple of 3D reconstructions of faces with color and texture (see stat.fsu.edu/∼vic/Kouadio/collected-by-Davids/James and stat.fsu.edu/∼vic/Kouadio/collected-by-Davids/Mingfei). This software has not only a more visually appealing 3D reconstruction but would also allow for a much faster recovery of the 3D coordinates of our landmarks.

The work in Chapter 4 was born out of a question asked by Professor Patrangenaru about the hypothesis testing technique developed in [7]. More specifically, for $a = 1, 2$, $X_{a,1}, \ldots, X_{a,n_a}$ i.i.d. random objects on Lie group $(\mathcal{G}, \odot)$, and the hypothesis problem given as follows

$$H_0 : \ \mu_{2,E}^{-1} \odot \mu_{1,E} = \delta \ vs. \ H_1 : \ \mu_{2,E}^{-1} \odot \mu_{1,E} \neq \delta \tag{1.5}$$

we would like to have the asymptotic behavior of

$$\tan_{j(\mu_{2,E}^{-1} \odot \mu_{1,E})} \left(j(\overline{X}_{2,E}^{-1} \odot \overline{X}_{1,E}) - j(\mu_{2,E}^{-1} \odot \mu_{1,E}) \right) \tag{1.6}$$

where $\mu_{1,E}, \mu_{2,E}$ are the extrinsic means and $\Sigma_{1,E}, \Sigma_{2,E}$ their respective corresponding extrinsic covariance matrices. The notation in (1.6) signifies the projection of the vector $\left(j(\overline{X}_{2,E}^{-1} \odot \overline{X}_{1,E}) - j(\mu_{2,E}^{-1} \odot \mu_{1,E}) \right)$ onto the tangent space of $j(\mathcal{G})$ at the point $j(\mu_{2,E}^{-1} \odot \mu_{1,E})$ and this results is given in Theorem 4.2.2 for some embedding $j : \mathcal{G} \to \mathbb{R}^N$ where $\overline{X}_{1,E}$, and $\overline{X}_{2,E}$ are our resulting extrinsic sample means. For a similar hypothesis testing problem as in [7] one of my goals was to take advantage of the CLT (Central Limit Theorem) framework for extrinsic sample means and the confidence regions one can construct from the given asymptotic behavior.

I started by giving a variation of the Delta Method [4] used in [7] which differs from the other one as it uses another extrinsic covariance matrix estimator, and also gives an explicit definition of it (see Lemma 4.1.1.) Let $\mathcal{M}$ and $\mathcal{N}$ be respectively, m-dimensional and n-dimensional smooth manifolds and let $G : \mathcal{M} \times \mathcal{M} \to \mathcal{N}$ be a smooth function between manifolds. In Theorem 4.2.1 I derived the following result;

$$n^{1/2} \ \tan_{j_2(G(\mu_{1,E}, \mu_{2,E}))} \left(j_2 \left(G(\overline{X}_{1,E}, \overline{X}_{2,E}) \right) - j_2 \left(G(\mu_{1,E}, \mu_{2,E}) \right) \right) \xrightarrow{\mathcal{L}} \mathcal{N}_n(0, \Sigma_{j_2,E}^{G}) \tag{1.7}$$

for $a = 1, 2$ let $f_1^{(a)}, \cdots, f_m^{(a)}$ orthonormal basis in $T_{\mu_{a,E}}(\mathcal{M})$. I was then able to have the asymptotic behavior of any smooth function G (between manifolds) and this is done in $T_{G(\mu_{1,E}, \mu_{2,E})}\mathcal{N}$, the tangent space on $\mathcal{N}$ at the point $G(\mu_{1,E}, \mu_{2,E})$ and with the corresponding extrinsic covariance matrix given in term of the extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$ at $\mu_{1,E}$ and $\mu_{2,E}$ respectively. Note that it is important to mention some of the benefits of using the extrinsic analysis framework, especially for computation purposes and more specifically for the sample extrinsic covariance matrix tied to $\mathbb{R}P^m$. For more on the extrinsic sample covariance matrix on $\mathbb{R}P^m$, see [6]. In section 4.3 I apply the new asymptotic results to

$\mathbb{R}P^3$. For $a = 1, 2$ let $[X_{a,1}], \cdots, [X_{a,n_a}]$ be independent random samples defined on $\mathbb{R}P^3$ from j-nonfocal distributions Q_a, with extrinsic means $\mu_{a,E}$ and extrinsic covariance matrices $\Sigma_{a,E}$ I get the following asymptotic behavior.

$$n^{1/2}\ \tan_{j(\mu_{2,E}^{-1}\odot\mu_{1,E})}\left(j(\overline{X}_{2,E}^{-1}\odot\overline{X}_{1,E}) - j(\mu_{2,E}^{-1}\odot\mu_{1,E})\right) \to_d N_m(0_m, \Sigma_E^{\iota G}) \tag{1.8}$$

where for $H(\mu_{2,E}^{-1}, \mu_{1,E}) = (\mu_{2,E}^{-1}\odot\mu_{1,E})$,

$$\Sigma_E^{\iota H} = \frac{1}{\pi}\,(dH^{(1)})\,\Sigma_{2,E}^{\iota}(dH^{(1)})^T + \frac{1}{1-\pi}(dH^{(2)})\Sigma_{1,E}(dH^{(2)})^T, \tag{1.9}$$

where π is the proportion of the first population relative to the total population. I was able to express $G_E^{\iota H}$ the consistent estimator of $\Sigma_E^{\iota H}$. This sample covariance matrix is expressed in a way that reduces the amount of computation by using in its expression the already computationally friendly formula of the sample covariance matrices $G_{1,E}$ and $G_{2,E}$ (see Battacharya and Patrangenaru (2005) [6]) and ,

$$G_E^{\iota H}(j, X_{1,1}, X_{2,1}) = \frac{1}{n_2}\,(d\Gamma^{(1)})\,G_{2,E}(d\Gamma^{(1)})^T + \frac{1}{n_1}(d\Gamma^{(2)})G_{1,E}(d\Gamma^{(2)})^T \tag{1.10}$$

for $d\Gamma^{(a)}, a = 1, 2$ are both diagonal matrices with our choice of basis on $S(4, \mathbb{R})$. One must also note that all the results about $\mathbb{R}P^3$ can be extended to $(\mathbb{R}P^3)^q$, the 3D projective shape space.

Chapter 5 is about extrinsic anti-mean. This chapter includes work I have recently published jointly with V. Patrangenaru and R. Guo (see [27] and [22]). In this chapter I introduce new location parameters, assuming that the object space $(\mathcal{M}, \rho)$ is compact. In particular, if ρ is the chord distance induced by an embedding $j : \mathcal{M} \to \mathbb{R}^N$, the extreme values of the Fréchet function are attained at points on $\mathcal{M}$. Note that the extrinsic mean is defined in fact on any complete metric space that is homeomorphically embedded in $\mathbb{R}^N$, therefore this chapter allows also for the situation when the extrinsic mean is a singular point. Let X be a random object for a distribution Q on $\mathcal{M}$, then we get a distribution for $j(X)$ on $j(Q)$ the ambient space. And we have an extrinsic mean often denoted $\mu_{j,E}$ provided we have a unique projection of μ denoted $P_j(\mu)$ onto the $j(\mathcal{M})$ and μ is called a j-nonfocal point. More specifically, μ j-nonfocal implies that we have $\rho_0(\mu, j(\mathcal{M})) = \rho_0(\mu, j(\mu_{j,E}))$ where $\rho_0(\mu, j(\mathcal{M}))$ is the distance between the point μ and the closest (unique) point on $j(\mathcal{M})$. The notion of anti-mean is motivated by the fact that, even when a distribution Q might not have an extrinsic mean, it may occur that the extrinsic anti-mean exists, thus an extrinsic analysis can still be performed. In case the extrinsic mean is a singular point, the asymptotic distributions of the extrinsic sample mean behave differently. In the case of a stratified space, such as an *open book* the extrinsic sample mean sticks to a lower dimensional stratum (see [3], [13]). The anti-means have a similar asymptotic behavior, thus offering a way to conduct nonparametric data analysis on not just smooth embedded manifolds but in a broader sense, on stratified spaces. In this chapter, I introduce the notion of αj-nonfocal distribution, and it is shown that a distribution has a unique extrinsic anti-mean if and only if it is αj-nonfocal (see Theorem 5.1.1). As a result, one also proves the existence and consistency of the extrinsic sample anti-mean set. In section 5.3, the focus is turned to $\mathbb{R}P^m$ with the VW embedding, and one gives a necessary and sufficient condition for a random axis $[X], X^T X = 1$ being α-VW-nonfocal in terms of eigenvalues of the expected matrix $E(XX^T)$. Further, in this chapter I develop a nonparametric methodology for addressing the hypothesis testing problem

$$H_0 : \alpha\mu_{2,j_q}^{-1} \odot_q \alpha\mu_{1,j_q} = 1_{(\mathbb{R}P^3)^q} \text{ vs. } H_a : \alpha\mu_{2,j_q}^{-1} \odot_q \alpha\mu_{1,j_q} \neq 1_{(\mathbb{R}P^3)^q}. \tag{1.11}$$

As it turns out, the framework developed by Crane and Patrangenaru in [7] can be adapted to the case of anti-means and provided certain general assumption on the VW anti-means $\alpha\mu_{a,j_q}, a = 1, 2$ I conduct, in section 5.5 two sample test to compare 3D projective shapes of two lily flowers, based on their digital camera images.

Chapter 6 is concerned with a new approach of hypothesis testing for the equality of extrinsic means of g random objects, $g \geq 3$. This is an extension of the classical MANOVA (Multivariate Analysis of Variance) problem (see Johnson and Wichern (2008)[15]), in nonparametric setting. This approach is motivated by the standard MANOVA hypothesis testing problem

$$\begin{aligned} &H_0 : \mu_1 = \mu_2 = ... = \mu_g = \mu \\ &H_a : \textit{at least one equation does not hold.} \end{aligned}$$

given the independent random vectors $X_a \sim N_p(\mu_a, \Sigma), a = 1, \dots, g$. We first consider a nonparametric test, based on the pooled sample mean, by dropping the normality assumption, and assuming that asymptotically the ratio between a group size and the total sample size converges to a positive constant, as the total sample size goes to infinity. I extended the ideas developped in the random variable case to object data, assuming that that $Q_a, a = 1, \dots, g$, are independent j- nonfocal probability measures on $\mathcal{M}$ and $X_{a,1}, \dots, X_{a,n_a}$ are i.i.d.r. objects from $Q_a, a = 1, 2, ..., g$. The extrinsic mean of Q_a if $\mu_{a,E}$ and corresponding extrinsic sample means is $\bar{X}_{a,E}$. To test

$$H_0 : \mu_{1,E} = \mu_{2,E} = ... = \mu_{g,E} = \mu_E, H_a : \textit{at least one equation does not hold,}$$

in general I consider the *pooled mean* given by $\mu_E = (j^{-1} \circ P_j)(\lambda_1 j(\mu_{1,E}) + \cdots + \lambda_g j(\mu_{g,E}))$ and the corresponding sample counterpart $\bar{X}_E \in \mathcal{M}$ given by

$$\bar{X}_E = (j^{-1} \circ P_j)\left(\frac{n_1}{n} j(\bar{X}_{1,E}) + \cdots + \frac{n_g}{n} j(\bar{X}_{g,E})\right)$$

where $\bar{X}_{a,E}$ is the extrinsic sample mean for $X_{a,1}$ and $n = \sum_{a=1}^{g} n_a$ and $\frac{n_a}{n} \to \lambda_a > 0$, as $n \to \infty$, with $\Sigma_{a=1}^{g} \lambda_a = 1$. From Theorem 6.2.1 I get two candidate statistics for testing (1.12) that have both asymptotically a χ^2_{gp} distribution. These are used for rejection regions in the large sample case. The small sample case is also addressed via nonparametric bootstrap in Corollary 6.2.2. In Section 6.3 I address the extrinsic MANOVA problem on the 3D projective shape space $(\mathbb{R}P^3)^q$ with the VW embedding. As an example I consider the equality of mean projective shapes of 3D landmark configurations in a number of individuals from digital camera images of their faces.

Chapter 7 is concerned with future directions in extrinsic data analysis it will involve using the 3D face data set I have reconstructed from digital images, to collect landmarks from the remaining faces in the database. Extend the work done in chapters 4, 5 and 6 to data analysis for VW antimeans including to MANOVA for such antimeans.

1.2 Description of contributions

In this section I clearly describe what are my contributions to the various research results in this dissertation, and which of these have been published. I start by recalling all my results that are theorems:

- In Theorem (4.1.1) I developed a new Delta method for a smooth function $F : \mathcal{M}_1 \to \mathcal{M}_2$ where for $a = 1, 2$ $\mathcal{M}_a$ are m_a-dimensional smooth manifolds. The aim was for me to express the resulting covariance matrix in an **explicit form.**

- Theorem (4.2.1) I develop the asymptotic behavior tied to a smooth function $G : \mathcal{M} \times \mathcal{M} \to \mathcal{N}$ between smooth manifolds. This result can certainly be used to get the asymptotic behavior in a case of a two sample hypothesis testing for extrinsic means because it can give the asymptotic behavior of a function G of two extrinsic sample means with an explicit expression of the resulting **extrinsic covariance matrix** written in term a linear combination of the extrinsic matrices tied to each of the two random samples whether they are of same size or not.

- For Theorem (4.2.2) I focus on Lie groups with a multiplicative operation $\odot$ and an inverse map ι. I give an asymptotic behavior for the tangential component $\tan_{j(\mu_{2,E}^{-1} \odot \mu_{1,E})}\left(j(\overline{X}_{2,E}^{-1} \odot \overline{X}_{1,E}) - j(\mu_{2,E}^{-1} \odot \mu_{1,E})\right)$. For this result, I use Theorem (4.1.1) to get the asymptotic behavior of $\tan_{j(\mu_{2,E}^{-1})}\left(j(\overline{X}_{2,E}^{-1}) - j(\mu_{2,E}^{-1})\right)$ and an explicit expression of its corresponding extrinsic covariance matrix $\Sigma_{2,E}^{\iota}$. I then used the results of Theorem (4.2.1) applied to the function $H : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ and given by $H(x_2, x_1) = x_2^{-1} \odot x_1$ to get the desired asymptotic behavior with an explicit expression of the extrinsic covariance matrix.

- In Theorems (5.1.1) and (5.1.2) I give the conditions for existence of the extrinsic anti-mean and the sample extrinsic antimeans. I applied these to a data analysis for anti-mean 3D projective shapes extracted from digital camera images.
- Theorem (6.2.1) I give the expression of two test statistic for the hypothesis testing problem of comparing multiple extrinsic means. One of the test statistic will be used to handle cases for which the extrinsic pooled mean is known and the other can be used whenever the extrinsic pooled mean is unknown.
- For Corollary (6.3.1) I used the results of Theorem (6.2.1) to expressed a couple of test statistic designed to test the 3D mean projective shape changes between multiple VW means.

And below I give a list of ideas I have developed.

- In chapter 4, I developed an idea that would allow anyone to conduct a two sample hypothesis testing involving random samples on smooth embedded manifolds whether the samples are of same sizes or not.
- The extrinsic pooled mean and sample mean inspired by the case for multiple random vectors give the possibility to develop and create a MANOVA for smooth embedded manifolds, allowing for the possibility to test for multiple extrinsic means.

My contribution to the data analysis has been in the form of well defined condition of existence of the extrinsic anti-mean. I also took advantage of the extrinsic CLT result about antimean developped in Patrangenaru et al (2016) [22] to conduct a two sample hypotheis testing method for change in antimean and therefore giving another effective way to differentiate between object via a landmark based approach.

My contribution to the publications listed is

- Patrangenaru, Yao and Guo (2016) [27] I my mork involve the whole of sections 2 through 5.
- Patrangenaru, Guo and Yao (2016) [22] For this publication, my work is featured in the whole of sections 4 and 5.
- For the paper Patrangenaru, Page, Yao, Qiu and Lester (2016) [24]) my work is featured in the whole of sections 4 and 5.
- (Patrangenaru et al (2016) [26]) my work is featured in subsections 3.1 and 3.2 and also in the whole of sections 4 and 5.

CHAPTER 2

PRELIMINARIES

Most of my analysis will be conducted on object spaces. These spaces consist of features measured from sample observations that can no longer be viewed as a values of random vectors if one wishes to conduct a proper statistical analysis on such said spaces. Examples of some object spaces I will consider are the space of points on a sphere and the space of projective shapes of configurations and for such a data set the associated object considered are points on the *projective shape space*. I will regard a *random object* X as a random point on a complete metric space $(\mathcal{M}, \rho_j)$ that has a manifold structure. In section 2.1 I give some relevant definitions and introduce some meaningful concepts we will use throughout the analysis. In the ensuing section I introduce the extrinsic mean and extrinsic sample mean as the unique minimizer of Fréchet functions on $(\mathcal{M}, \rho_j)$. Section 2.3 exposes the reader to a Central Limit Theorem for extrinsic sample means on embedded manifolds. In section 2.4 I present the m-D projective shape space of k-ads (labeled points, landmarks) in general position, which is denoted $P\Sigma_m^k$. I highlight the fact that for $P\Sigma_3^k$ can be identified with $(\mathbb{R}P^3)^q$ with $q = k - 5$. With this particular representation one can now view any elements of the 3-D projective shape space as a q-tuple of elements from the 3D projective space and $(\mathbb{R}P^3)^q$ is embedded via the Veronese-Whitney embedding (see Patrangenaru and Ellingson(2015)[21]). The final section introduce a two sample hypothesis testing problem for extrinsic means on Lie groups and the resulting bootstrap confidence region needed to conduct this test.

2.1 Some important concepts and definitions

The focus of our studies will revolve around metric spaces $(\mathcal{M}, \rho)$ with an additional smooth manifold structure. For that purpose we give the following definition of a *smooth manifold*. We start by giving the definition of a topological manifold.

DEFINITION 2.1.1. ***(Manifolds)***
A metric space $(\mathcal{M}, \rho)$ is a manifold of dimension m or a topological m-manifold if $\mathcal{M}$ is **second countable** *, i.e. there exists a countable basis for the metric topology of $\mathcal{M}$, and also $\mathcal{M}$ is* **locally Euclidean** *of*

dimension m, i.e. every point has a neighborhood that is homeomorphic to an open subset of $\mathbb{R}^m$. And the homeomorphism function $\varphi_U : U \to \varphi_U(U) \in \mathbb{R}^m$ is referred to as an m-dimensional chart on $\mathcal{M}$. We usually denote an m-dimensional chart by the pair (U, φ_U). (see Lee (2002) [18]).

Given a chart (U, φ_U) we call the set U a *coordinate domain*, or *coordinate neighborhood* of each of its points. If in addition $\varphi_U(U)$ is an open ball in $\mathbb{R}^m$, then U is called a *coordinate ball.* The map φ_U is also referred to as a local coordinate map, and its components $(x_U^1, \cdots, x_U^m)$, defined by $\varphi_U(p) = (x_U^1(p), \cdots, x_U^m(p))$ are called *local coordinates* on U. We will sometimes denote a chart by $(U, (x_U^i)_{i=1,\ldots,m})$ if we wish to emphasize the coordinate functions $(x_U^1, \cdots, x_U^m)$. (see Lee (2002) [18]).

Note that a homeomorphism is a bijective continuous function with a continuous inverse. The smooth structure of a manifold is established by a smooth *atlas* or $\mathcal{C}^\infty$ *atlas*.

DEFINITION 2.1.2. *A collection $\mathcal{A} = \{(U_\alpha,\ \varphi_\alpha)_{\alpha\in A}\}$ of $\mathbb{R}^m$-valued charts on the topological manifold $\mathcal{M}$ is called atlas of class C^r if the following conditions are satisfied:*

(i) $\displaystyle\bigcup_{\alpha\in A} U_\alpha = \mathcal{M}$

(ii) Whenever $U_\alpha \cap U_\beta \neq \emptyset$, then the (transition) map between $\varphi_\alpha(U_\alpha \cap U_\beta)$ and $\varphi_\beta(U_\alpha \cap U_\beta)$

$$\varphi_\beta \circ \varphi_\alpha^{-1}|_{\varphi_\alpha(U_\alpha\cap U_\beta)} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

is differentiable. Furthermore, this transition map must also have a differentiable inverse that has continuous partial derivatives up to order r.

(see Lee (2002) [18]).

DEFINITION 2.1.3. *An m-dimensional manifold of class $\mathcal{C}^r$ is a manifold $\mathcal{M}$ along with an $\mathbb{R}^m$-valued atlas of class $\mathcal{C}^r$ on $\mathcal{M}$. We will refer to a smooth manifold as an m-dimensional manifold of class C^∞.*

Example 1. *(i) Naturally, any open set in the Euclidean space $(\mathbb{R}^m, \rho_0)$, is an m-dimensional smooth manifold. Here $\rho_0(x, y) = \|x - y\|$, where $\|(u^1, \ldots, u^m)\|^2 = \sum_{i=1,\ldots,m}(u^i)^2$.*

(ii) The unit sphere $\mathbb{S}^m = \{x \in \mathbb{R}^{m+1} : \|x\| = 1\}$ is an example of m-dimensional smooth manifold.

(iii) The product of a p- dimensional manifold with a q- dimensional manifold is a $(p+q)$- dimensional manifold.

(iv) The space of 1-dimensional linear subspaces of $\mathbb{R}^{m+1}$, called the m-dimensional real projective space *and labeled $\mathbb{R}P^m$ is an example of a m - dimensional manifold that is not a subset of an Euclidean*

space. An element of $\mathbb{R}P^m$ *is often represented by* $[x]$ *where* $x \in \mathbb{R}^{m+1}$*. Here* $[x] = [y] \iff y = \lambda x$ *for some* $\lambda \neq 0$.

(see Lee (2002) [18]).

Note: A projective point $[x] \in \mathbb{R}P^m$ can also have a *spherical representation*, when thought of as a pair of antipodal points on S^m, and $[x] = \{x, -x\}$, with $\|x\| = 1$ and $x \in \mathbb{R}^{m+1}$. From this point on when referring to a projective point we will use this particular representation. (see [2] or [21])

The definitions of smoothness of diffeomorphism and differentiable curves will be needed for us to introduce tangent vectors and tangent spaces which are an integral part of the asymptotic analysis we will conduct later.

DEFINITION 2.1.4. *(smooth function) Let* $\mathcal{M}$ *be a smooth* m*-manifold, a function* $f : \mathcal{M} \to \mathbb{R}^k$ *is said to be* smooth *if for every* $p \in \mathcal{M}$*, there exists a smooth chart* (U, φ) *for* $\mathcal{M}$ *whose domain contains* p *and such that the composite function* $f \circ \varphi^{-1}$ *is smooth on the open subset* $\varphi(U) \subset \mathbb{R}^m$*. (see Lee (2002) [18]).*

DEFINITION 2.1.5. *(Smooth map between manifolds)*

A function $F : \mathcal{M} \to \mathcal{N}$ *between two smooth manifolds is differentiable, if for any charts* (U, φ_U) *on* $\mathcal{M}$ *and* (V, ϕ_V) *, on* $\mathcal{N}$*, the composite map,* $\phi_U \circ F \circ \varphi_V^{-1}\big|_{\phi(U \cap V)}$ *is differentiable of class* $\mathcal{C}^\infty$*. The composite map above is referred to as the local representative. (see Lee (2002) [18]).*

DEFINITION 2.1.6. *A* **diffeomorphism** *between (differentiable) manifolds* $\mathcal{M}$ *and* $\mathcal{N}$ *is a differentiable function* $F : \mathcal{M} \to \mathcal{N}$ *that has a differentiable inverse. Furthermore, we say that* $\mathcal{M}$ *and* $\mathcal{N}$ *are diffeomorphic if there exists a diffeomorphism between them. (see Lee (2002) [18]).*

DEFINITION 2.1.7. *A differentiable curve (path) on a smooth manifold* $\mathcal{M}$ *is a differentiable function from an interval to* $\mathcal{M}$*. Two such paths* c_1 *and* c_2*, defined on a neighborhood of* $0 \in \mathbb{R}$ *are tangent at* p *if* $c_1(0) = c_2(0) = p$ *and there is a chart* (U, φ_U) *around* p *such that*

$$(\varphi_U \circ c_1)'(0) = (\varphi_U \circ c_2)'(0)$$

(see Patrangenaru and Ellingson (2015) [21])

With the definition of differential curves we can now give a definition of tangent spaces which is more useful for object data analysis.

DEFINITION 2.1.8. *(Tangent vectors and tangent space)* *(i) The set of all paths tangent at p is called tangent vector ν_p at $p = c(0)$, and is labeled $\nu_p = \frac{dc}{dt}(0) = \frac{dc}{dt}\big|_0$.*

(ii) The tangent space $T_p\mathcal{M}$ at a point p of a manifold $\mathcal{M}$ is the set of all tangent vectors $\nu_p = \frac{dc}{dt}\big|_0$ to curves $c : (-\varepsilon, \varepsilon) \to \mathcal{M}$ with $p = c(0)$.

We will use the notations (p, ν), ν_p, and ν for a tangent vector in $T_p\mathcal{M}$, depending on how much emphasis we wish to give to the point p. (see Patrangenaru and Ellingson (2015) [21])

Example of tangent vectors

(E1) If $e_1, \cdots, e_m$ is the usual basis of $\mathcal{M} = \mathbb{R}^m$ and $p \in \mathcal{M}$ the following partial derivatives

$$\frac{\partial}{\partial x^1}\bigg|_p, \cdots, \frac{\partial}{\partial x^m}\bigg|_p$$

are tangent vectors in $T_p\mathbb{R}^m$. For $i = 1, \ldots, m$, $\frac{\partial}{\partial x^i}\big|_p$ is the tangent vector

$$e_i = \frac{dc_i}{dt}(0) = \frac{\partial}{\partial x^i}\bigg|_p,$$

where $c_i(t) = p + te_i$.

(E2) Similarly, if (U, φ) is a chart on $\mathcal{M}$, around p, $\frac{\partial}{\partial x^i}\big|_p^\varphi$ is the tangent vector

$$\frac{dc_i}{dt}(0) = \frac{\partial}{\partial x^i}\bigg|_p^\varphi,$$

where $c_i(t) = \varphi^{-1}(\varphi(p) + te_i)$.

(E3) In another example, consider $\mathcal{M} = \mathbb{S}^m$ regarded as a subset of $\mathbb{R}^{m+1}$, then the tangent space at $p \in \mathbb{S}^m$ can be described as

$$T_p\mathbb{S}^m = \{(p, v),\ v \ \in \mathbb{R}^{m+1}\ |v^T p = 0\} \tag{2.1}$$

(E4) Let $\mathbb{R}P^m$ be identified with antipodal points (spherical representation) then if $[x] = \{x, -x\} \in \mathbb{R}P^m$, the tangent space at $[x]$ is described as

$$T_{[x]}\mathbb{R}P^m = \{([x], \nu),\ \nu \ \in \mathbb{R}^{m+1}\ |\nu^T x = 0\} \tag{2.2}$$

(see Patrangenaru and Ellingson (2015) [21]).

PROPOSITION 2.1.1. *Let (U, φ) be a chart on $\mathcal{M}$. Then $T_p(\mathcal{M})$ has a basis of tangent vectors $\frac{\partial}{\partial x^1}\big|_p, \cdots, \frac{\partial}{\partial x^m}\big|_p$ where $(x^1, ..., x^m)$ is the system of **local coordinates** associated with the chart (U, ψ). Each vector $\nu_p \in T_p\mathcal{M}$ can be written uniquely as a linear combination of $\frac{\partial}{\partial x^1}\big|_p, \cdots, \frac{\partial}{\partial x^m}\big|_p$ and we have $\nu_p = \sum_{i=1}^m \nu^i \frac{\partial}{\partial x^i}\big|_p$ with any choice of charts on $\mathcal{M}$ and the numbers $(\nu^1, \nu^2, ..., \nu^m)$ are called the components of ν_p with respect to the given coordinate system. (see Lee (2002) [18])*

DEFINITION 2.1.9. *(Tangent Bundle).*

The tangent bundle $T\mathcal{M}$ of an m-dimensional manifold $\mathcal{M}$ is the disjoint union of the tangent spaces at all points of $\mathcal{M}$; it has a $2m$-dimensional manifold structure. The tangent bundle is often represented by the triple $(T\mathcal{M}, \Pi, \mathcal{M})$ where Π is a natural projection map *and $\Pi : T\mathcal{M} \to \mathcal{M}$ is a differentiable map which associates to each tangent vector its base point, $\Pi((p, \nu_p)) = p$. (see Lee (2002) [18] or Patrangenaru and Ellingson (2015) [21]).*

DEFINITION 2.1.10 (Vector Fields)**.** *If $\mathcal{M}$ is a smooth manifold, a* vector field *on $\mathcal{M}$ is a smooth section of the projection map Π, that is a smooth map $Y : \mathcal{M} \to T\mathcal{M}$ usually written $p \to Y(p)$, with the property that*

$$\Pi \circ Y = Id_{\mathcal{M}}, \tag{2.3}$$

or equivalently, $Y(p) \in T_p\mathcal{M}$ for each $p \in \mathcal{M}$. (see Lee (2002) [18] or Patrangenaru and Ellingson (2015) [21])

One may think of a vector field on $\mathcal{M}$ in the same way we think of vector fields in Euclidean spaces: as an arrow attached to each point of $\mathcal{M}$, chosen to be tangent to $\mathcal{M}$ and to vary smoothly from point to point. The value of a smooth vector field at the point p is a tangent vector at each point $p \in \mathcal{M}$.

Example 2. *If $(U, (x^i))$ is any smooth chart on $\mathcal{M}$, the assignment*

$$p \to \left.\frac{\partial}{\partial x^i}\right|_p \tag{2.4}$$

determines a smooth vector field on U, called the ith coordinate vector field *and denoted by $\frac{\partial}{\partial x^i}$. (see Lee (2002) [18])*

The set of all smooth vector fields on $\mathcal{M}$ often denoted by $\mathcal{T}(\mathcal{M})$ is an infinite-dimensional vector space under point wise addition and scalar multiplication:

$$(aY + bZ)(p) = aY(p) + bZ(p)$$

(see Lee (2002) [18])

DEFINITION 2.1.11. *Let $U \subset \mathcal{M}$ be an open subset of an m-dimensional smooth manifold. A local frame field is a system of m vector fields $(V_1, \ldots, V_m)$ of $T\mathcal{M}$ over U whose values $V_1(p), \ldots, V_m(p)$ are linearly independent in $T_p\mathcal{M}$ for each $p \in U$ (see Lee (2002) [18] or Patrangenaru and Ellingson (2015) [21]).*

Recall that for any smooth m-manifold $\mathcal{M}$, the tangent bundle has a natural topology and smooth structure that makes it into a smooth $2m$-dimensional manifold such that $\Pi : T\mathcal{M} \to \mathcal{M}$ is a smooth map. We can therefore have maps from one tangent bundle $T\mathcal{M}$ to another tangent bundle $T\mathcal{N}$. We now define a special map below.

DEFINITION 2.1.12. *(Tangent Map)*

(i) If $f : \mathcal{M}_1 \to \mathcal{M}_2$ is a differentiable function between manifolds, its tangent map is the function $df : T\mathcal{M}_1 \to T\mathcal{M}_2$, given by

$$df\left(\left.\frac{dc}{dt}\right|_{c(o)}\right) = \left.\frac{d(f\circ c)}{dt}\right|_{f(c(o))}$$

for all differentiable curves c defined on an interval containing $0 \in \mathbb{R}$.

(ii) The differential of f at the point p is the restriction of the tangent map, regarded as a linear function

$$d_pf : T_p\mathcal{M}_1 \to T_{f(p)}\mathcal{M}_2$$

$$df\left(\left.\frac{dc}{dt}\right|_{p}\right) = \left.\frac{d(f\circ c)}{dt}\right|_{f(p)} \tag{2.5}$$

For the definition above please refer to Patrangenaru and Ellingson (2015) [21]. Note that the restriction of df at the point p is a linear function that sends a tangent vector of $\mathcal{M}_1$ to a corresponding tangent vector of $\mathcal{M}_2$. Such a linear map is also referred to as a push forward see Lee (2002) [18].
Data analysis on embedded manifolds will be the focus of our study. On such manifolds we can define a distance with very useful properties.

DEFINITION 2.1.13. *(Embedding)*
An embedding of a manifold $\mathcal{M}$ in a Euclidean space $\mathbb{R}^k$ is a differentiable one-to-one function $j : \mathcal{M} \to \mathbb{R}^k$, for which

(i) the differential d_pj is a one-to-one function from $T_p\mathcal{M}$ to $T_{j(p)}\mathbb{R}^k$ at any point $p \in \mathcal{M}$, and

(ii) j is a homeomorphism from $\mathcal{M}$ to $j(\mathcal{M})$ with metric topology induced by the Euclidean distance.

(see Patrangenaru and Ellingson (2015) [21])

REMARK 2.1.1. *Given an embedded manifold $\mathcal{M}$ with embedding $j : \mathcal{M} \to j(\mathcal{M}) \subset \mathbb{R}^k$, we will, throughout this manuscript, consider the corresponding metric space $(\mathcal{M}, \rho_j)$ with the distance ρ_j being the chord distance defined in (1.2).*

Example 3. *The unit sphere $\mathbb{S}^m$ is a already embedded in $\mathbb{R}^{m+1}$ and the embedding is given by the inclusion, $\iota : \mathbb{S}^m \to \mathbb{R}^{m+1}$ given by $\iota(x) = x, \;\; \forall x \in \mathbb{S}^m$ with usual Euclidean metric $\rho_0^2(x, y) = \|x - y\|^2$*

Example 4. *The projective space $\mathbb{R}P^m$ is embedded in the space of symmetric $(m+1) \times (m+1)$ matrices, via the Veronese-Whitney embedding*

$$\begin{aligned} j : \mathbb{R}P^m &\to \mathcal{S}(m+1, \mathbb{R}), \\ j([x]) &= xx^T \end{aligned} \tag{2.6}$$

with the following metric on $Sym(m+1)$ given by $\rho_0^2(A, B) = Tr((A - B)^2)$, where Tr denotes the trace of the matrix $(A - B)^2$. (see Patrangenaru and Ellingson (2015) [21]) and Crane and Patrangenaru (2011) [7])

The definition below will allow us to set up a correspondence between a basis of tangent vectors in $T_p\mathcal{M}$ and an m-tuple of linearly independent tangent vectors in $T_j(p)\mathbb{R}^k$.

DEFINITION 2.1.14. *(Adapted frame field)*
Assume $p \to (f_1(p), ..., f_m(p))$ is a local frame field on an open subset of $\mathcal{M}$ such that, for each $p \in \mathcal{M}$, $(d_pj(f_1(p)), ..., d_pj(f_m(p)))$ are orthonormal vectors in $\mathbb{R}^k$. A local frame field $(e_1(y), ..., e_k(y))$ defined on an open neighborhood $U \subset \mathbb{R}^k$ is adapted to the embedding j if it is an orthonormal frame field and

$$e_r(j(p)) = d_pj(f_r(p)), \; r = 1, ..., m, \;\; \forall \, p \, \in \, j^{-1}(U) \tag{2.7}$$

(Patrangenaru and Ellingson (2015) [21])

2.2 Extrinsic means and sample means

The Fréchet function on a complete metric space is the main tool by which we will introduce means on embedded manifolds. It was introduced by Fréchet in 1948 [11]. Let X be a random vector from a probability measure Q on $\mathbb{R}^m$ with mean vector μ. The mean vector is also the value of $\mathbb{R}^m$ for which the expression $\mathbb{E}[\|X - p\|^2]$ (viewed as a function of p) is minimized. This function of p is none other than the Fréchet function on the metric space $(\mathbb{R}^m, \rho_0)$. Furthermore, for $X_1, ..., X_n$ iid random vectors from the distribution Q on $\mathbb{R}^m$ the sample mean is given by $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ with $\bar{X} \to_p \mu$. One thing we must note is that in the case of probability measures on Euclidean spaces we can easily estimate asymptotically the true mean via the sample mean as defined above. This will not be the case for most metric spaces we will encounter such as the sphere and the projective space, until we have a notion of mean, that is also a point on such object spaces . We must hence revisit the definition of the mean and sample mean and it will start with us thinking of it solely as the minimizer of some function, called Fréchet function. We will later give a more general definition of a Fréchet function but first we must mention that for this section, the reader may assume that a definition, an example, a theorem, property and most results can be found in the book by Patrangenaru and Ellingson (2015) [21].

2.2.1 Extrinsic mean

Let $\mathcal{M}$ be an m-dimensional manifold and let $\mathcal{B}_{\mathcal{M}}$ be the Borel σ-algebra generated by open sets of $\mathcal{M}$. Let $(\Omega, \mathcal{A},\ Pr)$ be a probability space. A random object (r.o.) on $\mathcal{M}$ is a function $X : \Omega \to \mathcal{M}$, such that for any Borel set $B \in \mathcal{B}_{\mathcal{M}}$, $X^{-1}(B) \in \mathcal{A}$. To each r.o. X we associate a probability measure $Q = P_X$ on $\mathcal{B}_{\mathcal{M}}$ given by $Q(B) = Pr(X^{-1}(B))$. In general, a natural index of location for a probability measure Q associated with a r.o. X on a complete metric space $\mathcal{M}$ with the distance metric ρ is the *Fréchet mean*. It is the unique minimizer of the *Fréchet function* (see Fréchet(1948) [11]), defined by

$$\mathcal{F}(p) = \mathbb{E}\left[\rho^2(p,x)\right] = \int \rho^2(p,x)\, Q(dx), \tag{2.8}$$

whenever such a unique minimizer exists. Generally two types of distance on a manifold $\mathcal{M}$ are considered:

1. A geodesic distance or arc distance. It is the Riemannian distance ρ_g associated with Riemannian structure g on $\mathcal{M}$.
2. A chord distance ρ_j associated with an embedding $j : \mathcal{M} \to \mathbb{R}^k$. (see Patrangenaru and Ellingson (2015) [21])

These two distances give rise to two types of statistical analysis on manifolds: an intrinsic analysis using an arc distance and an extrinsic analysis based on a chord distance. We will focus on the latter.
From this point on, we will assume that $(\mathcal{M}, \rho_j)$ is a complete metric space.

DEFINITION 2.2.1. *Let Q be a probability measure on $\mathcal{M}$ with a distance ρ_j. If $\mathcal{F}$ in* (2.8) *has a unique minimizer, this minimizer is called the extrinsic mean of Q and it is denoted $\mu_{j,E}(Q)$ or simply μ_E. If the minimizer is not unique, the set of all minimizers is the extrinsic mean set.*
(see Patrangenaru and Ellingson (2015) [21])

DEFINITION 2.2.2. *Let $X_1, X_2, ..., X_n$ be independent random variables with a common distribution Q on the metric space $(\mathcal{M}, \rho_j)$, and consider their empirical distribution $\hat{Q}_n = \dfrac{1}{n}\sum_{k=1}^{n} \delta(X_k)$.*
The extrinsic sample mean (set) is the extrinsic mean (set) of $\hat{Q}_n$ i.e. the (set of) minimizer(s) $\hat{p}$ of $\mathcal{F}_n(p) = \frac{1}{n}\sum_{j=1}^{n} \rho_j^2(X_j, p)$. (see Patrangenaru and Ellingson (2015) [21])

DEFINITION 2.2.3. *Assume ρ_0 is the Euclidean distance in $\mathbb{R}^k$. A point x of $\mathbb{R}^k$ such that there is a unique point p in $\mathcal{M}$ for which $\rho_0(x, j(\mathcal{M})) = \rho_0(x, j(p))$ is called j-nonfocal. A point which is not j-nonfocal is said to be j-focal.(see Patrangenaru and Ellingson (2015) [21])*

The only focal point of S^m with the inclusion in $\mathbb{R}^{m+1}$ is 0_{m+1}. Note that the probability measure Q induces a probability measure $j(Q)$ on $\mathbb{R}^k$.

DEFINITION 2.2.4. *A probability measure Q on $\mathcal{M}$ is said to be j-nonfocal if the mean μ of $j(Q)$ is a j-nonfocal point. If x is a j-nonfocal point, its projection on $j(\mathcal{M})$ is the unique point $y = P_j(x) \in j(\mathcal{M})$ with $\rho_0(x, j(M)) = \rho_0(x, y)$.(see Patrangenaru and Ellingson (2015) [21])*

THEOREM 2.2.1. *If μ is the mean of $j(Q)$ in $\mathbb{R}^k$, Then*

(a) the extrinsic mean set is the set of all points $p \in \mathcal{M}$, with $\rho_0(\mu, j(p)) = \rho_0(\mu, j(\mathcal{M})$ and

(b) If $\mu_{j,E}(Q)$ exists then μ exists and is j-nonfocal and $\mu_{j,E}(Q) = j^{-1}(P_j(\mu))$.

(see Patrangenaru and Ellingson (2015) [21])

THEOREM 2.2.2. *The set of focal points of a sub-manifold $\mathcal{M}$ of $\mathbb{R}^k$ is a closed subset of $\mathbb{R}^k$ of measure 0. (Patrangenaru and Ellingson (2015) [21])*

The 2D sphere and the 3D projective space are manifolds of interest to us. Their extrinsic means will appear and be used at various points in our study.

Example 5. *(Spheres) Lets assume that we have a random object X from a j-nonfocal probability measure Q on $S^m = \{x \in \mathbb{R}^{m+1} : \|x\| = 1\}$ an m-dimensional sphere. For this particular space, the j-nonfocal condition which guarantees the existence of a* **unique** *extrinsic mean is equivalent to requiring that the true mean $\mu_{\iota E} \neq 0 \in \mathbb{R}^{m+1}$.*
The embedding and its corresponding projection are two functions that are essential in finding and expressing our extrinsic mean. For S^m the embedding is the inclusion map $\begin{cases} \iota : S^m \to \mathbb{R}^{m+1} \\ \iota(x) = x \end{cases}$ and the projection

map is $\begin{cases} P_\iota : \mathcal{F}^c \to \iota(S^m) \\ P_\iota(y) = \frac{y}{\|y\|} \end{cases}$ *where* $\mathcal{F}^c = \mathbb{R}^{m+1}\backslash\{0\}$ *is the set of* ι*-nonfocal points in* $\mathbb{R}^{m+1}$*. Now, if* μ *is the mean of* $\iota(Q)$ *then the extrinsic mean is given by*

$$\mu_{\iota E} = \iota^{-1}\left(P_\iota(\mu)\right) = \frac{\mu}{\|\mu\|} \tag{2.9}$$

Example 6. *(Real projective spaces) We now assume that* $[X]$ *is a random object from a* j*-nonfocal probability measure* Q *on* $\mathbb{R}P^m$*. Much like in the example above we must have a clear expression of an embedding and its corresponding projection and for real projective spaces the embedding of choice is the* **VW** *(Veronese-Whitney) embedding mentioned in* (2.6)*. With this choice of embedding*

(i) The set $\mathcal{F}$ *of focal points of* $j(\mathbb{R}P^m) \in S_+(m+1,\ \mathbb{R})$ *is the set of matrices in* $S_+(m+1,\ \mathbb{R})$*(space of positive semi-definite symmetric matrices) whose largest eigenvalues are of multiplicity at least 2.*

(ii) The projection $P_j : S_+(m+1,\ \mathbb{R})\backslash\mathcal{F} \to j(\mathbb{R}P^m)$ *assigns to each positive semi-definite matrix* A *with a highest eigenvalue of multiplicity 1, the matrix* $j([m])$*, where* m *is a unit eigenvector of* A *corresponding to its largest eigenvalue.(see [6] or [21].)*

If $X^TX = 1$*, and in the ambient space the mean* $\mu = E\left[XX^T\right]$ *exists, then the VW mean is*

$$\begin{aligned} \mu_{jE} &= j^{-1}(P_j(\mu)) = j^{-1}(j([\gamma(m+1)])) \\ \mu_{j,E} &= [\gamma(m+1)] \end{aligned} \tag{2.10}$$

where $\lambda(a)$ *and* $\gamma(a),\ a = 1, \cdots, m+1$ *are eigenvalues in increasing order and corresponding eigenvectors of* $E\left[XX^T\right]$ *. (see Patrangenaru and Ellingson (2015) [21])*

In particular:

Example 7 (Extrinsic sample means for S^m and $\mathbb{R}P^m$.)**.** *(i) Assume* Q *is a nonfocal probability measure on the manifold* S^m *and* $X = \{X_1, ..., X_n\}$ *are i.i.d.r.o's from* Q*. Then the extrinsic sample mean is given by*

$$\overline{X}_{\iota\ n} = \frac{\bar{X}_n}{\|\bar{X}_n\|} \tag{2.11}$$

where $\bar{X}_n = \frac{1}{n}\Sigma_{i=1}^n X_i$

(ii) Now let Q *be V-W nonfocal probability measure on the manifold* $\mathbb{R}P^m$ *and* $[X] = \{[X_1], ..., [X_n]\}$ *are i.i.d.r.o's from* Q*. Then the V-W sample mean is given by;*

$$\overline{[X]}_{j\ n} = [g(m+1)] \tag{2.12}$$

where $d(a)$ and $g(a),\ a = 1, \cdots, m+1$ are eigenvalues in increasing order and corresponding unit eigenvectors of $J = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^T$

(Patrangenaru and Ellingson (2015) [21])

PROPOSITION 2.2.1. *Consider an embedding $j : \mathcal{M} \to \mathbb{R}^k$. Assume $(X_1, ..., X_n)$ is a random sample from a j-nonfocal probability measure Q on $\mathcal{M}$, and the sample mean vector $(j(\bar{X}))$ is j-nonfocal. Then this extrinsic sample mean is given by*

$$\overline{X}_E = j^{-1}\left(P_j(j(\bar{X}))\right) \tag{2.13}$$

(see Patrangenaru and Ellingson (2015) [21])

Remark: At this point it is important to note that for an embedded smooth manifold $\mathcal{M}$ into $j(\mathcal{M}) \subset \mathbb{R}^k$, one can analyze data from an unknown probability distribution Q, with help of the various widely known multivariate techniques and conduct inferences for extrinsic means, variances, etc.

THEOREM 2.2.3. *Assume Q is a j-nonfocal probability measure on the manifold $\mathcal{M}$ and $X = \{X_1, ..., X_n\}$ are i.i.d.r.o's from Q, then the extrinsic sample mean $\overline{X}_E$ is a strongly consistent estimator of the $\mu_{j,E}(Q)$. (see Patrangenaru and Ellingson (2015) [21])*

2.3 Central limit theorem for extrinsic sample means

A Central Limit Theorem for extrinsic sample means was given in Bhattacharya and Patrangenaru(2005)[6]. Let's assume Q is a j-nonfocal probability measure on the manifold $\mathcal{M}$ and $X = \{X_1, ..., X_n\}$ are i.i.d.r.o's from Q. Consider the embedded random variables $j(X) = \{j(X_1), ..., j(X_n)\}$ as random vectors from the probability measure $j(Q)$ with mean vector μ and assume $j(Q)$ has finite moments of order four. We can apply the usual (multivariate) Central Limit Theorem for our sample of embedded random objects and get the following convergence in distribution:

$$n^{1/2}\left(\overline{j(X)} - \mu\right) \to_d N(0,\ \Sigma) \tag{2.14}$$

where $\overline{j(X)} = \frac{1}{n}\sum_{i=1}^{n} j(X_i)$. Given the formula of the extrinsic sample mean, we will need to understand the asymptotic behavior of $P_j(\overline{j(X)}) = j(\overline{X}_{j,E})$. We do so by relying on the following theorem.

THEOREM 2.3.1 (Cramer's Delta Method)**.** *Let $Y_j,\ j \geq 1$ be i.i.d k-dimensional random vectors with mean vector μ and covariance matrix $\Sigma = (\sigma_{ij})$. For $H : \mathbb{R}^k \to \mathbb{R}^p$ a vector-valued and continuously differentiable function in a neighborhood of μ we have the following asymptotic behavior*

$$\sqrt{n}[H(\bar{Y}) - H(\mu)] \to_d D_\mu H \cdot V \sim N_p\left(0,\ D_\mu H\ \Sigma\ D_\mu H^T\right) \tag{2.15}$$

with $D_\mu H = \left(\frac{\partial H^j(z)}{\partial x_i}\Big|_{z=\mu} \right)_{i=1,\cdot,k;j=1,\cdot,p}$ *(see Patrangenaru and Ellingson (2015) [21], Theorem 2.8.5)*

Using the Cramer's Delta method for the real-valued and continuously differentiable function P_j we get the following for the random vectors $j(X) = \{j(X_1), ..., j(X_n)\}$

$$n^{1/2}\left(P_j(\overline{j(X)}) - P_j(\mu)\right) \to_d D_\mu P_j \cdot V \sim N_k\left(0,\ \Sigma_\mu\right), \tag{2.16}$$

where $\Sigma_\mu = D_\mu P_j\ \Sigma\ D_\mu P_j^T$. Here $P_j : \mathcal{F}^c \to j(\mathcal{M})$ where $\mathcal{F}$ is the set of focal points in $j(\mathcal{M})$. Note that since $\mathcal{F}$ is a closed subset of $\mathbb{R}^k$ thus $\mathcal{F}^c$ is an open subset of $\mathbb{R}^k$ a smooth k-manifolds and is itself a smooth k-manifold. Let $e_1, e_2, ..., e_n$ be the canonical basis of $\mathbb{R}^k$ and assume that $(e_1(y), ..., e_k(y))$ is an adapted frame field around $P_j(\mu) = j(\mu_E)$ i.e $e_r(P_j(\mu)) = e_r(j(\mu_E)) = d_{\mu_E} j(f_r(p)),\ r = 1, \ldots, m$ where $p \to (f_1(p), \ldots, f_m(p)$ is our local frame field of interest. Then $d_\mu P_j(e_b) \in T_{P_j(\mu)} j(\mathcal{M})$ and we can now represent this vector as a linear combination of $e_1(P_j(\mu)), ..., e_m(P_j(\mu)) \in T_{P_j(\mu)}\mathbb{R}^k$;

$$d_\mu P_j(e_b) = \sum_{a=1}^{m} [d_\mu P_j(e_b) \cdot e_a(P_j(\mu))]\, e_a(P_j(\mu)),\ \ \forall\, b = 1, ..., k \tag{2.17}$$

$$d_\mu P_j(e_b) = \sum_{a=1}^{m} \alpha_{a,b}\ e_a(P_j(\mu)) \ \text{ where } \alpha_{a,b} = [d_\mu P_j(e_b) \cdot e_a(P_j(\mu))]$$

Recall that using Cramer's Delta Method we have that $n^{1/2}\left(P_j(\overline{j(X)}) - P_j(\mu)\right)$ converges weakly to a random vector $D_\mu P_j \cdot V \approx \mathcal{N}_k(0, \Sigma_\mu)$, with $\Sigma_\mu = D_\mu P_j\ \Sigma\ D_\mu P_j^T$ where Σ is the covariance matrix of $j(X_1)$ w.r.t the canonical basis $e_1, ..., e_k$. We can now express our covariance matrix Σ_μ using the new representation of vectors $d_\mu P_j(e_b),\ \ \forall\, b = 1, ..., k$

$$\Sigma_\mu = \left[\sum_{a=1}^{m} \alpha_{a,b}\ e_a(P_j(\mu))\right]_{b=1,...,k} \Sigma\ \left[\sum_{a=1}^{m} \alpha_{a,b}\ e_a(P_j(\mu))\right]^T_{b=1,..,k} \tag{2.18}$$

And note that

$$d_\mu P_j(e_b) \cdot e_a(P_j(\mu)) = 0, \quad \text{for }\ a = m+1, ..., k$$

It is important to remember that $n^{1/2}\left(P_j(\overline{j(X)}) - P_j(\mu)\right)$ is a vector in $\mathbb{R}^k$ with origin at $P_j(\mu) = j(\mu_E)$ and as such it can be decomposed into component in the tangent space $T_{j(\mu_E)}j(\mathcal{M})$ and component of the orthogonal complement of the tangent space at $j(\mu_E)$. If we take the component in the tangent space then asymptotic distribution we obtain is a distribution on $T_{P_j(\mu)}j(\mathcal{M})$, a linear space. To illustrate this point we start by defining tangential components which corresponds to tangent vectors in $T_p\mathbb{R}^k$ and are dependent on the choice of basis elements of the tangent space of interest.

DEFINITION 2.3.1. *The tangential component* $\tan(\nu)$ *of* $\nu \in \mathbb{R}^k$ *w.r.t. the basis* $e_a(P_j(\mu)) \in T_{P_j(\mu)}j(\mathcal{M}), a = 1, 2, ..., m$ *given by*

$$\tan(\nu) = \begin{bmatrix} e_1(P_j(\mu))^T \\ \vdots \\ e_m(P_j(\mu))^T \end{bmatrix} \nu = [e_1(P_j(\mu)) \cdot \nu, \ldots, e_m(P_j(\mu)) \cdot \nu]^T \tag{2.19}$$

(Patrangenaru and Ellingson (2015) [21])

We now get the following asymptotic for the tangential component of $P_j\left(\overline{j(X)}\right) - P_j(\mu)$

$$n^{1/2}\ \tan_{j(\mu_E)}\left(P_j\left(\overline{j(X)}\right) - P_j(\mu)\right) \to_d \mathcal{N}_m(0, \Sigma_{j,E}) \tag{2.20}$$

where

$$\Sigma_{j,E} = A^T\Sigma_\mu A = \begin{bmatrix} e_1(P_j(\mu))^T \\ \vdots \\ e_m(P_j(\mu))^T \end{bmatrix} \Sigma_\mu \begin{bmatrix} e_1(P_j(\mu)) & \cdots & e_m(P_j(\mu)) \end{bmatrix} \tag{2.21}$$

The tangential component of $P_j\left(\overline{j(X)}\right) - P_j(\mu)$ is a tangent vector in $T_{j(\mu_E)}j(\mathcal{M})$ and therefore its corresponding random vector $(d_{\mu_E}j)^{-1}\tan(P_j\left(\overline{j(X)}\right) - P_j(\mu)) \in T_{\mu_E}\mathcal{M}$ converges asymptotically to a multivariate normal with mean vector 0 and covariance matrix w.r.t. the basis $f_1(\mu_E), ..., f_m(\mu_E)$ given by

$$\Sigma_{j,E} = (A^T D_\mu P_j)\ \Sigma\ (A^T D_\mu P_j)^T \tag{2.22}$$

where under the new basis

$$(A^T D_\mu P_j)_{ab} = [d_\mu P_j(e_b) \cdot e_a(P_j(\mu))] = \begin{bmatrix} d_\mu P_j(e_1) \cdot e_1(P_j(\mu)) & \ldots & d_\mu P_j(e_m) \cdot e_1(P_j(\mu)) \\ \vdots & \ddots & \vdots \\ d_\mu P_j(e_1) \cdot e_m(P_j(\mu)) & \ldots & d_\mu P_j(e_m) \cdot e_m(P_j(\mu)) \end{bmatrix} \tag{2.23}$$

DEFINITION 2.3.2. *The matrix $\Sigma_{j,E}$ given by* (2.22) *is the extrinsic covariance matrix of the j-nonfocal distribution Q of X_1 w.r.t. the basis $f_1(\mu_E), ..., f_m(\mu_E)$. When j is fixed in a specific context, the subscript j will be omitted. If in addition, Σ_E is invertible (of rank m) we can define the j-standardized mean vector*

$$\overline{Z}_{j,n} := n^{\frac{1}{2}} \Sigma_E^{-\frac{1}{2}} \left(\overline{X}_j^1 ... \overline{X}_j^m \right)^T, \tag{2.24}$$

where $\left(\overline{X}_j^1 ... \overline{X}_j^m \right)^T$ are the coordinates of the tangent component of $j(\overline{X}_{j,E}) - \jmath(\mu_{j,E}(Q))$, w.r.t the basis $e_a(P_j(\mu)) \in T_{P_j(\mu)} j(\mathcal{M}), a = 1, 2, ..., m$. (Patrangenaru and Ellingson (2015) [21])

PROPOSITION 2.3.1. *Assume $\{X_r\}_{r=1}^n$ are i.i.d.r.o's from the j-nonfocal distribution Q, with finite mean $\mu = E(j(X_1))$, and assume the extrinsic covariance matrix $\Sigma_{j,E}$ of Q is finite. Let $(e_1(y), ..., e_k(y))$ be an orthonormal frame field adapted to j. Then*

(a) the tangential component of the difference between $j(\overline{X}_{j,E})$ and the $\jmath(\mu_{j,E}(Q))$ has asymptotically a distribution that is approximately multivariate normal the tangent space to $\mathcal{M}$ at $\mu_{j,E}(Q)$ with mean 0 and covariance matrix $n^{-1}\Sigma_{j,E}$. and

(b) if $\Sigma_{j,E}$ is nonsingular, the standardized mean vector $\overline{Z}_{j,n}$ given in (2.24) *converges weakly to a $\mathcal{N}_m(0_m, I_m)$-distributed random vector.*

(Patrangenaru and Ellingson (2015) [21])

The CLT for extrinsic sample means stated in Proposition 2.3.1 cannot be used to construct confidence regions for extrinsic means since the population extrinsic covariance matrix is unknown. In order to define our confidence regions we will need to have the following consistent estimator for $\Sigma_{j,E}$.

$$S_{j,E,n} = \left[d_{j(\overline{X})} P_j(e_b) \cdot e_a(P_j(j(\overline{X}))) \right]_{a=1,...,m} S_{j,n} \left[d_{j(\overline{X})} P_j(e_b) \cdot e_a(P_j(\jmath(\overline{X}))) \right]_{a=1,...,m}^T \tag{2.25}$$

is a consistent estimator of $\Sigma_{j,E}$. With

$$S_{j,n} = n^{-1} \sum_{r=1}^{n} \left(j(X_r) - j(\overline{X}) \right) \left(j(X_r) - j(\overline{X}) \right)^T \tag{2.26}$$

a consistent estimator of Σ the covariance matrix of $j(X_1)$ and $d_{j(\overline{X})} P_j(e_b)$ consistent estimator of $d_\mu P_j(eb)$ and $e_a(P_j(j(\overline{X})))$ a consistent estimator of $e_a(P_j(\mu))$.(see Bhattacharya and Patrangenaru [6] also Patrangenaru and Ellingson (2015) [21]).

THEOREM 2.3.2. *Assume $j : \mathcal{M} \to \mathbb{R}^k$ is a closed embedding of $\mathcal{M}$ in $\mathbb{R}^k$. Let $\{X_r\}_{r=1}^n$ be a random sample from the j-nonfocal distribution Q, and let $\mu = \mathbb{E}[j(X_1)]$ and assume $j(X_1)$ has finite second order moments and the extrinsic covariance matrix $\Sigma_{j,E}$ of X_1 is nonsingular. Let $(e_1(y), ..., e_k(y))$ be an orthonormal frame field adapted to j. If $S_{j,E,n}$ is given by (2.25), then for n large enough $S_{j,E,n}$ is nonsingular (with probability converging to one) and*

(a) the statistic

$$n^{\frac{1}{2}} S_{j,E,n}^{-\frac{1}{2}} \ \tan(P_j(\overline{j(X)}) - P_j(\mu) \tag{2.27}$$

converges weakly to $\mathcal{N}_m(0_m, I_m)$, so that

$$n \left\| S_{j,E,n}^{-\frac{1}{2}} \ \tan(P_j(\overline{j(X)}) - P_j(\mu) \right\|^2 \tag{2.28}$$

converges weakly to χ^2_m and

(b) the statistic

$$n^{\frac{1}{2}} S_{j,E,n}^{-\frac{1}{2}} \ \tan_{P_j(\overline{j(X)})}(P_j(\overline{j(X)}) - P_j(\mu) \tag{2.29}$$

converges weakly to $\mathcal{N}_m(0_m, I_m)$, so that

$$n \left\| S_{j,E,n}^{-\frac{1}{2}} \ \tan_{P_j(\overline{j(X)})}(P_j(\overline{j(X)}) - P_j(\mu) \right\|^2 \tag{2.30}$$

converges weakly to χ^2_m and

(Patrangenaru and Ellingson (2015) [21])

COROLLARY 2.3.1. *Under the hypothesis of Theorem (2.3.2) , a confidence region for μ_E of asymptotic level $1-\alpha$ is given by*

(a) $C_{n,\alpha} = j^{-1}(U_{n,\alpha})$ where $U_{n,\alpha} = \{P_j(\mu) \in j(\mathcal{M}) : \ n \left\| S_{j,E,n}^{-\frac{1}{2}} \ \tan\left(P_j(\overline{j(X)}) - P_j(\mu)\right) \right\|^2 \leq \chi^2_{m,1-\alpha}\}$ or by

(b) $D_{n,\alpha} = j^{-1}(V_{n,\alpha})$ where $V_{n,\alpha} = \{P_j(\mu) \in j(\mathcal{M}) : \ n \left\| S_{j,E,n}^{-\frac{1}{2}} \ \tan_{P_j(\overline{j(X)})}\left(P_j(\overline{j(X)}) - P_j(\mu)\right) \right\|^2 \leq \chi^2_{m,1-\alpha}\}$

(Patrangenaru and Ellingson (2015) [21])

For small samples, we use nonparametric bootstrap confidence regions. Now lets recall that if $\{X_r\}_{r=1}^n$ is a random sample from an unknown distribution Q, and $\{X_r^*\}_{r=1}^n$ is a (bootstrap) random sample from the empirical distribution $\hat{Q}_n$, conditionally given by $\{X_r\}_{r=1}^n$, then the statistic in Theorem 2.3.2 (a),

$$T(X,Q) = n \left\| S_{j,E,n}^{-\frac{1}{2}} \ \tan(P_j(\overline{j(X)}) - P_j(\mu) \right\|^2 \tag{2.31}$$

has the bootstrap analog

$$T(X^*,Q) = n \left\| {S^*}_{j,E,n}^{-\frac{1}{2}} \ \tan_{P_j(\overline{j(X)})}(P_j(\overline{j(X^*)}) - P_j(\overline{j(X)}) \right\|^2 \tag{2.32}$$

Where $T(X^*,Q)$, ${S^*}_{j,E,n}$ is obtained by substituting $\{X_r\}_{r=1}^n$ by $\{X_r^*\}_{r=1}^n$ and also by replacing μ by $\overline{j(X)}$. From this point on, we will assume that $j(Q)$, , has finite moment of sufficiently high order. This result is automatic for compact manifolds such as S^m and $\mathbb{R}P^m$. The following theorem addresses the order of convergence related to our bootstrap statistic.

THEOREM 2.3.3. *Let $\{X_r\}_{r=1}^n$ be a random sample from he j-nonfocal distribution Q which has a nonzero absolutely continuous component w.r.t. the volume measure on $\mathcal{M}$ induced by j. Let $\mu = E[j(X_1)]$ and assume the covariance matrix Σ of $j(X_1)$ is defined and the extrinsic covariance matrix $\Sigma_{j,E}$ is non-singular and let $p \to (e_1(p), \dots, e_N(p))$ an orthonormal frame field adapted to j. Then the distribution of*

$$n \left\| S_{j,E,n}^{-\frac{1}{2}} \ \tan(P_j(\overline{j(X)}) - P_j(\mu) \right\|^2$$

can be approximated by the bootstrap extrinsic Hotelling distribution of

$$n \left\| {S^*}_{j,E,n}^{-\frac{1}{2}} \ \tan_{P_j(\overline{j(X)})}(P_j(\overline{j(X^*)}) - P_j(\overline{j(X)}) \right\|^2$$

with a coverage error $O_p(n^{-2})$. (Patrangenaru and Ellingson (2015) [21])

We will encounter cases when $S_{j,E,n}$ is difficult to compute and for such situations,we will rely on the following result.

PROPOSITION 2.3.2. *on the asymptotic distribution of $n \ \left\| \tan(P_j(\overline{j(X)}) - P_j(\mu) \right\|^2$ can be approximated uniformly by the bootstrap distribution of*

$$n \ \left\| \tan(P_j(\overline{j(X^*)}) - P_j(\overline{j(X)}) \right\|^2 \tag{2.33}$$

to provide a confidence region for μ_E with coverage error no more than $O_p(n^{-\frac{m}{m+1}})$. (see Patrangenaru and Ellingson (2015) [21])

REMARK 2.3.1. *For bootstrap confidence regions in Theorem 2.3.3 the bootstrap analog of Corollary 6.2.1 (a) is preferable. The corresponding* $100(1-\alpha)\%$ *confidence region is* $C^*_{n,\alpha} := j^{-1}(U^*_{n,\alpha})$ *with* $U^*_{n,\alpha}$ *given by*

$$U^*_{n,\alpha} = \{P_j(\nu) \in j(\mathcal{M}) : n\|S^{-1/2}_{j,E,n}\tan(P_j(\overline{j(X)}) - P_j(\mu)\|^2 \leq c^*_{1-\alpha}\}, \tag{2.34}$$

where $c^*_{1-\alpha}$ *is the upper* $100(1-\alpha)\%$ *point of the values*

$$\|S^{*}{}^{-1/2}_{j,E,n}\tan_{P_j(\overline{j(X)})}(P_j(\overline{j(X^*)}) - P_j(\overline{j(X)}\|^2 \tag{2.35}$$

among the bootstrap re samples. And the region given by 2.34 has a coverage error $O_p(n^{-2})$.

2.4 Projective shape space

The bulk of our analysis will directly involve $P\Sigma^k_3$ the 3D projective shape space of k-ads (landmarks) in general position. We will conduct a landmark based analysis which will involve recovering the 3D coordinates of our labeled points.

2.4.1 Representation of projective shapes

We associate a *shape* to a configuration of k labeled points. We are interested in conducting our analysis on *projective shapes* but first we start with defining the a projective transformation of elements in a Euclidean space.

DEFINITION 2.4.1. *Generally, a projective transformation* ν *of* $\mathbb{R}^m$ *is defined in terms of a matrix* $A = (a^j_i) \in GL(m+1,\mathbb{R})$, *via* $\nu(x^1,\dots,x^m) = (y^1,\dots,y^m)$,

$$y^j = \frac{\sum_{i=1}^m a^j_i x^i + a^j_{m+1}}{\sum_{i=1}^m a^{m+1}_i x^i + a^{m+1}_{m+1}} = \frac{A^j \cdot \mathbf{u}}{A^{m+1}\cdot \mathbf{u}},\ \forall j = 1,\dots,m. \tag{2.36}$$

where A^j *is the* j*-th column of* A *and* $\mathbf{u} = (x^1,\dots,x^m,1)^T$.

(Patrangenaru and Ellingson (2015) [21])

REMARK 2.4.1. *Two configurations of points in* $\mathbb{R}^m$ *have the same* 3D shape *if they differ by a projective transformation of* $\mathbb{R}^3$. *However, in applications, such projective transformations act only on subsets of* $\mathbb{R}^3$ *and consequently they do not have a group structure under composition.*

Note that if one multiplies the matrix A by a nonzero constant, then the equation (2.36) does not change; therefore the group $PGL(m)$ of projective transformations of $\mathbb{R}^m$ has dimension $(m+1)^2 - 1 = m(m+2)$. Furthermore, $\mathbb{R}^m$ can be identified with an open affine subset of $\mathbb{R}P^m$, any configuration of points $\{x_1, \ldots, x_k\}$ in $\mathbb{R}^m$ can be regarded as a configuration projective points $\{p_1, \ldots, p_k\}$ in $\mathbb{R}P^m$. An example of such an identification is the affine embedding $h : \mathbb{R}^m \to \mathbb{R}P^m$ given by

$$h(x) = h((x^1, \ldots, x^m)) = [x^1 : \cdots : x^m : 1] \tag{2.37}$$

(see Patrangenaru and Qiu (2014) [25]).

The pseudo group action by projective transformations on open dense subsets of $\mathbb{R}^m$ is extended to a group action of the projective group $PGL(m)$. And the group action is given by

$$\begin{aligned} &\alpha : PGL(m) \times \mathbb{R}P^m \to \mathbb{R}P^m \\ &\alpha([A], [x]) = [Ax],\ \forall A\ \in GL(m+1, \mathbb{R}),\ \forall\, x \in \mathbb{R}^{m+1} \end{aligned} \tag{2.38}$$

Note that given the matrix A in the projective transformation ν in 2.36 and $\mathbf{u}$ we have the following vector $\tilde{\mathbf{u}} = A\mathbf{u} = ((A^1 \cdot \mathbf{u}), \ldots, (A^m \cdot \mathbf{u}), (A^{m+1} \cdot \mathbf{u}))^T$ we now get the following equality

$$[A\mathbf{u}] = [\tilde{\mathbf{u}}^{\mathbf{1}} : \cdots : \tilde{\mathbf{u}}^{\mathbf{m}} : \tilde{\mathbf{u}}^{\mathbf{m+1}}] = \left[\frac{\tilde{\mathbf{u}}^{\mathbf{1}}}{\tilde{\mathbf{u}}^{\mathbf{m+1}}} : \cdots : \frac{\tilde{\mathbf{u}}^{\mathbf{m}}}{\tilde{\mathbf{u}}^{\mathbf{m+1}}} : \mathbf{1} \right] \tag{2.39}$$

where $\frac{\tilde{\mathbf{u}}^i}{\tilde{\mathbf{u}}^{m+1}} = y^i$ for $i = 1, \ldots, m$. And we refer to $(y^1, \ldots, y^m)$ as the inhomogeneous (affine) coordinates of the point $[\tilde{\mathbf{u}}] \in \mathbb{R}P^m$.

Therefore, rather then considering projective shapes of configurations in $\mathbb{R}^m$ we consider projective shapes of configurations in the projective space $\mathbb{R}P^m$.

DEFINITION 2.4.2. *Two sets of labeled points* $\{[x_{a,1}], \ldots, [x_{a,k}]\} \subset \mathbb{R}P^m,\ a = 1, 2$ *have the same projective shape if there is a projective transformation* $\beta : \mathbb{R}P^m \to \mathbb{R}P^m$*, such that* $\beta([x_{1,j}]) = [x_{2,j}],\ \forall\, j = 1, \ldots, k$. *(see Patrangenaru and Qiu (2014) [25]).*

In projective shape analysis it is preferable to employ coordinates invariant with respect to the group PGL(m). To create such coordinates we will need to use a projective frame.

DEFINITION 2.4.3. *A* projective frame $\pi = (p_1, \ldots, p_{m+2})$ *in* $\mathbb{R}P^m$ *is an ordered set of* $m+2$ *projective points in general position. Note that* k *points in* $\mathbb{R}P^m$ *are in general position if their linear span is* $\mathbb{R}P^m$. *For* $p_i,\ i = 1, \ldots, m+2$ *with the spherical representation* $p_i = \{x_i, -x_i\}$ $x_i \in \mathbb{R}^{m+1}$*, this means that for*

$\{x_1, \ldots, x_{m+2}\}$ *any subset of size* $m+1$ *form a linear span of* $\mathbb{R}^{m+1}$. *(Patrangenaru and Ellingson (2015) [21])*

An example of projective frame in $\mathbb{R}P^m$ is the *standard projective frame* $\pi_0 = ([e_1], \ldots, [e_{m+1}], [e_1 + ... + e_{m+1}])$.

PROPOSITION 2.4.1. *Given two projective frames* $\pi_1 = (p_{1,1}, \ldots, p_{1,m+2})$ *and* $\pi_2 = (p_{2,1}, \ldots, p_{2,m+2})$, *there is a unique* $\beta \in PGL(m)$ *with* $\beta(p_{1,j}) = p_{2,j},\ j = 1, \ldots, m+2$. *(see Mardia and Patrangenaru (2005) [20]).*

A projective transformation takes a projective frame to a projective frame, and its action on $\mathbb{R}P^m$ is determined by its action on a projective frame.

DEFINITION 2.4.4. *The* projective coordinate(s) *of a point* $p = [x^1 : \cdots : x^{m+1}] \in \mathbb{R}P^m$ *w.r.t. a projective frame* $\pi = (p_1, \ldots, p_{m+2})$ *as being given by*

$$p^{\pi} = \beta^{-1}(p) \tag{2.40}$$

where β *is a projective (transformation) map taking the standard projective frame* π_0 *to* π, *these coordinates have automatically the invariance property. (Patrangenaru and Ellingson (2015) [21])*

PROPOSITION 2.4.2. *Assume* $u_1, \ldots, u_k$ *are points in* $\mathbb{R}^m$. *We then identify the first* $m+2$ *points with* $\tilde{u}_1, \ldots, \tilde{u}_{m+2}$ *in* $\mathbb{R}P^3$ *where* $\tilde{u}_i = [u_1^i : u_2^i : \cdots : u_3^i : 1]$ *for* $i = 1, \ldots, m+2$. *If we consider the* $m+1$ *by* $m+1$ *matrix* $U_m = [\tilde{u}_1^T, \ldots, \tilde{u}_{m+1}^T]$, *the projective coordinate of* $[\tilde{u}]$ *with respect to* π *are given by*

$$\begin{gathered} p^{\pi} = [y^1(u)\ :\ \ldots\ : y^{m+1}(u)], \\ where\ \ y^i(u) = \frac{v^i(u)}{v^i(u_m+2)}\ \ with\ v(u) = U_m^{-1}\tilde{u}^T \end{gathered} \tag{2.41}$$

(Patrangenaru and Ellingson (2015) [21])

DEFINITION 2.4.5. *A* projective shape *of a k-ad (configuration of k labeled points) is the orbit of that k-ad under projective transformations. If the k-ad is regarded as a point on* $(\mathbb{R}P^m)^k$, *then such a transformation acts at the same time on each point of the k-ad; therefore the action of PLG*(m) *is the diagonal action of this group on* $(\mathbb{R}P^m)^k$,

$$\alpha_k(p_1, ..., p_k) = (\alpha(p_1), ..., \alpha(p_k))$$

(Patrangenaru and Ellingson (2015) [21])

Now, lets consider the set $G(k, m)$ of k-ads $(p_1, ..., p_k)$ with $k > m + 2$ for which $\pi = (p_1, ..., p_{m+2})$ is a projective frame. Once the first $m + 2$ points are used to create a projective frame, we now use the remaining projective coordinates $(p^\pi_{m+3}, \ldots, p^\pi_k)$ to uniquely represent our projective shape of k-ads with respect to its projective frame π. The m-dimensional projective shape space of a *generic k-ad* is determined by the *projective coordinates* $(p^\pi_{m+3}, ..., p^\pi_k)$ of $k - m - 2$ of its points, relative to other $(m + 2)$ of its points that form a projective frame. Using the projective coordinates $(p^\pi_{m+3}, \ldots, p^\pi_k)$ on can show that $P\Sigma^k_m$ is a manifold diffeomorphic to $(\mathbb{R}P^m)^{k-m-2}$. The drawback of this representation is that the resulting analysis may depend on the projective frame selection. But on the other hand the projective shape space has a manifold structure allowing us to use the asymptotic theory for means on manifolds we introduced in the previous subsections.

REMARK 2.4.2. *We will now use interchangeably the notation $P\Sigma^k_m$ and $(\mathbb{R}P^m)^{k-m-2}$ to refer to the projective shape space of k-ads in m-dimensions. Furthermore, we will now represents an element $\mathbf{y} \in P\Sigma^k_m$ by $\mathbf{y} = ([x_1], \ldots, [x_q])$ where $q = k - m - 2$ and $[x_i] = p^\pi_j$ is a projective coordinate with respect to $\pi = (p_1, \ldots, p_{m+2})$.*

2.4.2 VW mean and sample mean on $(\mathbb{R}P^3)^{k-5}$

We will look at samples of random projective shapes of k-ad $(k \geq 5)$ in general position including a projective frame in $\mathbb{R}P^3$. The corresponding 3D projective space of k-ad is given by $P\Sigma^k_3 = (\mathbb{R}P^3)^{k-5}$ and is an embedded manifold. The embedding of choice is the Veronese-Whitney embedding on $(\mathbb{R}P^m)^q$ with $q = k - m - 2$ and the embedding is denoted j_k. But before we formally define this map, we will recall the VW embedding on $\mathbb{R}P^m$ is defined by

$$j : \mathbb{R}P^m \to S_+(m + 1, \mathbb{R})$$
$$j([x]) = xx^T, \ \ \|x\| = 1, \ and \ x \in \mathbb{R}^{m+1}$$

j maps $\mathbb{R}P^m$ into a $\left(\frac{1}{2}(m + 1)(m + 2)\right)$-dimensional Euclidean hypersphere in the space $S(m + 1, \mathbb{R})$, where the Euclidean distance between two symmetric matrices A and B is

$$\rho_0(A, B) = Tr((A - B)^2) \tag{2.42}$$

(see Bhattacharya and Patrangenaru (2005) [6]).

PROPERTY 2.4.1. *The VW embedding on $\mathbb{R}P^m$ is an equivariant embedding. It means that the special orthogonal group $SO(m+1)$ of orthogonal matrices with determinant $+1$ acts as a group of isometries on $\mathbb{R}P^m$ and it also acts on the left on $S_+(m+1,\mathbb{R})$, the set of nonnegative definite symmetric matrices with real coefficients. This left action is given by $W \cdot A = WAW^T$ for $W \in SO(m+1)$ and $A \in S_+(m+1,\mathbb{R})$ (see Bhattacharya and Patrangenaru (2005) [6]). Also*

$$j(W \cdot [x]) = W \cdot j([x]), \ \ \forall\, W \in SO(m+1), \ \ \forall\, [x] \in \mathbb{R}P^m \tag{2.43}$$

DEFINITION 2.4.6. *The VW embedding on $(\mathbb{R}P^m)^q$ is an equivariant embedding given by*

$$\begin{aligned} j_k &: (\mathbb{R}P^m)^q \to (S_+(m+1,\mathbb{R}))^q \\ j_k(\mathbf{y}) &= (j([x_1]), ..., j([x_q])), \ \ \mathbf{y} = ([x_1], \ldots, [x_q]) \end{aligned} \tag{2.44}$$

where $[x_s] \in \mathbb{R}P^m$ for $s = 1, \ldots, q$ with $\|x_s\| = 1$ and $x_s \in \mathbb{R}^{m+1}$ and j is the VW embedding on $\mathbb{R}P^m$. This function embed the manifold $(\mathbb{R}P^m)^q$ in the Euclidean space $E = ((S(m+1,\mathbb{R}))^q, \langle\langle\ ,\ \rangle\rangle)$ with scalar product and metric given by

$$\begin{aligned} \langle\langle \mathbf{A}, \mathbf{B} \rangle\rangle &= \sum_{i=1}^{q} Tr(A_i B_i) \\ d_0^q(\mathbf{A}, \mathbf{B}) &= \sum_{i=1}^{q} Tr((A_i - B_i)^2) \end{aligned} \tag{2.45}$$

with $\mathbf{A} = (A_1, \ldots, A_q)$ and $\mathbf{B} = (B_1, \ldots, B_q)$. (see Crane and Patrangenaru (2011) [7].)

For our Extrinsic analysis we will require a definition of the projection of the VW embedding of the projective shape space.

DEFINITION 2.4.7. *Let $\mathcal{F}^q \subset (S_+(m+1,\mathbb{R}))^q$ be the set of focal points of $j_k\,((\mathbb{R}P^m)^q)$, the projection $P_{j_k} : (S_+(m+1,\mathbb{R}))^q \backslash \mathcal{F}^q \to j_k\,(\mathbb{R}P^m)^q)$ is given by*

$$P_{j_k}(\mathbf{A}) = (P_j(A^1), ..., P_j(A^q)) = j_k([m_1], ..., [m_q]) \tag{2.46}$$

where for $i = 1, \ldots, q$ the projection $P_j : S_+(m+1,\ \mathbb{R}) \backslash \mathcal{F} \to j(\mathbb{R}P^m)$ assigns to each positive semi-definite matrix A_i with a highest eigenvalue of multiplicity 1, the matrix $j([m_i])$, where m_i is a unit eigenvector of A_i corresponding to its largest eigenvalue. And $\mathcal{F} \subset S_+(m+1,\mathbb{R})$ is the set of focal points of $j(\mathbb{R}P^m)$. (see Crane and Patrangenaru (2011) [7].)

Now that we have properly, define an embedding j_k and its corresponding projection P_{j_k} we will introduce the Extrinsic mean and sample mean on the projective shape space.

DEFINITION 2.4.8. *Let $Y = ([X_1], \dots, [X_q])$ with be a random object from a j_k-nonfocal probability measure Q on $(\mathbb{R}P^m)^q$ where $q = k - m - 2$. The corresponding VW mean is given by*

$$\mu_{j_k} = ([\gamma_1(4)], ..., [\gamma_q(4)]) \tag{2.47}$$

$\forall\ s = 1, .., q,\ (\lambda_s(a), \gamma_s(a)),\ \ a = 1, ..., m+1$ *are eigenvalues in increasing order and corresponding eigenvectors of $E(X_s(X_s)^T)$. (see Crane and Patrangenaru (2011) [7].)*

DEFINITION 2.4.9. *Let $\{Y_r\}_{r=1}^n$ be an i.i.d. random sample defined on $(\mathbb{R}P^m)^q$ from Veronese-Whitney-nonfocal distribution $\mathcal{Q}$. The corresponding sample mean extrinsic projective shape, in the multi-axial representation, is given by*

$$\overline{Y}_{j_k,n} = ([g_1(4)], ..., [g_q(4)]) \tag{2.48}$$

where for $s = 1, \dots, q$ $(d_s(a), g_s(a)),\ \ a = 1, ..., 4$ are the eigenvalues in increasing order and corresponding eigenvectors of $J_s = \dfrac{1}{n}\sum_{r=1}^{n} X_r^s (X_r^s)^T$. (see Crane and Patrangenaru (2011) [7].)

2.4.3 Lie group structure of the 3D projective shape space

In this section we introduce a very useful feature of the 3D projective shape space under our usual projective frame representation. Unlike in other dimensions, the $3D$ real projective space $\mathbb{R}P^3$ has a *Lie group* structure. This additional property is important and will allows to perform useful binary operations we would not generally have for most smooth manifolds. we now define this group structure on manifolds.

DEFINITION 2.4.10. *A Lie group is a smooth manifold $\mathcal{G}$ that is also a group in the algebraic sense, with the property the the multiplication map $\odot$ and the inversion map $i : \mathcal{G} \to \mathcal{G}$ are both smooth. (see Lee (2002) [18]*

Note that under our spherical representation, $\mathbb{R}P^3$ is the quotient $\mathbb{S}^3/\{x \sim -x\}$ and if $x, y\ \in \mathbb{S}^3$(a group of quaternions of norm one) then if follows that the multiplication

$$[p_1] \odot [p_2] = [p_1 \cdot p_2],\ \ for\ p_1, p_2\ \in \mathbb{S}^3. \tag{2.49}$$

where $(\cdot)$ is the quaternion multiplication is a well defined Lie group multiplication on $\mathbb{R}P^3$. For more on the quaternion multiplication please refer to Crane and Patrangenaru (2011) [7]. And for $[p_i] = [x_1 : y_1 : z_1 : t_1],\ i = 1, 2$ an explicit formula for our Lie group multiplication is given by

$$[p_1] \odot [p_2] = [(t_1x_2 - x_1t_2 + y_1z_2 - z_1y_2) : (t_1y_2 - y_1t_2 + z_1x_2 - x_1z_2) \\ : (t_1z_2 - z_1t_2 + x_1y_2 - y_1x_2) : (t_1t_2 - x_1x_2 - y_1y_2 - z_1z_2)] \tag{2.50}$$

Also for $[p] = [x : y : z : t] \in \mathbb{R}P^3$ with $\|p\| = 1$, its conjugate is $[\bar{p}] = [-x : -y : -z : t] \in \mathbb{R}P^3$, the inverse map on $\mathbb{R}P^3$ is given by

$$[p]^{-1} = [\bar{p}], \tag{2.51}$$

and the identity of this Lie group is $1_{\mathbb{R}P^3} = [0 : 0 : 0 : 1]$. Recall that the projective shape space is diffeomorphic to $(\mathbb{R}P^3)^q,\ (q = k - 5)$. Therefore with this identification, $P\Sigma_3^k$ inherits a Lie group structure from the group structure of $\mathbb{R}P^3$. The Lie group multiplication in $(\mathbb{R}P^3)^q$ is given by

$$([p_1], \dots, [p_q]) \odot_q ([p'_1], \dots, [p'_q]) = ([p_1] \odot [p'_1], \dots, [p_q] \odot [p'_q]) \tag{2.52}$$

And the identity element of this group is given by

$$1_{(\mathbb{R}P^3)^q} = ([0 : 0 : 0 : 1], \dots, [0 : 0 : 0 : 1]), \tag{2.53}$$

and given $\mathbf{p} = ([p_1], \dots, [p_q])$ the inverse is

$$\mathbf{p}^{-1} = \bar{\mathbf{p}} = ([\bar{p}_1], \dots, [\bar{p}_q]) \tag{2.54}$$

(see Crane and Patrangenaru (2011) [7].)

2.5 Homogeneous spaces and two sample means tests for unmatched pairs

The benefits of an added Lie group structure have been exploited especially in hypothesis testing for two sample means of matched pairs see Crane and Patrangenaru (2011) [7]. Recall that for a large sample of observations from a matched pair (X, Y) of random vectors in $\mathbb{R}^m$, one may estimates the difference vector $D = Y - X$ to eliminate much of the influence of extraneous unit to unit variation without increasing the dimensionality. Crane and Patrangenaru extended this technique to paired r.o.'s on an embedded Lie group that is not necessarily commutative. Assuming X and Y are paired r.o.'s on a Lie group $(\mathcal{G}, \odot)$. The *change* from X to Y was defined to be $C = X^{-1} \odot Y$. And a test for no mean change from X to Y is one for the null hypothesis

$$H_0 : \mu_j = 1_{\mathcal{G}}$$

where $1_{\mathcal{G}}$ is the identity of $\mathcal{G}$ and μ_j is the extrinsic mean of C with respect to the embedding j (see Patrangenaru and Qiu (2014) [25] and Crane and Patrangenaru (2011) [7]). In Mathematical Statistics it makes sense to consider the equality of means on a smooth object space $\mathcal{M}$, with an action of a Lie group $\mathcal{G}$, only for those means that lie on the same orbit (see Patrangenaru and Ellingson (2015) [21], Chapter 3), which a good reason of considering smooth object spaces made of one orbit only.

For pairs of unmatched random objects X and Y on Lie groups we cannot use the new random object C mentioned above. To circumvent this difficulties, we look to *homogeneous spaces*.

DEFINITION 2.5.1. *(see Patrangenaru and Qiu (2014) [25])*
A left action of a group $\mathcal{G}$ on a $\mathcal{M}$, is a function $\alpha : \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ such that

$$\begin{aligned} \alpha(1_{\mathcal{G}}, x) = x, \ \forall \, x \in \mathcal{M}, \\ \alpha(g, \alpha(h, x)) = \alpha(g \odot h, x), \ \forall \, g \ \in \ \mathcal{G}, \ \forall x \ \in \ \mathcal{M} \end{aligned} \tag{2.55}$$

DEFINITION 2.5.2 (Homogeneous space)**.** *(see Patrangenaru and Qiu (2014) [25])*
Assume $\alpha : \mathcal{G} \times \mathcal{M} \to \mathcal{M}$ is a left action of a group $\mathcal{G}$ on $\mathcal{M}$ and define the orbit $\mathcal{G}(x)$ of a point $x \in \mathcal{M}$ as the set $\{\alpha(k, x), k \in K\}$. Then $\mathcal{M}$ is a $\mathcal{G}$- homogeneous space if there is a point x s.t. $\mathcal{G}(x) = \mathcal{M}$.

In the case $\mathcal{M}$ is a manifold, we assume in addition that $(\mathcal{G}, \odot)$ is a Lie group and the action α is smooth. A Lie group $(\mathcal{G}, \odot)$ is automatically a $\mathcal{G}$-homogeneous space, for the action $\alpha = \odot$. Examples of object spaces that are homogeneous spaces:

- spaces of directions ($\mathcal{M} = \mathbb{S}^m, m = 1, 2$), spaces of dihedral angles ($\mathcal{M} = (\mathbb{S}^1)^k$),
- the spaces of shapes of planar k-ad's ($\mathcal{M} = \mathbb{C}P^{k-2}$. (see [16])
- spaces of shapes 2D contours ($\mathcal{M} = (P(\mathbb{H}), \mathbb{H}$ Hilbert space), spaces of cell filaments ($\mathcal{M} = \mathbb{R}P^2 \times (0, \infty)$ (see Huckemann [14].)

DEFINITION 2.5.3. *(see Patrangenaru and Qiu (2014) [25])*
$\mathcal{M}$ has a simply transitive Lie group $\mathcal{G}$, if there is a Lie group action $\alpha : \mathcal{G} \times \mathcal{M} \to \mathcal{M}$, with the property that given $x \in \mathcal{M}$, for any object $y \in \mathcal{M}$, there is a unique $g \in \mathcal{G}$ such that $\alpha(g, x) = y$.

Let $\mathcal{M}$ be a $\mathcal{G}$-homogeneous space, where $\mathcal{M}$ is an embedded manifold and $(\mathcal{G}, \odot)$ a Lie group that acts simply transitively on $\mathcal{M}$ via a smooth left action $\alpha : \mathcal{G} \times \mathcal{M} \to \mathcal{M}$. For $a = 1, 2$, let $X_{a,1}, \cdots, X_{a,n_a}$ be independent random samples defined on $\mathcal{M}$, from a distribution Q_a, with the extrinsic means $\mu_{1,j}, \mu_{2,j}$ and with the corresponding extrinsic covariance matrices $\Sigma_{1,j}, \Sigma_{2,j}$, where $j : \mathcal{M} \to \mathbb{R}^N$ is the embedding. Then, a two-sample hypothesis testing problem can be formulated as follows

$$H_0 : \mu_{1,j} = \alpha(\mu_{2,j}, \delta) \quad \text{vs.} \quad H_1 : \mu_{1,j} \neq \alpha(\mu_{2,j}, \delta),$$

for $\delta \in \mathcal{G}$. Now for a fixed object $\mu_{2,j}$ the mapping $\alpha^{\mu_{2j}} : \mathcal{G} \to \mathcal{M}, \;\; \alpha^{\mu_{2j}}(g) = \alpha(\mu_{2j}, g), \; \forall\, g \in \mathcal{G}$ is one-to-one, and we can now rewrite the hypothesis problem from above as follows

$$H_0 : (\alpha^{\mu_{2,j}})^{-1}(\mu_{1,j}) = \delta \quad \text{vs.} \quad H_1 : (\alpha^{\mu_{2,j}})^{-1}(\mu_{1,j}) \neq \delta, \tag{2.56}$$

(see Patrangenaru and Qiu (2014) [25]) We recall the following

THEOREM 2.5.1. *(see Patrangenaru and Qiu (2014) [25])*
For $a = 1, 2$, let $X_{a,1}, \cdots, X_{a,n_a}$ identically independent distributed random objects (i.i.d.r.o.'s) from the independent j_a-nonfocal probability measures Q_a with finite extrinsic moments of order $s,\ s \leq 4$ on the m dimensional manifold $\mathcal{M}$ on which the Lie group $\mathcal{G}$ acts simply transitively. Let $n = n_1 + n_2$ and assume $\lim_{n\to\infty} \frac{n_1}{n} \to \pi \in (0, 1)$. Let φ be an affine chart defined on an open neighborhood of $1_{\mathcal{G}}$ with $\varphi(1_{\mathcal{G}}) = 0_{\mathbf{g}}$, and L_δ the left translation by $\delta \in \mathcal{G}$. Then under H_0 (2.56),

(i) The sequence of random vectors $n^{1/2}\left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right)$ converges weakly to $N_m(0_m, \Sigma_J)$, for some covariance matrix Σ_J that depends linearly on the extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$.

(ii) If (i) holds and Σ is positive definite, then the sequence $n\ \left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right)^T \Sigma_J^{-1}\ \left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right)$ converges weakly to χ^2_m distribution.

Furthermore, assuming that Σ_J is positive definite, given that $\hat{\Sigma}_J$ is a consistent estimator for Σ_J, the asymptotic p-value for the hypothesis testing problem H_0 is given by $p = P(T \geq t^2_\delta)$ where

$$t^2_\delta = \ n\ \left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right)^T \hat{\Sigma}_J^{-1}\ \left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right) \tag{2.57}$$

and T has a χ^2_m distribution. (see Patrangenaru and Qiu (2014) [25])
If the distributions are unknown and the samples are small an alternative nonparametric bootstrap technique

(see [8]) may be used. If $\max(n_1, n_2) \leq \frac{m}{2}$, the pulled sample covariance $\hat{\Sigma}_J$ in 2.57 does not have an inverse, and pivotal nonparametric bootstrap methodology can not be applied. In this case one can use non pivotal bootstrap methodology for the two sample problem H_0 which involves a bootstrap confidence region.

THEOREM 2.5.2. *(see Patrangenaru and Qiu (2014) [25])*
Under hypothesis of Theorem $3.1(i)$, assume in addition, that for $a = 1, 2$ the support of the distribution of $X_{a,1}$ and the extrinsic mean $\mu_{a,E}$ are included in the domain of the chart φ and $\varphi(X_{a,1})$ has absolutely continuous component and finite moments of sufficiently high order. Then the joint distribution of

$$V = n^{1/2} \left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right) \tag{2.58}$$

can be approximated by the bootstrap joint distribution of

$$V^* = n^{1/2} \left(\varphi \circ L_\delta^{-1}(H(\bar{X}^*_{1,E}, \bar{X}^*_{2,E}))\right) \tag{2.59}$$

with an error $O_p(n^{-1/2})$, where, for $a = 1, 2$ $\bar{X}^*_{a,E}$ are the extrinsic means of the bootstrap re samples $X^*{}_{a,r_a},\ r_a = 1, \ldots, n_a$. given $X_{a,r_a},\ r_a = 1, \ldots, n_a$.

COROLLARY 2.5.1. *The large sample p-value for the hypothesis testing problem H_0* (2.56) *is given by $p = Pr(T > nV^T\hat{\Sigma}_J V)$ where T has a $\chi^2_m)$ distribution and V is given by equation* (2.58) *and $\hat{\Sigma}_J$ is consistent estimator of the extrinsic covariance matric of $H(\bar{X}_{1,E}, \bar{X}_{2,E})$.*

When the sample size is small, we use Efron's bootstrap , and the hypothesis problem in (2.56) can be solved by using the following $100(1-\alpha)\%$ bootstrap confidence region for $\varphi \circ L_\delta^{-1}(H(\mu_{1,j}, \mu_{2,j}))$.

The concepts presented in sections 2.2 through 2.4 are essential to our statistical analysis in object spaces. We will be able to take advantage of the asymptotic theory developed in section 2.3 (i.e CLT for extrinsic sample means and confidence regions) to conduct hypothesis testing problems on manifolds. Recall from section 2.4 that this space has a Lie group structure with the multiplication operation inherited from the quaternion multiplication on $\mathbb{S}^3 \subset \mathbb{R}^4$. Therefore a 3D object analysis based on landmarks can make use of the recently developed nonparametric techniques for two sample tests on Lie groups (see [25, 21]). We emphasize that the reconstructed configuration of 3D landmarks obtained from pairs of non calibrated camera images, is unique up to a projective transformation in 3D, as noticed in [23]; this allows to analyze without ambiguity the projective shapes of such configurations (see [23]). The developed statistical analysis is performed for samples of pictures of faces, without making any distributional assumption for the corresponding 3D projective shapes of human facial surfaces.

CHAPTER 3

TWO SAMPLE TEST FOR UNMATCHED PAIRS OF 3D PROJECTIVE SHAPES

In this chapter I use the two sample hypothesis testing method for extrinsic means, to differentiate between two 3D scenes of the same kind (faces, flowers, etc...), within the framework of 3D projective shape analysis as developed in [7, 21, 25], based on small samples of digital camera images. The analysis is conducted on the space of 3D projective shapes of k-ads in general position $P\Sigma_3^k$ that contain a projective frame at given landmarks labels, which is homeomorphic to $\mathcal{M} = (\mathbb{R}P^3)^{k-5}$ (see Mardia and Patrangenaru [20]).
In section 3.1 I apply the theory presented in section 2.5 to conduct a two sample test for unmatched pairs on $(\mathbb{R}P^3)^{k-5}$, viewed as a Lie group. In section 3.2 I perform the statistical analysis for sets of pictures of faces along with conveniently selected anatomical landmarks. I make no distributional assumptions for our hypothesis testing methods . The data consist of three sets of images, one female face and two male faces. In Section 3.3 I discuss the process involved in collecting the data sets via MATLAB and introduce a new data collection tool named Agisoft which offers significant benefits and improve the speed and accuracy involved in data collection.

3.1 Two sample test for VW means for unmatched pairs on $(\mathbb{R}P^3)^q$

For a statistical analysis of 3D projective shapes, we are lead into considering r.o.'s Y on $(\mathbb{R}P^3)^q$ that have a VW-mean (have an extrinsic mean w.r.t. the VW-embedding j_k). And since $\mathcal{M} = (\mathbb{R}P^3)^q, q = k-5$ has a Lie group structure (see Chapter 2), and that a Lie group is a homogeneous manifold with a simply transitive Lie group action, we can take advantage of the methodology introduced in the previous chapter. The large sample distribution of the tangential component of the mean change between the extrinsic sample means of two random objects on an embedded Lie group $\mathcal{M}$ can be found in [25]. The probability measure P_Y on $(\mathbb{R}P^3)^q$, associated with such a r.o. is said to be *VW-nonfocal probability measure* on $(\mathbb{R}P^3)^q$. The VW-mean of a VW-nonfocal probability measure $P_Y, Y = ([X^1], ..., [X^q]), \quad (X^s)^T X^s = 1, \ \forall\, s = 1, \dots, q$, is given by

$$\mu_{j_k,E} = (\gamma_1(4), \dots, \gamma_q(4)), \tag{3.1}$$

where $(\lambda_s(a), \gamma_s(a))$, $a = 1,2,3,4$ are the eigenvalues in increasing order, and the corresponding unit eigenvectors of the matrix $E[X^s(X^s)^T]$, respectively (see [23], [20]). In particular, given a random sample of 3D projective shapes $y_1, \dots, y_n$, with $y_i = [x_i], x_i^T x_i = 1, \forall i = 1, \dots, n$, their sample VW-mean is

$$\bar{y}_{j_q} = (g_1(4), \dots, g_q(4)), \tag{3.2}$$

where $(d_s(a), g_s(a))$, $a = 1,2,3,4$ are the eigenvalues in increasing order, and the corresponding unit eigenvectors of the matrix

$$\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T.$$

The particular smooth Lie group action we will use in our analysis is $\alpha \equiv \otimes$, the Lie group multiplication on $(\mathbb{R}P^3)^q$, and if for simplicity we label the VW-means of the two populations by $\mu_{1,E}, \mu_{2,E}$, the null hypothesis in (2.56) can be expressed,

$$H_0 : \ \mu_{1,E} = \mu_{2,E} \quad \text{vs.} \quad H_1 : \ \mu_{1,E} \neq \mu_{2,E} \tag{3.3}$$

where for $a = 1, 2, \mu_{a,E}$ are extrinsic means from VW distributions Q_a on $(\mathbb{R}P^3)^q$. We can rewrite the hypothesis in (4.1) as follows

$$H_0 : \ \mu_{2,E}^{-1} \otimes \mu_{1,E} = 1_{(\mathbb{R}P^3)^q} \quad \text{vs.} \quad H_1 : \ \mu_{2,E}^{-1} \otimes \mu_{1,E} \neq 1_{(\mathbb{R}P^3)^q} \tag{3.4}$$

We further define the smooth map $H : \mathcal{M}^2 \to M$ by $H(x_1, x_2) = (\alpha^{x_2})^{-1}(x_1)$. We now have (4.2) expressed as follow that the expression found in the hypothesis above

$$H_0 : \ H(\mu_{1,E}, \mu_{2,E}) = 1_{(\mathbb{R}P^3)^q} \quad \text{vs.} \quad H_1 : \ H(\mu_{1,E}, \mu_{2,E}) \neq 1_{(\mathbb{R}P^3)^q} \tag{3.5}$$

For $a = 1, 2$, let $Y_{a,1}, \cdots, Y_{a,n_a}$ be independent random samples from VW distributions Q_a on $(\mathbb{R}P^3)^q$ with the extrinsic means $\mu_{1,E}, \mu_{2,E}$ and the corresponding extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$. We are led into characterizing the asymptotic behavior of $\bar{Y}_{2,E}^{-1} \otimes \bar{Y}_{1,E}$, where $\bar{Y}_{1,E}, \bar{Y}_{2,E}$ are the sample extrinsic mean estimators corresponding to the two random samples.

DEFINITION 3.1.1. *The affine chart φ_q defined on an open neighborhood U of $1_{(\mathbb{R}P^3)^q}$ with $\varphi_q(U) \subset (\mathbb{R}^3)^q$ and it is given by*

$$\varphi_q([x_1], \dots, [x_q]) = (\varphi([x_1]), \dots, \varphi([x_q])). \tag{3.6}$$

where φ is an affine chart defined on an affine open neighborhood of $1_{\mathbb{R}P^3}$, given by $\varphi([(x^1, x^2, x^3, x^4)^T]) = (\frac{x^1}{x^4}, \frac{x^2}{x^4}, \frac{x^3}{x^4})$.

Note that $\varphi_q\left(1_{(\mathbb{R}P^3)^q}\right) = (0_3, \ldots, 0_3)$ in $\mathbb{R}^{3q}$ From Patrangenaru et al.(2016)[26] we have the following

PROPOSITION 3.1.1. *For $a = 1, 2$, let $Y_{a,1}, \cdots, Y_{a,n_a}$ identically independent distributed random objects (i.i.d.r.o.'s) from the independent j_k-nonfocal probability measures Q_a. Let $n = n_1 + n_2$ and assume $\lim_{n\to\infty} \frac{n_1}{n} \to \pi \in (0,1)$. Then under H_0 in (3.4),*

(i) The sequence of random vectors $n^{1/2}\left(\varphi_q(\bar{Y}_{2,E}^{-1} \otimes \bar{Y}_{1,E})\right)$ converges weakly to $N_{3q}(0_{3q}, \Sigma_{J_k})$, for some covariance matrix Σ_{J_k} that depends linearly on the extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$.

(ii) If (i) holds and Σ_{J_k} is positive definite, then the sequence $n\left(\varphi_q(\bar{Y}_{2,E}^{-1} \otimes \bar{Y}_{1,E})\right)^T \Sigma_{J_k}^{-1}\left(\varphi_q(\bar{Y}_{2,E}^{-1} \otimes \bar{Y}_{1,E})\right)$ converges weakly to χ^2_{3q} distribution.

(iii) If (i) holds and assume in addition, that for $a = 1, 2$ the support of the distribution of $Y_{a,1}$ and the extrinsic mean $\mu_{a,E}$ are included in the domain of the chart φ_q and $\varphi_q(Y_{a,1})$ has absolutely continuous component and finite moments of sufficiently high order. Then the joint distribution of

$$D = \varphi_q(\bar{Y}_{2,E}^{-1} \otimes \bar{Y}_{1,E})$$

can be approximated by the bootstrap joint distribution of

$$D^* = \varphi_q(\bar{Y^*}_{2,E}^{-1} \otimes \bar{Y^*}_{1,E}) \tag{3.7}$$

with an error $O_p(n^{-1/2})$, where, for $a = 1, 2$ $\bar{Y}^*_{a,E}$ are the extrinsic means of the bootstrap resamples $Y^*{}_{a,r_a},\ r_a = 1, \ldots, n_a$. given $Y_{a,r_a},\ r_a = 1, \ldots, n_a$.

COROLLARY 3.1.1. *For $a = 1, 2$, let $Y_{a,1}, \cdots, Y_{a,n_a}$ identically independent distributed random objects (i.i.d.r.o.'s) from the independent VW probability measures Q_a. Form random resamples with repetition $(Y^*_{a,1}, \cdots, Y^*_{a,n_a})$ from $(Y_{a,1}, \cdots, Y_{a,n_a})$, for $a = 1, 2$. The corresponding approximate $100(1-\alpha)\%$ bootstrap confidence region for $\varphi_q^{-1}(\theta) = \varphi_q(\mu_{2,E}^{-1} \otimes \mu_{1,E})$ is $C^*_\alpha = \varphi_q^{-1}(U^*_\alpha)$, where $U^*_\alpha \in (\mathbb{R}^3)^q$ is the Cartesian product of 3q intervals at $100(1 - \frac{\alpha}{3q})\%$ confidence level for the components of $\theta = \varphi_q(\mu_{2,E}^{-1} \otimes \mu_{1,E})$. This simultaneous confidence intervals yield a confidence region of at least $100(1-\alpha)\%$ level, of coverage error $O_P(n^{-1/2})$. We reject our null hypothesis if $0_{3q} \notin U^*_\alpha$, that is, if at least one of these intervals does not contain 0.*

3.2 Data set and hypothesis testing results

In this section we analyze the 3D projective mean shape changes to differentiate between faces (see Patrangenaru et.al.(2016)[24]). We conduct two sample hypothesis testing on unmatched pairs (i.e different sample sizes $n_1 \neq n_2$.) The analyzed data set consists of images of the faces shown below

Figure 3.1: Faces used for analysis

For our landmark based analysis we first recover a 3D configuration of $k = 10$ landmarks from each pairs of uncalibrated pictures of the same face (see Ma et. l.(2005)[19]). This will result, for the female face,in 8 projective shapes (3-D configurations of labeled points), for the first male we have 10 projective shapes and finally for the last data set we have 11 projective shapes. The collections and reconstructions of all of our landmark configurations were done in Matlab. The landmarks are shown in figure 3.2:

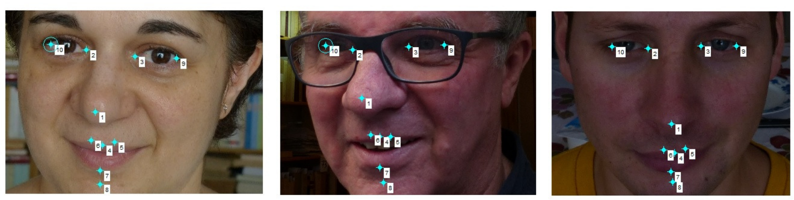


Figure 3.2: Landmark placements for all faces

For a given face, and a single set of landmarks $\{u_1, \dots, u_{10}\}$ the first five labeled points $u_1, \dots, u_5$ are used to construct a projective frame $\pi = (\tilde{u}_1, \dots, \tilde{u}_5)$ where $\tilde{u}_i = [u_1^i : u_2^i : u_2^i : 1]$. Throughout the data we use the same landmarks for our projective frame and they are, in increasing order; *pronasale, right and left Endocathion, Labiale Superius, left Crista Philtri*. The resulting $k-5$-tuple of projective coordinates $(p_6^\pi, \dots, p_{10}^\pi) \in (\mathbb{R}P^3)^5$ represents one observation used in our analysis. The resulting $k-5$-tuple of projective coordinates $(p_6^\pi, \dots, p_{10}^\pi) \in (\mathbb{R}P^3)^5$ represents one observation used in our analysis. In other word, the projective shape of the 3D 10-ad, is determined by the 5 projective coordinates of the remaining landmarks of the reconstructed configurations.

3.2.1 2 sample test for facial data

Given two faces, we assume that the sets $Y_{1,1},\dots,Y_{1,n_1}$ and $Y_{2,1},\dots,Y_{2,n_2}$ of 3D projective shapes recovered from data sets consisting of n_1 and n_2 pairs of images respectively are coming from a VW Q_1 and Q_2 distribution on $(\mathbb{R}P^3)^5$. We statistically differentiate between faces if we reject the following null hypothesis ;

$$H_0 : \mu_{1,E}^{-1} \otimes \mu_{2,E} = 1_{(\mathbb{R}P^3)^5}$$

For our result we used the simultaneous confidence intervals mentioned in Corollary (3.1.1). We failed to reject the null hypothesis if all of our confidence intervals contain the value 0.

Results for comparing Male faces:

For the two male faces with data sets of sizes $n_1 = 10$ and $n_2 = 11$ we conduct our two sample hypothesis testing and we get the following simultaneous intervals

Simultaneous confidence intervals for changes between the 2 mean projective shapes of the two faces landmarks 6 to 8			
	LM6	LM7	LM8
x	$(-1.111498, 0.805386)$	$(-1.117512, 1.099536)$	$(-1.296547, 0.966296)$
y	$(-1.215218, 0.710931)$	$(-1.355167, 1.336021)$	$(-0.635282, 1.372627)$
z	$(-1.161234, 1.150762)$	$(-1.432217, 1.349541)$	$(-1.394141, 1.349442)$

Simultaneous confidence intervals for changes between the 2 mean projective shapes of the two faces landmarks 9 and 10		
	LM9	LM10
x	$(0.952164, 0.996354)$	$(-0.962541, 1.005917)$
y	$(-0.760124, 1.129782)$	$(-1.070631, 0.982195)$
z	$(-0.817503, 1.319117)$	$(-1.319374, 1.089272)$

Another good set of visual tools we use in our analysis are the *Bootstrap marginals* boxes which can be found in figure 3.3.

We notice that one of the simultaneous confidence intervals for landmark 9, corresponding to the right *Exocanthion*, does not contain 0. We then reject the null hypothesis, showing that there is significant projective shape change between the two male faces. And for the bootstrap marginal boxes we notice that the first three landmarks have a pretty dense concentration around the center, indicating no significant mean change which is not the case for the last two boxes.

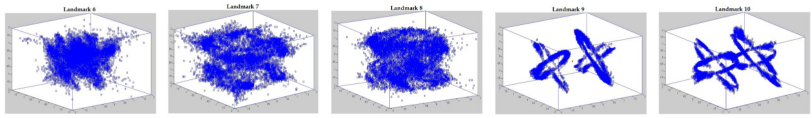


Figure 3.3: Bootstrap projective shape marginals for male face data

Result for cross gender comparison:

For samples of sizes $n_1 = 11$ (male) and $n_2 = 8$ (female) conduct the following null hypothesis H_0 : $\mu_{1,11}^{-1} \otimes \mu_{2,8} = 1_{(\mathbb{R}P^3)^5}$, and in the figure below 3.4 we indicate the two faces being analyzed.

Figure 3.4: Faces used in cross gender analysis

We then get the following bootstrap marginals boxes (figure 3.5) for our cross gender analysis along with the simultaneous confidence intervals.

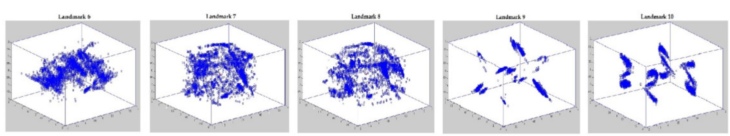


Figure 3.5: Bootstrap projective shape marginals for cross gender data

Simultaneous confidence intervals for cross gender landmarks 6 to 8			
	LM6	LM7	LM8
x	$(-1.251984, 1.202986)$	$(-1.228628, 1.234229)$	$(-1.273092, 1.332798)$
y	$(-0.633834, 0.902621)$	$(-0.928523, 0.995304)$	$(-0.226587, 0.865510)$
z	$(-0.231190, 0.432009)$	$(-0.684483, 1.045302)$	$(-0.590623, 1.132418)$

Simultaneous confidence intervals for cross gender landmarks 9 and 10		
	LM9	LM10
x	$(0.998446, 1.028374)$	$(-0.988191, -0.931250)$
y	$(-0.702335, 0.540613)$	$(-1.162803, 1.008259)$
z	$(-1.057821, 0.849069)$	$(-0.118635, 0.969739)$

The landmarks 9 and 10 corresponding to the right and left Exocanthion have intervals not containing 0. We reject the null hypothesis, and conclude that there is a significant projective shape change between the two faces.

Results for cross validation:

We separate the original sample into two smaller data sets of sizes $n_1 = 5$ and $n_2 = 6$. They are displayed in Figs (3.6).

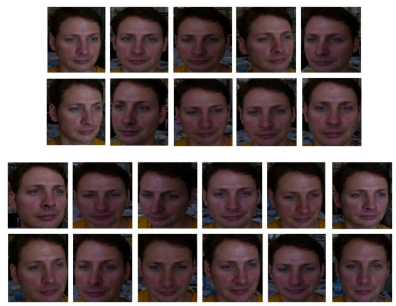


Figure 3.6: Cross validation samples

The bootstrap axial marginals (Fig 3.7) and simultaneous confidence regions for cross validation are given below.

Simultaneous confidence interval for cross validation face 2 for landmarks 6 to 8			
	LM6	LM7	LM8
x	$(-17.496785, 3.552070)$	$(-4.027879, 4.860970)$	$(-1.990796, 7.497709)$
y	$(-10.967285, 4.340129)$	$(-3.776026, 9.830274)$	$(-7.558584, 0.865119)$
z	$(-2.724184, 13.093615)$	$(-3.006049, 5.891478)$	$(-0.698745, 4.293201)$

Simultaneous confidence intervals for cross validation face2 for landmarks 9 and 10		
	LM9	LM10
x	$(-2.459882, 1.230096)$	$(-3.264292, 1.036499)$
y	$(-1.631839, 0.983147)$	$(-1.387133, 2.942318)$
z	$(-1.451487, 1.196335)$	$(-0.916768, 1.658124)$

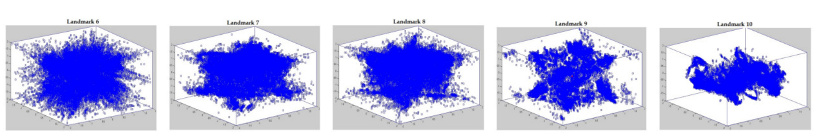


Figure 3.7: Bootstrap marginals for crossvalidation of male face 2

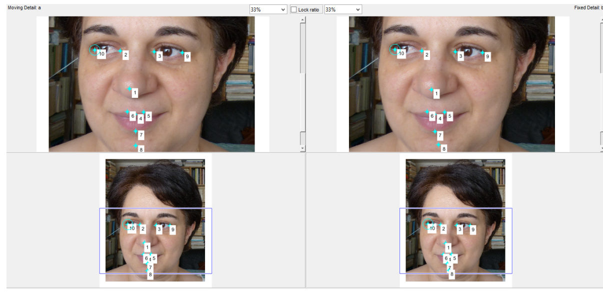


Figure 3.8: Landmark placements in Matlab

All the simultaneous intervals (affine coordinates) contain 0. We fail to reject the null hypothesis; there no statistically significant mean projective shape change. Furthermore, the bootstrap marginals all show values that are concentrated around 0_3.

3.3 Landmark coordinates from ideal non calibrated camera images

Our data sets are built from sets of digital camera images of faces and other objects. The 3D face analysis we are conducting is a landmark based analysis. Our landmarks are composed of reconstructed 3D points in a particular configuration and the collection of our landmarks in Matlab is done in a few stages.

3.3.1 Matlab data set

For any one reconstruction of a particular 3D object (faces, flowers, leaves, etc...) two pictures from two different angles are needed. Once the pair of pictures are stored and saved in the an appropriate window within the Matlab platform, the digital images are loaded using the **imread** command in Matlab. The landmarks are manually selected using the function **cpselect()**. We illustrate a set of landmarks in Fig 3.8.

Generally, a finite configuration of eight or more points in general position in 3D can be reconstructed, by using the fundamental matrix of the coordinates of the images of these points provided by two ideal non calibrated digital camera views. We assign the same landmarks throughout our whole data sample; the images from below show the placement of our matching points.

By this method we usually get very reliable 3D coordinates for our landmarks. However, one drawback associated with this technique is that it is hard to visualize the reconstructed 3D configurations. In fact, to get a descent visualization of our reconstruction may require the collection of a large amount of landmarks, which can be time consuming.

To illustrate this particular situation we have the following 3D reconstruction involving 80 landmarks placed on a pair of pictures of an oak leaf and resulting in the following 3D images without color and/or texture.(see Fig 3.9)

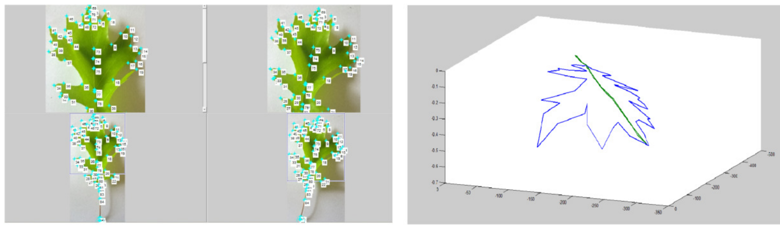


Figure 3.9: Oak leaf reconstruction with midriff

3.3.2 Advanced 3D data collection methods from digital camera outputs

Recently for our data analysis we started using a professional version of Agisoft, which extracts the 3D image of a surface from two or more non-calibrated digital camera views, based on RGB texture matching followed by a 3D reconstruction algorithm. This software gives us a much better visualization of our reconstructed data set without relying on landmark collection and the use of an eight point algorithm to estimate the fundamental matrix (see Ma et al.(2005)[19]).

Although the reconstruction could be done with just two uncalibrated camera images, we get a better resolution and complete reconstruction of the surface of a head or face, by increasing the number of images of the same individual. A training data set of fifteen surfaces of faces including texture was collected from digital images (see ani.stat.fsu.edu/∼vic/Davids-PhDs). An additional sample of three samples of 3D faces was collected along with facial landmark coordinates; this will be used in Chapter 6 (see ani.stat.fsu.edu/∼vic/Davids-PhDs/MANOVA) We illustrate this fact we use set of pictures in Fig. 3.10. After the reconstruction is done, we may visualize our result and also indicate the relative camera placement in Fig. 3.11. The Agisoft output then gives us the 3D coordinates of our ten landmarks in Figs. 3.12-3.13.

In this chapter we took advantage of the fact the $\mathcal{M} = (\mathbb{R}P^3)^q$ being a Lie group acts simply transitively on itself with the action being the left multiplication $\otimes$. We can then use the recent work on asymptotic behavior on homogeneous space to have an expression of the convergence of $\left(\varphi \circ L_\delta^{-1}(H(\bar{X}_{1,E}, \bar{X}_{2,E}))\right)$. This allows us to perform hypothesis testing on random samples of different sizes defined on $\mathcal{M}$. The theory

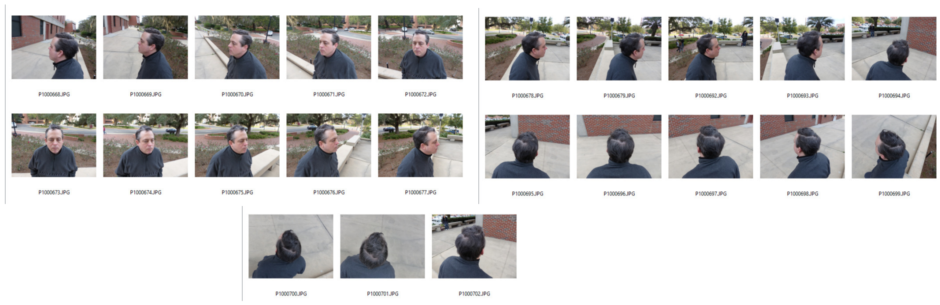


Figure 3.10: Pictures used for 3D reconstruction

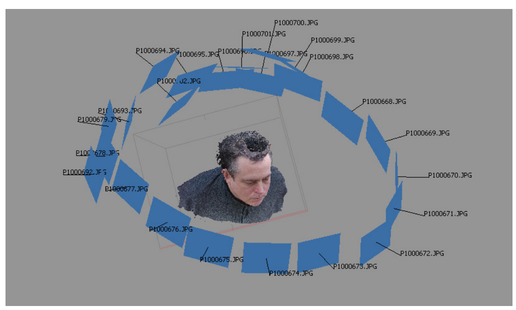
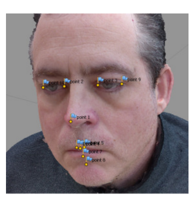


Figure 3.11: 3D face reconstruction with camera placement



s	X est	Y est	Z est
point 1	-0.056717	-0.181218	0.342859
point 2	-1.545557	-1.970076	3.936616
point 3	1.577311	-2.008529	4.018979
point 4	0.103276	2.888047	1.164003
point 5	0.478865	2.788800	1.214435
point 6	-0.224090	2.844635	1.227195
point 7	0.228277	3.734438	1.206160
point 8	0.288231	4.821007	1.547070
point 9	3.685678	-1.958006	4.334273
point 10	-3.676733	-1.767309	4.242580

Figure 3.12: Landmark placement and coordinates

Figure 3.13: Pictures for 3D reconstructions

involves applying a *Cramer's delta method* for functions between manifolds that will depend heavily on the choice of a convenient chart φ. The expression of the covariance matrix Σ_J we obtain depends linearly on the extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$.. Recall that an extrinsic matrix Σ_E is always defined with respect to a basis $f_1(\mu_E), ..., f_m(\mu_E)$ of local frame field referred to as orthoframe (see definition 2.3.2). In the next chapter we will work on developing an asymptotic theory that builds on the work in [25] but is not dependent on the choice of a chart. The work in this chapter led to a couple of publications *" 3D face analysis from digital camera images"* (see [26]) and *"Projective shape analysis of contours and finite 3D configurations from digital camera images "*(see [24]).

CHAPTER 4

A TWO SAMPLE TEST FOR MEAN CHANGE BASED ON A DELTA METHOD ON MANIFOLDS

I introduce a new method of two sample tests for 3D mean projective shapes. This method builds upon the various results of the two sample hypothesis testing methods, as developed in Patrangenaru et al. (2010)[23], Crane and Patrangenaru et al.(2011) [7], and Patrangenaru et al.(2014) [25].

In section 4.1 I start by expressing a version of the Cramer's delta method for a function $F : \mathcal{M}_1 \to \mathcal{M}_2$ that depends on a compositions of functions involving the embeddings of both the domain and co domain space. In section 4.2 I will use the results of our new version of the Cramer's delta method to construct an asymptotic behavior for $\mu_{2,E}^{-1} \odot \mu_{1,E}$ with explicit definition of the corresponding extrinsic covariance matrix.The result in this section can also be applied to any smoth function between manifolds. In the last section I express the some asymptotic behaviors for the space $\mathbb{R}P^3$.

4.1 Cramer's delta method for data on manifolds

Recall that $(\mathcal{G}, \odot)$, a Lie group is a manifold with a group structure and for which the multiplication map $(g, h) \to g \odot h$ and the inverse map $g \to g^{-1}$ are smooth as maps between manifolds.

We consider the following null hypothesis

$$\begin{aligned} H_0 :\ & \mu_{1,E} = \mu_{2,E} \odot \delta \\ H_1 :\ & \mu_{1,E} \neq \mu_{2,E} \odot \delta \end{aligned} \tag{4.1}$$

Since for $a = 1, 2, X_{a,1}, \ldots, X_{a,n_a}$ i.i.d. random objects on $\mathcal{G}$ we can rewrite the hypothesis in (4.1) as follows

$$H_0 :\ \mu_{2,E}^{-1} \odot \mu_{1,E} = \delta \ vs.\ H_1 :\ \mu_{2,E}^{-1} \odot \mu_{1,E} \neq \delta \tag{4.2}$$

For that we will need to know the asymptotic behavior of $\bar{X}_{2,E}^{-1} \odot \bar{X}_{1,E}$, where $\bar{X}_{1,E}, \bar{X}_{2,E}$ are the sample extrinsic mean estimators corresponding to the two random samples. To address this problem, we are first considering an extension of Cramer's delta method, in the context of manifold valued data. An initial extension can be found in Patrangenaru et al.(2014)[25]. Here we are interested in a method which applies to embeddings $j_a : \mathcal{M}_a \to \mathbb{R}^{N_a}, a = 1, 2$. Let $X_1, \ldots, X_n$ be i.i.d. random objects on $(\mathcal{M}_a, \rho_{j_a})$ and assume μ_E, Σ_E are respectively the extrinsic mean, and extrinsic covariance matrix of X_1 (see Bhattacharya and Patrangenaru (2005)). Let $\mathcal{F} \subset \mathbb{R}^{N_1}$ be the set of j_1-focal points then P_{j_1} is the corresponding projection with $P_{j_1} : \mathcal{F}^c \to j_1(\mathcal{M}_1) \subset \mathbb{R}^{N_1}$.

THEOREM 4.1.1. *(Delta method for embedded manifolds). Assume $F : \mathcal{M}_1 \to \mathcal{M}_2$ is a differentiable function between manifolds, and let $(f_1^{(a)}, \ldots, f_{m_a}^{(a)})$ be orthonormal bases in $T_{\mu_{a,E}}(\mathcal{M}_a)$, where $\mu_{1,E} = \mu_E, \mu_{2,E} = F(\mu_E)$. For $a = 1, 2$, assume $dim M_a = m_a$ with j_1 and j_2 as previously defined. Let $X_1, \ldots, X_n$ be a sequence of random objects on $\mathcal{M}_1$ such that*

$$n^{1/2} \ \tan_{j_1(\mu_E)} (j_1(X_n) - j_1(\mu_E)) \to_d \mathcal{N}_{m_1}(0, \Sigma_E).$$

$$Then$$

$$n^{1/2} \ \tan_{j_2(F(\mu_E))} (j_2 (F(X_n)) - j_2 (F(\mu_E))) \to_d \mathcal{N}_{m_2}(0, \Sigma^F_{j_2,E})$$

where $\Sigma^F_{j_2,E} = dF \ \Sigma_E \ (dF)^T$ with dF given by

$$dF = [(dF)_{ab}] = \left[d_\mu \tilde{F}_{12}(e_b) \ \cdot \tilde{e}_a(\tilde{F}_{12}(\mu)) \right], \ for \ a = 1, ..., m_2; \ and \ b = 1, ..., m_1.$$

where $j_2 \circ F \circ j_1^{-1} \circ P_{j_1} = \tilde{F}_{12}$.

Proof. Now recall from Bhattacharya and Patrangenaru (2005)[6] that

$$\Sigma_E = A^T \Sigma_\mu A = \begin{bmatrix} e_1(P_{j_1}(\mu))^T \\ \vdots \\ e_{m_1}(P_{j_1}(\mu))^T \end{bmatrix} \Sigma_\mu \ \left[e_1(P_{j_1}(\mu)) \quad \cdots \quad e_{m_1}(P_{j_1}(\mu)) \right] \tag{4.3}$$

where $\Sigma_\mu = (D_\mu P_{j_1}) \ \Sigma \ (D_\mu P_{j_1})^T$ and Σ is the covariance matrix of $j_1(X_1)$ with respect to the standard basis $e_1, ..., e_{N_1}$ of $\mathbb{R}^{N_1}$. By the CLT, we have

$$n^{1/2} \ (j_1(X_n) - j_1(\mu_E)) \to_d \mathcal{N}_{N_1}(0, \Sigma_\mu).$$

Let us define the following map $\tilde{F} = j_2 \circ F \circ j_1^{-1}$; this is a map from $j_1(M_1) \to j_2(M_2)$ and acts as follows

$$\tilde{F}(j_1(x)) = \tilde{F}(P_{j_1}(j_1(x))) = j_2 (F(x)), \forall x \in \mathcal{M}_1.$$

Note that $\tilde{F} \circ P_{j_1}$ is a smooth function from $\mathcal{F}^c \subset \mathbb{R}^{N_1}$ to $j_2(\mathcal{M}_2) \subset \mathbb{R}^{N_2}$. We can now apply the Cramer's delta method and get

$$n^{1/2}\ (j_2\left(F(X_n)\right) - j_2\left(F(\mu_E)\right)) \to_d \mathcal{N}_{N_2}(0, \Sigma_{j_2})$$

where $\Sigma_{j_2} = (D_\mu(\tilde{F} \circ P_{j_1}))\ \Sigma\ (D_\mu(\tilde{F} \circ P_{j_1}))^T = (D_{P_{j_1}(\mu)}\tilde{F})\ \Sigma_\mu\ (D_{P_{j_1}(\mu)}\tilde{F})^T$.

Now assume that V_2 is an open neighborhood of $F(\mu_E)$ in $\mathcal{M}_2$, and $V_1 = F^{-1}(V_2)$. Assume $U_2 \subset \mathbb{R}^{N_2}$, is an open subset, such that $U_2 \cap j_2(\mathcal{M}_2) = j(V_2)$, and $p_2 \to (\tilde{e}_1(p_2), \ldots, \tilde{e}_{N_2}(p_2))$ is an orthonormal frame field on U_2, which is adapted to the embedding j_2. Define the local frame field $y \to (f_{2,1}(y)), \ldots, f_{2,m_2}(y))$ on V_2, such that

$$\forall\, y \in V_2,\ \ \tilde{e}_s(j_2(y)) = d_y j_2(f_{2,s}(y)),\ \ s = 1, \ldots, m_2.$$

Now let $\Big(\ \tilde{e}_1(\tilde{F}(p_1)), \ldots, \tilde{e}_{N_2}(\tilde{F}(p_1))\ \Big)$ be the value of this adapted frame field at a point $\tilde{F}(p_1)$ on $j_2(V_2)$ around $j_2 \circ F(\mu_E)$ and for $p_1 \in j_1(\mathcal{M}_1) \subset \mathbb{R}^{N_1}$. Note that $d_\mu(\tilde{F} \circ P_{j_1})(e_b) \in T_{\tilde{F}(P_{j_1}(\mu))} j_2(M_2)$, while $(e_1, \ldots, e_{N_1})$ is the standard basis in $\mathbb{R}^{N_1}$.

To ease notation we let $\tilde{F} \circ P_{j_1} = \tilde{F}_{12}$ and $\tilde{F}_{12} : \mathcal{F}^c \to j_2(\mathcal{M}_2)$, where $\mathcal{F}^c$ represents j_1-nonfocal set, and we now have:

$$d_\mu \tilde{F}_{12}(e_b) = \sum_{a=1}^{m_2} \left[\Big(d_\mu \tilde{F}_{12}(e_b)\Big) \cdot \tilde{e}_a(\tilde{F}_{12}(\mu))\right] \tilde{e}_a(\tilde{F}_{12}(\mu)) \tag{4.4}$$

And, for $e_b \in \mathbb{R}^{N_1}$ with $b = 1, \ldots, N_1$, we have

$$\begin{aligned}
\Sigma_{j_2} &= (D_{P_j(\mu)}\tilde{F})\ \Sigma_\mu\ (D_{P_j(\mu)}\tilde{F})^T \\
\Sigma_{j_2} &= \left[\left[\sum_{a=1}^{m_2} d_\mu \tilde{F}_{12}(e_b) \cdot \tilde{e}_a(j_2(F(\mu_E))\tilde{e}_a(j_2(F(\mu_E))\right]_{b=1,..,N_1}\right] \Sigma_\mu \\
&\quad \left[\left[\sum_{a=1}^{m_2} d_\mu \tilde{F}_{12}(e_b) \cdot \tilde{e}_a(j_2(F(\mu_E))\tilde{e}_a(j_2(F(\mu_E))\right]_{b=1,..,N_1}\right]^T
\end{aligned}$$

Note that $\Sigma_{j_2} \in M(N_2, N_2, \mathbb{R})$, while $\Sigma_\mu \in M(N_1, N_1, \mathbb{R})$.

If we set $\nu = j_2(F(\mu_E))$, then the tangential component $\tan(\nu)$ of $\nu \in \mathbb{R}^{N_2} = T_{\tilde{F}_{12}(\mu)} j_2(\mathcal{M}_2) \oplus (T_{\tilde{F}_{12}(\mu)} j_2(M_2))^\perp$, w.r.t the basis $e_a(\tilde{F}_{12}(\mu)) \in T_{\tilde{F}_{12}(\mu)} j_2(\mathcal{M}_2)$ has the following asymptotic behavior

$$\begin{aligned}
&\tan_{j_2(F(\mu_E))}\Big(\tilde{F}_{12}(j_1(X_{n_1}) - \tilde{F}_{12}(\mu)\Big) \to_d \mathcal{N}_{m_2}(0, \Sigma^F_{j_2,E}) \\
&\tan_{j_2(F(\mu_E))}\left(j_2\left(F(X_{n_1})\right) - j_2\left(F(\mu_E)\right)\right) \to_d \mathcal{N}_{m_2}(0, \Sigma^F_{j_2,E})
\end{aligned} \tag{4.5}$$

with

$$\Sigma^F_{j_2,E} = \begin{bmatrix} \tilde{e}_1(\tilde{F}_{12}(\mu))^T \\ \vdots \\ \tilde{e}_{m_2}(\tilde{F}_{12}(\mu))^T \end{bmatrix} (D_{P_j(\mu)}\tilde{F})\ \Sigma_\mu\ (D_{P_j(\mu)}\tilde{F})^T \left[\tilde{e}_1(\tilde{F}_{12}(\mu)) \quad \cdots \quad \tilde{e}_{m_2}(\tilde{F}_{12}(\mu))\right]$$

$$\Sigma^F_{j_2,E} = B\ \Sigma_\mu B^T$$

were $B = \begin{bmatrix} \tilde{e}_1(\tilde{F}_{12}(\mu))^T \\ \vdots \\ \tilde{e}_{m_2}(\tilde{F}_{12}(\mu))^T \end{bmatrix} \left[\left[\sum_{a=1}^{m_2} d_\mu \tilde{F}_{12}(e_b) \cdot \tilde{e}_a(\tilde{F}_{12}(\mu))\tilde{e}_a(\tilde{F}_{12}(\mu))\right]_{b=1,..,N_1}\right]$

$$B = \begin{bmatrix} \Big(d_\mu \tilde{F}_{12}(e_1)\Big) \cdot \tilde{e}_1(\tilde{F}_{12}(\mu)) & \ldots & \Big(d_\mu \tilde{F}_{12}(e_{N1})\Big) \cdot \tilde{e}_1(\tilde{F}_{12}(\mu)) \\ \vdots & & \\ \Big(d_\mu \tilde{F}_{12}(e_1)\Big) \cdot \tilde{e}_{m_2}(\tilde{F}_{12}(\mu)) & \ldots & \Big(d_\mu \tilde{F}_{12}(e_{N_1})\Big) \cdot \tilde{e}_{m_2}(\tilde{F}_{12}(\mu)) \end{bmatrix} \tag{4.6}$$

Note that,$A^T A = I_{N_1}$ and

$\Sigma^F_{j_2,E} = BA\ A^T\ \Sigma_\mu\ A\ A^T B^T = (BA)\ \Sigma_E\ (BA)^T$, and

$$BA = \begin{bmatrix} \Big(d_\mu \tilde{F}_{12}(e_1)\Big) \cdot e_1(\tilde{F}_{12}(\mu)) & \ldots & \Big(d_\mu \tilde{F}_{12}(e_{m_1})\Big) \cdot e_1(\tilde{F}_{12}(\mu)) \\ \vdots & & \\ \Big(d_\mu \tilde{F}_{12}(e_1)\Big) \cdot e_{m_2}(\tilde{F}_{12}(\mu)) & \ldots & \Big(d_\mu \tilde{F}_{12}(e_{m_1})\Big) \cdot e_{m_2}(\tilde{F}_{12}(\mu)) \end{bmatrix} \tag{4.7}$$

and letting $B\ A = dF$ we have our desired result. □

4.2 Asymptotic behavior for Lie group

For $a = 1, 2$, let $X_{a,1}, \cdots, X_{a,n_a}$ be independent random samples defined on $\mathcal{G}$, a Lie group, from a distribution Q_a, with the extrinsic means $\mu_{1,E}, \mu_{2,E}$ and corresponding extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$. Let $j : \mathcal{G} \to \mathbb{R}^N$ be an embedding. We are interested in the asymptotic behavior of

$$\tan_{j(\mu^{-1}_{2,E} \odot \mu_{1,E})} \Big(j(\overline{X}^{-1}_{2,E} \odot \overline{X}_{1,E}) - j(\mu^{-1}_{2,E} \odot \mu_{1,E})\Big)$$

Recall that the map $(g_1, g_2) \to g_1 \odot g_2$, for $g_1, g_2 \in \mathcal{G}$ is a smooth map from $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$. Theorem 4.2.1 below, focuses on a more general case involving manifolds $\mathcal{M}$ and $\mathcal{N}$ along with their corresponding embedding $j_1 : \mathcal{M} \to \mathbb{R}^{N_1}$ and $j_2 : \mathcal{N} \to \mathbb{R}^{N_2}$ and corresponding chord distances ρ_{j_1} and ρ_{j_2}.

THEOREM 4.2.1. *Let $\mathcal{M}$ and $\mathcal{N}$ be respectively, m-dimensional and n-dimensional smooth manifolds with embeddings $j_1 : \mathcal{M} \to \mathbb{R}^{N_1}$ and $j_2 : \mathcal{N} \to \mathbb{R}^{N_2}$ .Let $G : \mathcal{M} \times \mathcal{M} \to \mathcal{N}$ be a smooth function between manifolds. For $a = 1, 2$ let $f_1^{(a)}, \cdots, f_m^{(a)}$ be orthonormal basis in $T_{\mu_{a,E}}(\mathcal{M})$ where $\mu_{a,E}$ are extrinsic means of j_1-nonfocal probability distribution Q_a on $\mathcal{M}$ with corresponding extrinsic covariance matrices $\Sigma_{a,E}$ and $\overline{X}_{a,E}$ are their respective extrinsic sample means.*

(i) Let $n = n_1 + n_2$, if $\frac{n_1}{n} \to \pi$ as $n_a \to \infty$, and for $a = 1, 2$ we have the following asymptotic behavior,

$$n_a^{1/2} \ \tan_{j_1(\mu_{a,E})} \left(j_1(\overline{X}_{a,E}) - j_1(\mu_{a,E}) \right) \xrightarrow{\mathcal{L}} \mathcal{N}_m(0, \Sigma_{a,E})$$

$$Then$$

$$n^{1/2} \ \tan_{j_1^{(2)}(\mu_{1,E},\mu_{2,E})} \left(j_1^{(2)}(\overline{X}_{1,E}, \overline{X}_{2,E}) - j_1^{(2)}(\mu_{1,E}, \mu_{2,E}) \right) \xrightarrow{\mathcal{L}} \mathcal{N}_{2m}(0, \Sigma_{j_1,E}^{(2)}), \tag{4.8}$$

where $\Sigma_{j_1,E}^{(2)} = \begin{pmatrix} \frac{1}{\pi}\Sigma_{1,E} & 0_m \\ 0_m & \frac{1}{1-\pi}\Sigma_{2,E} \end{pmatrix}$ and $j_1^{(2)} : \mathcal{M} \times \mathcal{M} \to j_1(\mathcal{M}) \times j_1(\mathcal{M})$.

(ii) Let $(g_1, \cdots, g_n)$ be an orthonormal basis in $T_{G(\mu_{1,E},\mu_{2,E})}\mathcal{N}$, if the result in (i) holds we have

$$n^{1/2} \ \tan_{j_2(G(\mu_{1,E},\mu_{2,E}))} \left(j_2 \left(G(\overline{X}_{1,E}, \overline{X}_{2,E}) \right) - j_2 \left(G(\mu_{1,E}, \mu_{2,E}) \right) \right) \xrightarrow{\mathcal{L}} \mathcal{N}_n(0, \Sigma_{j_2,E}^G) \tag{4.9}$$

with

$$\Sigma_{j_2,E}^G = \frac{1}{\pi} \, (dG^{(1)}) \, \Sigma_{1,E} (dG^{(1)})^T + \frac{1}{1-\pi} (dG^{(2)}) \Sigma_{2,E} (dG^{(2)})^T \tag{4.10}$$

and $dG_{ab}^{(1)} = d_{\mu_1,\mu_2}\tilde{G}(\hat{e}_b) \cdot \tilde{e}_a(\tilde{G}(\mu_1, \mu_2)$; $dG_{ab}^{(2)} = d_{\mu_1,\mu_2}\tilde{G}(\hat{e}_{N_1+b}) \cdot \tilde{e}_a(\tilde{G}(\mu_1, \mu_2)$ for $a = 1, ..., n$ and $b = 1, ..., m$. And $\tilde{G} = j_2 \circ G \circ j_1^{-1}(P_{j_1}) \times j_1^{-1}(P_{j_1})$.

Proof. For part (i), it follows from Bhattacharya and Patrangenaru (2005) [6] that

$$n_a^{1/2} \ \left(P_{j_1}(\overline{j(X_{a,1})}) - P_{j_1}(\mu_a) \right) \to_d \mathcal{N}_{N_1} \left(0, \Sigma_{\mu_a} \right), \tag{4.11}$$

where, for $a = 1, 2$ $\Sigma_{\mu_a} = (D_{\mu_a} P_{j_1}) \, \Sigma_a \, (D_{\mu_a} P_{j_1})^T$ and Σ_a is the covariance matrix for the random vector $j_1(X_{a,1}) \in j_1(\mathcal{M})$. And the projection $P_{j_1} : \mathcal{F}^c \to j_1(\mathcal{M})$ where $\mathcal{F}$ is the set of j_1-focal points. Since $n_1/n \to \pi$ as $n_1 \to \infty$ it then follows that

$$n^{1/2} \ \left(P_{j_1} \times P_{j_1}(\overline{j(X_{1,1})}, \overline{j(X_{2,1})}) - P_{j_1} \times P_{j_1}(\mu_1, \mu_2) \right) \to_d \mathcal{N}_{2N_1}(0, \Sigma^\star) \tag{4.12}$$

with $\Sigma^\star = \begin{pmatrix} \frac{1}{\pi}\Sigma_{\mu_1} & 0_{N_1} \\ 0_{N_1} & \frac{1}{1-\pi}\Sigma_{\mu_2} \end{pmatrix}$ since the samples are independents.
Recall that from Bhattacharya and Patrangenaru (2005) [6] , that for $a = 1, 2$ $\Sigma_{a,E}$ are the extrinsic covariance matrices of the j-nonfocal distributions Q_a of X_a w.r.t. $(f_1^{(a)}(\mu_{a,E}), \ldots, f_m^{(a)}(\mu_{a,E}))$ the special

orthonormal frame fields around $\mu_{a,E}$. For each of these local frame fields there is a corresponding adapted frame field $(e_1^{(a)}(P_{j_1}(\mu_a)), \dots, e_{N_1}^{(a)}(P_{j_1}(\mu_a))$ around $P_{j_1}(\mu_a) = j_1(\mu_{a,E})$ (for a definition see section (2.2)). Now from the two local frame fields we have above, we can construct the following local frame field in $\mathcal{M} \times \mathcal{M}$ around the point $(\mu_{1,E}, \mu_{2,E})$;

$$\begin{aligned} &[f_1(x_1,x_2), \dots, f_m(x_1,x_2), f_{m+1}(x_1,x_2), \dots, f_{2m}(x_1,x_2)] \\ &\qquad = \\ &\left[\left(f_1^{(1)}(x_1), \zeta(x_2)\right), \dots, \left(f_m^{(1)}(x_1), \zeta(x_2)\right), \left(\zeta(x_1), f_1^{(2)}(x_2)\right), \dots, \left(\zeta(x_1), f_m^{(2)}(x_2)\right)\right], \end{aligned} \tag{4.13}$$

where $\zeta(x)$ is the zero section of T_pU with $U \in \mathcal{M}$ and U contains $\mu_{a,E}$ for $a = 1, 2$.

For ease of notation we let j be the embedding $j \equiv j_1^{(2)} : \mathcal{M} \times \mathcal{M} \to j_1(\mathcal{M}) \times j_1(\mathcal{M})$ then we get, for the local frame field in equation (4.13) on an open subset of $\mathcal{M} \times \mathcal{M}$ containing $(\mu_{1,E}, \mu_{2,E})$, the following vectors in $\mathbb{R}^{N_1} \times \mathbb{R}^{N_1}$

$$\left[d_{\mu_{1,E},\mu_{2,E}} j(f_1(x_1,x_2)), \dots, d_{\mu_{1,E},\mu_{2,E}} j(f_m(x_1,x_2)), d_{\mu_{1,E},\mu_{2,E}} j(f_{m+1}(x_1,x_2)), \dots, d_{\mu_{1,E},\mu_{2,E}} j(f_{2m}(x_1,x_2))\right]$$

which is expressed in more details as follow;

$$\begin{aligned} &\left[\left(d_{\mu_{1,E}} j_1(f_1^{(1)}(x_1)), d_{\mu_{2,E}} j_1(\zeta(x_2))\right), \dots, \left(d_{\mu_{1,E}} j_1(f_m^{(1)}(x_1)), d_{\mu_{2,E}} j_1(\zeta(x_2))\right), \right. \\ &\left. \left(d_{\mu_{1,E}} j_1(\zeta(x_1)), d_{\mu_{2,E}} j_1(f_1^{(2)}(x_2))\right), \dots, \left(d_{\mu_{1,E}} j_1(\zeta(x_1)), d_{\mu_{2,E}} j_1(f_m^{(2)}(x_2))\right)\right], \end{aligned} \tag{4.14}$$

where $d_{\mu_{a,E}} j_1(\zeta(x_a))$ is the zero section of $T_{j_1(p)} j_1(U)$ which corresponds to the zero vector in $\mathbb{R}^{N1}$.

It follows that the expression in (4.14) represents a set of orthonormal vectors in $\mathbb{R}^{N_1} \times \mathbb{R}^{N_1}$ and they are represented below as follow;

$$\begin{bmatrix} d_{\mu_{1,E}} j_1(f_1^{(1)}(x_1)) \\ 0_{N_1} \end{bmatrix}, \begin{bmatrix} d_{\mu_{1,E}} j_1(f_2^{(1)}(x_1)) \\ 0_{N_1} \end{bmatrix} \dots, \begin{bmatrix} d_{\mu_{1,E}} j_1(f_m^{(1)}(x_1)) \\ 0_{N_1} \end{bmatrix}, \dots \begin{bmatrix} 0_{N_1} \\ d_{\mu_{2,E}} j_1(f_1^{(2)}(x_2)) \end{bmatrix}, \dots \begin{bmatrix} 0_{N_1} \\ d_{\mu_{2,E}} j_1(f_m^{(2)}(x_2)) \end{bmatrix}$$

For $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_{2N_1}$ be the canonical basis of $\mathbb{R}^{N_1} \times \mathbb{R}^{N_1}$, let $(\hat{e}_1(p_1,p_2), \hat{e}_2(p_1,p_2), \dots, \hat{e}_{2N_1}(p_1,p_2))$ be a local frame field on an open neighborhood $U \subset \mathbb{R}^{2N_1}$ containing $(j_1(\mu_{1,E}), j_1(\mu_{2,E}))$ such that $\forall\ (x_1, x_2) \in j^{-1}(U)$

$$\hat{e}_r(j(x_1,x_2)) = d_{\mu_{1,E},\mu_{2,E}} j(f_r(x_1,x_2)), \text{ for } r = 1, ..., m \tag{4.15}$$

and

$$\hat{e}_{N_1+r}(j(x_1,x_2)) = d_{\mu_{1,E},\mu_{2,E}} j(f_{m+r}(x_1,x_2)), \text{ for } r =, ..., m \tag{4.16}$$

Note that these vectors are orthonormal to each other by results of equation (4.14). Since the other elements of the local frame field $(\hat{e}_1(p_1,p_2), \hat{e}_2(p_1,p_2), \ldots, \hat{e}_{2N_1}(p_1,p_2))$ can be orthogonalized and normalized, we may now assume that $(\hat{e}_1(p_1,p_2), \hat{e}_2(p_1,p_2), \ldots, \hat{e}_{2N_1}(p_1,p_2))$ is an orthonormal frame field with elements ranging from 1 to m and from N_1+1 to N_1+m defined as in (4.15) and (4.16). It then follows that for $p=(p_1,p_2)$, $(\hat{e}_1(p), \hat{e}_2(p), \ldots, \hat{e}_{2N_1}(p))$ is an adapted frame field around $(j_1(\mu_{1,E}), j_1(\mu_{2,E})) = (P_{j_1}(\mu_1), P_{j_1}(\mu_2)) = P_j(\mu_1,\mu_2) = P_j(\hat{\mu})$. The vectors
$\hat{e}_1(P_j(\hat{\mu})), \hat{e}_2(P_j(\hat{\mu})), \ldots, \hat{e}_m(P_j(\hat{\mu})), \hat{e}_{N_1+1}(P_j(\hat{\mu})), \ldots, \hat{e}_{N_1+m}(P_j(\hat{\mu}))$ are represented below as follow;

$$\begin{bmatrix} e_1^{(1)}(P_{j_1}(\mu_1)) \\ 0_{N_1} \end{bmatrix}, \begin{bmatrix} e_2^{(1)}(P_{j_1}(\mu_1)) \\ 0_{N_1} \end{bmatrix} \ldots, \begin{bmatrix} e_m^{(1)}(P_{j_1}(\mu_1)) \\ 0_{N_1} \end{bmatrix}, \cdots \begin{bmatrix} 0_{N_1} \\ e_1^{(2)}(P_{j_1}(\mu_2)) \end{bmatrix}, \cdots \begin{bmatrix} 0_{N_1} \\ e_m^{(2)}(P_{j_1}(\mu_2)). \end{bmatrix} \tag{4.17}$$

Then

$$d_{\mu_1,\mu_2} P_j(\hat{e}_b) = (d_{\mu_1} P_{j_1}(e_b), 0_{N_1}) \in T_{P_j(\mu_1,\mu_2)} j(\mathcal{M}, \mathcal{M}), \text{ for } b = 1, \cdots, N_1$$

and

$$d_{\mu_1,\mu_2} P_j(\hat{e}_{N_1+b}) = (0_{N_1}, d_{\mu_2} P_{j_1}(e_b)) \in T_{P_j(\mu_1,\mu_2)} j(\mathcal{M}, \mathcal{M}), \text{ for } b = 1, \cdots, N_1$$

are linear combinations of $\hat{e}_1(P_j(\hat{\mu})), \hat{e}_2(P_j(\hat{\mu})), \ldots, \hat{e}_m(P_j(\hat{\mu})), \hat{e}_{N_1+1}(P_j(\hat{\mu})), \ldots, \hat{e}_{N_1+m}(P_j(\hat{\mu}))$
Note that

$$(d_{\mu_1} P_{j_1}(e_b), 0_{N_1}) \cdot \hat{e}_a(P_j(\hat{\mu})) = 0$$

for $a = m+1, \cdots, 2N_1$ and $b = 1, \cdots, m$

$$(0_{N_1}, d_{\mu_2} P_{j_1}(e_b)) \cdot \hat{e}_a(P_j(\hat{\mu})) = 0$$

$a = N_1 + m + 1, \cdots, 2N_1$ and $a = 1, ..., N_1$ and $b = 1, \cdots, m$
It then follow that the tangential component of $\left(P_j(\overline{j(X_{1,1})}, \overline{j(X_{2,1})}) - P_j(\mu_1,\mu_2)\right) \in \mathbb{R}^{2N_1}$ with respect to the basis $\hat{e}_1(P_j(\hat{\mu})), \hat{e}_2(P_j(\hat{\mu})), \ldots, \hat{e}_m(P_j(\hat{\mu})), \hat{e}_{N_1+1}(P_j(\hat{\mu})), \ldots, \hat{e}_{N_1+m}(P_j(\hat{\mu}))$ has the following asymptotic behavior;

$$n^{1/2} \, \tan_{P_j(\hat{\mu})} \left(P_j(\overline{j(X_{1,1})}, \overline{j(X_{2,1})}) - P_j(\mu_1,\mu_2)\right) \to_d \mathcal{N}_{2m}(0_{2m}, \Sigma^{(2)}_{j_1,E}), \tag{4.18}$$

where

$$\Sigma^{(2)}_{j_1,E} = [A^{(2)}]^T \Sigma^\star \, A^{(2)}$$

where $A^{(2)}$ is a $2N_1 \times 2m$ matrix given by;

$$A^{(2)} = \begin{pmatrix} e_1^{(1)}(P_{j_1}(\mu_1)) & \cdots & e_m^{(1)}(P_{j_1}(\mu_1)) & | & 0_{N_1} & \cdots & 0_{N_1} \\ -- & -- & -- & | & -- & -- & -- \\ 0_{N_1} & \cdots & 0_{N_1} & | & e_1^{(2)}(P_{j_1}(\mu_2)) & \cdots & e_m^{(2)}(P_{j_1}(\mu_2)) \end{pmatrix} \tag{4.19}$$

$$A^{(2)} = \begin{pmatrix} A_1 & | & A_2 \end{pmatrix}$$

And we have

$$\Sigma_{j_1,E}^{(2)} = \begin{pmatrix} \frac{1}{\pi}\Sigma_{1,E} & 0_m \\ 0_m & \frac{1}{1-\pi}\Sigma_{2,E} \end{pmatrix} \tag{4.20}$$

For part (ii), we will rely on colorblue Theorem (4.1.1) with $f_1(x_1, x_2), \ldots, f_{2m}(x_1, x_2)$ defined in (4.13), as our orthonormal basis in $T_{(\mu_{1,E},\mu_{2,E})}(\mathcal{M}^2)$ and its corresponding embedding is $j : \mathcal{M}^2 \to \mathbb{R}^{2N_1}$. We will also let $(g_1, \cdots, g_n)$ be an orthonormal basis in $T_{G(\mu_{1,E},\mu_{2,E})}(\mathcal{N})$ with embedding $j_2 : \mathcal{N} \to \mathbb{R}^{N_2}$ and $(\tilde{e}_1(\tilde{G}(\mu_1, \mu_2)), \cdots, \tilde{e}_n(\tilde{G}(\mu_1, \mu_2)))$ is adapted to the embedding j_2 on $\mathcal{N}$ and is such that;

$$\tilde{e}_s(\tilde{G}(\mu_1, \mu_2)) = d_y j_2(g_s),\ with\, y = G(\mu_{1,E}, \mu_{2,E}),\ and\, s = 1, ...n,\ with\, \tilde{G} = j_2 \circ G \circ j_1^{-1}(P_{j_1}) \times j_1^{-1}(P_{j_1})$$

With our result in part (i) we now appeal to the Theorem and we get the following asymptotic behavior;

$$n^{1/2}\ \tan_{j_2(G(\mu_{1,E},\mu_{2,E}))}\left(j_2\left(G(\overline{X}_{1,E}, \overline{X}_{2,E})\right) - j_2\left(G(\mu_{1,E}, \mu_{2,E})\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}_n(0, \Sigma_{j_2,E}^G)$$

and $\Sigma_{j_2,E}^G = (B^\star A^{(2)})\, \Sigma_{j_1,E}^{(2)}\, (B^\star A^{(2)})^T$ with $B^\star A^{(2)} = [B^{(1)} A_1 \mid B^{(2)} A_2]$ and for

$$\tilde{G} = j_2 \circ G \circ j_1^{-1}(P_{j_1}) \times j_1^{-1}(P_{j_1}) : \mathcal{F}^c \times \mathcal{F}^c \to j_2(\mathcal{N})$$

where $\mathcal{F}^c$ is the set of j_1-nonfocal points. Let $\hat{e}_1, ..., \hat{e}_{2N_1}$ be the canonical basis of $\mathbb{R}^{2N_1}$. And for $\tilde{e}_1(p_2), ..., \tilde{e}_n(p_2)$, for $p_2 \in j_2(V_2)$.

$$B^{(1)} A_1 = \begin{bmatrix} \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_1)\right) \cdot \tilde{e}_1(\tilde{G}(\hat{\mu})) & ... & \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_m)\right) \cdot \tilde{e}_1(\tilde{G}(\hat{\mu})) \\ \vdots & & \\ \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_1)\right) \cdot \tilde{e}_n(\tilde{G}(\hat{\mu})) & ... & \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_m)\right) \cdot \tilde{e}_n(\tilde{G}(\hat{\mu})) \end{bmatrix}$$

and

$$B^{(2)} A_2 = \begin{bmatrix} \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_{N_1+1})\right) \cdot \tilde{e}_1(\tilde{G}(\hat{\mu})) & ... & \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_{N_1+m})\right) \cdot \tilde{e}_1(\tilde{G}(\hat{\mu})) \\ \vdots & & \\ \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_{N_1+1})\right) \cdot \tilde{e}_n(\tilde{G}(\hat{\mu})) & ... & \left(d_{\hat{\mu}}\tilde{G}(\hat{e}_{N_1+m})\right) \cdot \tilde{e}_n(\tilde{G}(\hat{\mu})) \end{bmatrix} \tag{4.21}$$

Letting $dG^{(1)} = B^{(1)}A_1$ and $dG^{(2)} = B^{(2)}A_2$ we have

$$\Sigma^G_{j_2,E} = \frac{1}{\pi}\,(dG^{(1)})\,\Sigma_{1,E}(dG^{(1)})^T + \frac{1}{1-\pi}(dG^{(2)})\Sigma_{2,E}(dG^{(2)})^T \tag{4.22}$$

□

DEFINITION 4.2.1. *The matrix $\Sigma^G_{j_2,E}$ given in* (4.22) *is the extrinsic covariance matrix of the j_2-nonfocal distribution Q_2 (of $G(X_{1,1}, X_{2,1})$) w.r.t the orthonormal basis $g_1(G(\mu_{E,1}, \mu_{E,2})), \dots, g_n(G(\mu_{E,1}, \mu_{E,2}))$ written in term of the extrinsic covariance matrices $\Sigma_{1,E}$ and $\Sigma_{2,E}$ of $X_{1,1}$ and $X_{2,1}$ respectively and where for $a = 1, 2$ $\Sigma_{a,E}$ is expressed w.r.t the orthonormal basis $f^{(a)}_1(\mu_{a,E}), \dots, f^{(a)}_m(\mu_{a,E})$.*

THEOREM 4.2.2. *For $a = 1, 2$, let $X_{a,1}, \cdots, X_{a,n_a}$ be independent random samples defined on $\mathcal{G}$, an m-dimensional Lie group, from a distribution Q_a, with the extrinsic means $\mu_{1,E}, \mu_{2,E}$ and corresponding extrinsic covariance matrices $\Sigma_{1,E}, \Sigma_{2,E}$ and respective extrinsic sample mean $\overline{X}_{1,E}$ and $\overline{X}_{2,E}$. Let $\hat{j} : \mathcal{G} \to \mathbb{R}^N$ be an embedding on $\mathcal{G}$ and for $a = 1, 2$ let $f^{(a)}_1, \cdots, f^{(a)}_m$ be orthonormal basis in $T_{\mu_{a,E}}(\mathcal{G})$. Furthermore for $n = n_1 + n_2$, if $\frac{n_1}{n} \to \pi$ as $n_a \to \infty$. Let $g_1, \cdots, g_m$ be an orthonormal basis in $T_{\mu^{-1}_{2,E}\odot\mu_{1,E}}(\mathcal{G})$ we have the following*

$$n^{1/2}\tan_{\hat{j}(\mu^{-1}_{2,E}\odot\mu_{1,E})}\left(\hat{j}(\overline{X}^{-1}_{2,E}\odot\overline{X}_{1,E}) - \hat{j}(\mu^{-1}_{2,E}\odot\mu_{1,E})\right) \to_d N_m(0_m, \Sigma^{\iota H}_E) \tag{4.23}$$

were $H : \mathcal{G}\times\mathcal{G} \to \mathcal{G}$ and is given by $H(\overline{X}^{-1}_{2,E}, \overline{X}_{1,E}) = \overline{X}^{-1}_{2,E}\odot\overline{X}_{1,E}$, then we have,

$$\Sigma^{\iota H}_E = \frac{1}{\pi}\,(dH^{(1)})\,\Sigma^{\iota}_{2,E}(dH^{(1)})^T + \frac{1}{1-\pi}(dH^{(2)})\Sigma_{1,E}(dH^{(2)})^T \tag{4.24}$$

where

$$\begin{aligned} dH^{(1)} &= \left(dH^{(1)}_{a,b}\right) = \left(d_{\hat{\mu}}\hat{H}(\hat{e}_b)\cdot\tilde{e}_a(\hat{H}(\hat{\mu}))\right) \\ dH^{(2)} &= \left(dH^{(2)}_{a,b}\right) = \left(d_{\hat{\mu}}\hat{H}(\hat{e}_{N_1+b})\cdot\tilde{e}_a(\hat{H}(\hat{\mu}))\right),\quad for\ a,b = 1,...,m \end{aligned}$$

where $\hat{H} \equiv \hat{j}\circ H\circ\hat{j}^{-1}(\tilde{\iota}\circ P_{\hat{j}})\times\hat{j}^{-1}(P_{\hat{j}}) : \mathcal{F}^c\times\mathcal{F}^c \to \hat{j}(\mathcal{M})$.

Proof. Recall that for $X_{1,1}, \cdots, X_{1,n_1}$ independent random samples defined on $\mathcal{G}$ we have the following asymptotic behavior

$$\tan_{\hat{j}(\mu_{1,E})}\left(\hat{j}(\overline{X}_{1,E}) - \hat{j}(\mu_{1,E})\right) \to_d N_m(0_m, \Sigma_{1,E}) \tag{4.25}$$

and for the other independent random samples, $X_{2,1}, \cdots, X_{2,n_2}$ we have, after applying Theorem (4.1.1), the following asymptotic behavior;

$$\tan_{\hat{j}(\mu^{-1}_{2,E})}\left(\hat{j}(\overline{X}^{-1}_{2,E}) - \hat{j}(\mu^{-1}_{2,E})\right) \to_d N_m(0_m, \Sigma^{\iota}_{2,E}) \tag{4.26}$$

$$\Sigma_{2,E}^{\iota} = (dI)\, \Sigma_{2,E}\, (dI)^T$$

and

$$dI = \begin{bmatrix} (d_{\mu_2}\tilde{\iota} \circ P_j(e_1)) \cdot \tilde{e}_1(\tilde{\iota} \circ P_j(\mu_2)) & \ldots & (d_{\mu_2}\tilde{\iota} \circ P_j(e_m)) \cdot \tilde{e}_1(\tilde{\iota} \circ P_j(\mu_2)) \\ & \vdots & \\ (d_{\mu_2}\tilde{\iota} \circ P_j(e_1)) \cdot \tilde{e}_m(\tilde{\iota} \circ P_j(\mu_2)) & \ldots & (d_{\mu_2}\tilde{\iota} \circ P_j(e_m)) \cdot \tilde{e}_m(\tilde{\iota} \circ P_j(\mu_2)) \end{bmatrix}$$

Not that for $a = 1, 2$ μ_a is the mean of $j(Q_a)$ and where $\hat{j} \circ \iota \circ \hat{j}^{-1} = \tilde{\iota}$ and the new covariance matrix $\Sigma_{2,E}^{\iota}$ is the extrinsic covariance matrix with respect to the local frame field $(f_1^{\iota}, ..., f_m^{\iota})$ defined on $W_2 \in \mathcal{G}$. Note that W_2 is an open neighborhood of $\iota(\mu_{2,E}) = \mu_{2,E}^{-1}$ and $V_2 = \iota^{-1}(W_2)$ is the open neighborhood of $\mu_{2,E}$ on which the local frame field $(f_1^{(2)}, ..., f_m^{(2)})$ is defined. Furthermore, for points $p_1 \in \hat{j}(V_1)$, and $p_2 \in \hat{j}(V_2)$, with $\tilde{\iota}(p_2) \in \hat{j}(W_2)$, we have

$$\begin{gathered} e_1^{(1)}(p_1), \cdots, e_N^{(1)}(p_1) \\ e_1^{(2)}(\tilde{\iota}(p_2)), \cdots, e_N^{(2)}(\tilde{\iota}(p_2)) \end{gathered}$$

respectively the adapted frame field around $\hat{j}(\mu_{1,E})$ and $\hat{j}(\mu_{2,E}^{-1})$.

We then get the following combined asymptotic behavior;

$$n^{1/2}\ \tan_{\hat{j}^{(2)}(\mu_{2,E}^{-1}, \mu_{1,E})} \left(\hat{j}^{(2)}(\overline{X}_{2,E}^{-1}, \overline{X}_{1,E}) - \hat{j}^{(2)}(\mu_{2,E}^{-1}, \mu_{1,E}) \right) \xrightarrow{\mathcal{L}} \mathcal{N}_{2m}(0, \Sigma_E^{(2)})$$

where $\Sigma_E^{(2)} = \begin{pmatrix} \frac{1}{\pi}\Sigma_{2,E}^{\iota} & 0_m \\ 0_m & \frac{1}{1-\pi}\Sigma_{1,E} \end{pmatrix}$

Here, $\Sigma_E^{(2)}$ is the extrinsic covariance matrix with respect to the local frame field $f_1(y_2, x_1), \cdots, f_{2m}(y_2, x_1)$ around $(\mu_{2,E}^{-1}, \mu_{1,E}) \in \mathcal{G} \times \mathcal{G}$. And $(\hat{e}_1(\tilde{\iota}(p_2), p_1), \hat{e}_2(\tilde{\iota}(p_2), p_1), \ldots, \hat{e}_{2N}(\tilde{\iota}(p_2), p_1))$ is the adapted frame field around $(\hat{j}(\mu_{2,E}^{-1}), \hat{j}(\mu_{1,E}))$. And now for $P_{\hat{j}}^{\iota} = \tilde{\iota} \circ P_{\hat{j}}$ with $\hat{e}_1, ..., \hat{e}_N, ..., \hat{e}_{2N}$ the canonical basis in $\mathbb{R}^{2N}$ we have,

$$d_{\mu_2,\mu_1} P_{\hat{j}}^{\iota} \times P_{\hat{j}}(\hat{e}_b) = (d_{\mu_2}\tilde{\iota} \circ P_{j_1}(e_b), 0_N) = (d_{\mu_{2,E}^{-1}} \hat{j}(f_b^{\iota}(y_2)), 0_N) \in T_{P_{\hat{j}}^{\iota} \times P_{\hat{j}}(\mu_{21})} \hat{j}^{(2)}(\mathcal{G}, \mathcal{G}),$$

and

$$d_{\mu_2,\mu_1} P_{\hat{j}}^{\iota} \times P_{\hat{j}}(\hat{e}_{N_1+b}) = (0_N, d_{\mu_1} P_{\hat{j}}(e_b)) \in T_{P_{\hat{j}}^{\iota} \times P_{\hat{j}}(\mu_{21})} \hat{j}^{(2)}(\mathcal{M}, \mathcal{M}),\ \text{for}\ \ b = 1, \cdots, N$$

And $e_b,\ b=1,\cdots,N$ represent the canonical basis for $\mathbb{R}^N$. These tangent vectors in $T_{P^{\iota}_{\hat{j}}\times P_{\hat{j}}(\mu_2,\mu_1)}\hat{j}^{(2)}(\mathcal{M},\mathcal{M})$ are linear combinations of the vectors

$$\hat{e}_1(P^{\iota}_{\hat{j}}\times P_{\hat{j}}(\mu_2,\mu_1)),\ldots,\hat{e}_m(P^{\iota}_{\hat{j}}\times P_{\hat{j}}(\mu_2,\mu_1)),\hat{e}_{N+1}(P^{\iota}_{\hat{j}}\times P_{\hat{j}}(\mu_2,\mu_1)),\ldots,\hat{e}_{N+m}(P^{\iota}_{\hat{j}}\times P_{\hat{j}}(\mu_2,\mu_1))$$

Now we may use the results from part (ii) of Theorem (4.2.1). Let $g_1,\cdots,g_m$ be an orthonormal basis in $T_{\mu_{2,e}^{-1}\odot\mu_{1,E}}(\mathcal{G})$ and a local frame field $\tilde{e}_1(\hat{H}(\hat{\mu})),\cdots,\tilde{e}_N(\hat{H}(\hat{\mu}))$ adapted to the embedding $\hat{j}$ with

$$\tilde{e}_s(\hat{H}(\hat{\mu}))=d_{\mu_{2,E}^{-1}\odot\mu_{1,E}}\hat{j}(g_s),\quad s=1,\cdots,m$$

We have the following asymptotic behavior,

$$n^{1/2}\tan_{\hat{j}(\mu_{2,E}^{-1}\odot\mu_{1,E})}\left(\hat{j}(\overline{X}_{2,E}^{-1}\odot\overline{X}_{1,E})-\hat{j}(\mu_{2,E}^{-1}\odot\mu_{1,E})\right)\to_d N_m(0_m,\Sigma_E^{\iota H}) \tag{4.27}$$

$$\Sigma_E^{\iota H}=\frac{1}{\pi}\,(dH^{(1)})\,\Sigma_{2,E}^{\iota}(dH^{(1)})^T+\frac{1}{1-\pi}(dH^{(2)})\Sigma_{1,E}(dH^{(2)})^T \tag{4.28}$$

And for $\hat{H}=\hat{j}\circ H\circ\hat{j}^{-1}(\tilde{\iota}\circ P_{\hat{j}})\times\hat{j}^{-1}(P_{\hat{j}}):\mathcal{F}^c\times\mathcal{F}^c\to\hat{j}(\mathcal{M})$.

$$\begin{aligned}dH^{(1)}&=\left(dH^{(1)}_{a,b}\right)=\left(d_{\hat{\mu}}\hat{H}(\hat{e}_b)\cdot\tilde{e}_a(\hat{H}(\hat{\mu}))\right)\\ dH^{(2)}&=\left(dH^{(2)}_{a,b}\right)=\left(d_{\hat{\mu}}\hat{H}(\hat{e}_{N_1+b})\cdot\tilde{e}_a(\hat{H}(\hat{\mu}))\right),\ \ for\ a,b=1,...,m\end{aligned}$$

□

Recall the following hypothesis testing problem,

$$H_0:\ \mu_{2,E}^{-1}\odot\mu_{1,E}=\delta\ vs.\ H_1:\ \mu_{2,E}^{-1}\odot\mu_{1,E}\neq\delta$$

we get the following corollary.

COROLLARY 4.2.1. *Under the assumptions of Theorem 4.2.2 and also assuming that $j(X_{a,1})$ for $a=1,2$ have finite second order moments and the extrinsic covariance matrices $\Sigma_{a,E}$ are nonsingular, then for $n=n_1+n_2$ large enough the sample extrinsic covariance matrices S_{a,E,n_a} are nonsingular (with probability converging to one) and*

(a) The statistics

$$n\|S_{\iota,H}^{-1/2}\ \tan_{j(\delta)}\left(j\left(\overline{X}_{2,E}^{-1}\odot\overline{X}_{2,E}\right)-j(\delta)\right)\|^2\xrightarrow{\mathcal{L}}\chi^2_n \tag{4.29}$$

$$n\|S_{\iota,H}^{-1/2}\ \tan_{j_2(\overline{X}_{2,E}^{-1}\odot\overline{X}_{2,E})}\left(j\left(\overline{X}_{2,E}^{-1}\odot\overline{X}_{2,E}\right)-j(\delta)\right)\|^2\xrightarrow{\mathcal{L}}\chi^2_n \tag{4.30}$$

(b) and a confidence region for $\mu_{2,E}^{-1} \odot \mu_{1,E}$ *of asymptotic level* $1-\alpha$ *is given by*
$(i) C_{n,\alpha}^{\iota,H} := j^{-1}(U_{n,\alpha}^{\iota,H})$,
where $U_{n,\alpha}^{\iota,H} = \{\nu \in j(\mathcal{G}) : n\|S_{\iota,H}^{-1/2}\ \tan_{\nu}\left(j\left(\overline{X}_{2,E}^{-1} \odot \overline{X}_{2,E}\right) - \nu\right)\|^2 \leq \chi^2_{n,1-\alpha}\}$.
Another such confidence region can also be given by
$(ii) D_{n,\alpha}^{\iota,H} := j^{-1}(V_{n,\alpha}^{\iota,H})$ *where*
$V_{n,\alpha}^{\iota,H} = \{\nu \in j(\mathcal{G}) : n\|S_{\iota,H}^{-1/2}\ \tan_{j(\overline{X}_{2,E}^{-1} \odot \overline{X}_{2,E})}\left(j\left(\overline{X}_{2,E}^{-1} \odot \overline{X}_{2,E}\right) - \nu\right)\|^2 \leq \chi^2_{n,1-\alpha}\}$.

where $S_{\iota,H} = \frac{1}{n_2}\,(dH_e^{(1)})\,G_{2,E}^{\iota}(dH_e^{(1)})^T + \frac{1}{n_1}(dH_e^{(2)})G_{1,E}(dH_e^{(2)})^T$ and

$$dH_e^{(1)} = \left(d_{\hat{\mathbf{x}}_{\hat{j}}}\hat{H}(\hat{e}_b) \cdot \tilde{e}_a(\hat{H}(\hat{\mathbf{x}}_{\hat{j}}))\right)$$
$$dH_e^{(2)} = \left(d_{\hat{\mathbf{x}}_{\hat{j}}}\hat{H}(\hat{e}_{N_1+b}) \cdot \tilde{e}_a(\hat{H}(\hat{\mathbf{x}}_{\hat{j}}))\right)$$

For $a, b = 1, ..., m$ and $\hat{\mathbf{x}}_{\hat{j}} = \left(\overline{j(X_2)}, \overline{j(X_1)}\right)$

4.3 3D real projective space $\mathbb{R}P^3$

For $[X_r],\ \|X_r\| = 1,\ r = 1, ..., n$, a random sample from a VW-nonfocal probability measure Q on $\mathbb{R}P^3$, let μ_E be the VW mean and $[\overline{X}_E]$ its VW sample mean with the corresponding extrinsic covariance matrix Σ_E. We have the following asymptotic behavior

$$\tan_{\hat{j}(\mu_E^{-1})}\left(\hat{j}([\overline{X}_E]^{-1}) - \hat{j}(\mu_E^{-1})\right) \to_d N_m(0_m, \Sigma_E^{\iota})$$

where $\Sigma_E^{\iota} = (dI)\ \Sigma_E\ (dI)^T$ and $dI_{a,b} = (d_\mu \tilde{\iota} \circ P_j(e_b)) \cdot \tilde{e}_b(\tilde{\iota} \circ P_j(\mu))\ \ a, b = 1, 2, 3$. And ι is the inverse map of the Lie group $\mathbb{R}P^3$.

PROPOSITION 4.3.1. *Assume* $[X_r],\ \|X_r\| = 1,\ r = 1, ..., n$, *is a random sample from a VW-nonfocal probability measure* Q *on* $\mathcal{G} = \mathbb{R}P^3$ *a* 3*-dimensional Lie group. Also let* $\iota : \mathbb{R}P^3 \to \mathbb{R}P^3$ *be the inverse map on that manifold. The sample covariance matrix* $G_E^{\iota}(j, X)$*, which is the consistent estimator of* Σ_E^{ι}*, has entries given by;*

$$G_E^{\iota}(j, X)_{a,b} = n^{-1}(\eta_4 - \eta_a)^{-2}(\eta_4 - \eta_b)^{-2} \times \sum_r (m_a \cdot X_r)(m_b \cdot X_r)(m_4 \cdot X_r)^2 \qquad (4.31)$$

where $\eta_a,\ a = 1, .., 4$ *are eigenvalues of* $K = n^{-1}\sum_{r=1}^n X_r X_r^T$ *in increasing order and* $m_a = 1, ..., 4$*, are corresponding linearly independent unit eigenvectors.*

Proof. Note that since $\overline{j([X])}$ is a consistent estimator of μ the mean of $j([X_1]) \in S(4, \mathbb{R})$. Also for the orthonormal frame field $(e_1(P_j(\mu)), e_2(P_j(\mu)), e_3(P_j(\mu)))$ on a subset of $\mathbb{R}P^3$ with $P_j(\mu) = j(\overline{X}_E)$ we have that for $a = 1, 2, 3$, $e_a(P_j(j(\overline{[X]}))$ is a consistent estimator of $e_a(P_j(\mu))$. Similarly, $d_{\overline{j([X])}}P_j$is a consistent estimator of $d_\mu P_j$.

For the orthonormal frame field $(\tilde{e}_1(\tilde{\iota} \circ P_j(\mu)), \tilde{e}_2(\tilde{\iota} \circ P_j(\mu)), \tilde{e}_3(\tilde{\iota} \circ P_j(\mu)))$ we also have the corresponding consistent estimator $(\tilde{e}_1(\tilde{\iota} \circ P_j(\overline{j([X])})), \tilde{e}_2(\tilde{\iota} \circ P_j(\overline{j([X])})), \tilde{e}_3(\tilde{\iota} \circ P_j(\overline{j([X])})))$. And $d_\mu \tilde{\iota} \circ P_j$ has the following consistent estimator $d_{\overline{j([X])}} \tilde{\iota} \circ P_j$

Now recall that

$$\Sigma_E^\iota = (dI)\ \Sigma_E\ (dI)^T$$

$$(dI)_{a,b} = d_\mu \tilde{\iota} \circ P_j(e_b) \cdot \tilde{e}_a(\tilde{\iota} \circ P_j(\mu))$$

for $a, b = 1, 2, 3$. And Σ_E is the extrinsic covariance matrix. Let $j([\overline{X}_E]) = P_j(\overline{j([X])})$ then we would like to first investigate the case for which $\overline{j([X])} = D$ be a diagonal matrix. If this matrix is diagonal we get $[m_4] = [e_4] = [\overline{X}_E]$ and we get the consistent estimator of Σ_E denoted $G_E(j, X)$ and with entries given by

$$G_E(j, X)_{ab} = n^{-1}(\eta_4 - \eta_a)^{-1}(\eta_4 - \eta_b)^{-1} \sum_r X_r^a X_r^b (X_r^4)^2 \tag{4.32}$$

where $\eta_a,\ a = 1, .., 4$ are eigenvalues of $K = n^{-1} \sum_{r=1}^n X_r X_r^T$ in increasing order and $m_a = 1, ..., 4$, are corresponding linearly independent unit eigenvectors. We can now express our consistent estimator $G_E^\iota(j, X)$ as follow

$$G_E^\iota(j, X) = (d\psi)\ G_E(j, X)\ (d\psi)^T$$

where $d\psi$ is a matrix with entries

$$d\psi_{a,b} = d_D \tilde{\iota} \circ P_j(e_b) \cdot \tilde{e}_a(\tilde{\iota} \circ P_j(D))$$

for $a, b = 1, 2, 3$. $S(4, \mathbb{R})$ has the orthonormal basis $F_a^b,\ b \leq a$, where, for $a < b$, the matrix F_a^b has all entries zeros except for those in the positions $(a, b), (b, a)$ that are equal to $2^{-1/2}$; also $F_a^a = j([e_a])$. Recall from proposition 4.2 in Battacharya and Patrangenaru 2005, that we have

$$\begin{aligned} & d_D P_j(F_a^b) = 0, \forall\, b \leq a < 4 \\ \Longrightarrow\ & d_D \tilde{\iota} \circ P_j(F_a^b) = d_{P_j(D)} \tilde{\iota}\ d_D P_j(F_a^b) = 0, \forall\, b \leq a < 4 \end{aligned}$$

Note that $[\overline{X}_E] = [m_4] = [e_4]$ and the other unit eigenvectors of $D = \overline{j([X])}$ are $m_a = e_a, \ \forall\, a = 1, 2, 3$. Since $j([\overline{X}_E]^{-1}) = \tilde{\iota} \circ P_j(D)$, we want to evaluate $d_D \tilde{\iota} \circ P_j(F_a^b) \in T_{\tilde{\iota} \circ P_j(D)} j(\mathcal{G})$. But given that

$$[\overline{X}_E]^{-1} = [e_4]^{-1} = [\bar{e}_4] = [e_4] = [\overline{X}_E]$$

we then have the following choice of orthonormal frame

$$\tilde{e}_a(\tilde{\iota} \circ P_j(D)) = \tilde{e}_a(j([\overline{X}_E]^{-1})) = d_{\overline{X}_E} j(e_a) = d_{[e_4]} j(e_a)$$

We will now compute the remaining 3 tangent vectors in $T_{P_j(D)} j(\mathbb{R}P^3)$ of interest, namely;

$$d_D \ \tilde{\iota} \circ P_j(e_a) = d_D \ \tilde{\iota} \circ P_j(F_4^a), \ for \ a = 1, 2, 3.$$

And for $a = 1, 2, 3$, direct computations

$$d_\mu \tilde{\iota} \circ P_j(F_N^a) = \left. \frac{d}{dt} \tilde{\iota} \circ P_j(D + tF_N^a) \right|_{t=0}$$

will yield

$$d_D \ \tilde{\iota} \circ (e_1) = (\eta_1 - \eta_4)^{-1} \ \tilde{e}_1(P_j(D))$$
$$d_D \ \tilde{\iota} \circ (e_2) = (\eta_2 - \eta_4)^{-1} \ \tilde{e}_2(P_j(D))$$
$$d_D \ \tilde{\iota} \circ (e_3) = (\eta_3 - \eta_4)^{-1} \ \tilde{e}_3(P_j(D))$$

we then have the following

$$d\psi = \begin{bmatrix} (\eta_1 - \eta_4)^{-1} & 0 & 0 \\ 0 & (\eta_2 - \eta_4)^{-1} & 0 \\ 0 & 0 & (\eta_3 - \eta_4)^{-1} \end{bmatrix}$$

Hence, the matrix $G_E^\iota(j, X)$ has entries;

$$G_E^\iota(j, X)_{a,b} = n^{-1}(\eta_4 - \eta_a)^{-2}(\eta_4 - \eta_b)^{-2} \times \sum_r X_r^a X_r^b (X_r^4)^2$$

□

For $a = 1, 2$ let $[X_{a,1}], \cdots, [X_{a,n_a}]$ be independent random samples defined on $\mathbb{R}P^3$ from j-nonfocal distributions Q_a, with extrinsic means $\mu_{a,E}$ and extrinsic covariance matrices $\Sigma_{a,E}$. Also let $n = n_1 + n_2$ such that $n_1/n \to \pi$ as $n_a \to \infty$ $a = 1, 2$. Then using the result of Lemma 4.2.2 we have for, $\iota : \mathbb{R}P^3 \to$

$\mathbb{R}P^3$ the inverse map and $H : \mathbb{R}P^3 \times \mathbb{R}P^3 \to \mathbb{R}P^3$ the Lie group multiplication, the following asymptotic behavior.

$$n^{1/2}\ \tan_{j(\mu_{2,E}^{-1} \odot \mu_{1,E})} \Big(j(\overline{X}_{2,E}^{-1} \odot \overline{X}_{1,E}) - j(\mu_{2,E}^{-1} \odot \mu_{1,E}) \Big) \to_d N_m(0_m, \Sigma_E^{\iota G}) \tag{4.33}$$

where for $H(\mu_{2,E}^{-1}, \mu_{1,E}) = (\mu_{2,E}^{-1} \odot \mu_{1,E})$,

$$\Sigma_E^{\iota H} = \frac{1}{\pi}\,(dH^{(1)})\,\Sigma_{2,E}^{\iota}(dH^{(1)})^T + \frac{1}{1-\pi}(dH^{(2)})\Sigma_{1,E}(dH^{(2)})^T \tag{4.34}$$

PROPOSITION 4.3.2. *For $a = 1, 2$, let $\{[X_{r_a}]\}_{r_a=1}^{n_a}$, $\|X_{r_a}\| = 1$, be independent random samples from j-nonfocal probability measures Q_a on $\mathbb{R}P^3$. Then the consistent estimator of Σ_E^{ι} is denoted $G_E^{\iota}(j, X_{1,1}, X_{2,1})$ with extrinsic means and covariance respectively $\mu_{a,E}$ and $\Sigma_{a,E}$. Also let $\iota : \mathbb{R}P^3 \to \mathbb{R}P^3$ be the inverse map on that manifold and $\circ$ denote the Lie group multiplication on $\mathbb{R}P^3$. The sample covariance matrix $G_E^{\iota}(X)$, which is the consistent estimator of Σ_E^{ι}, has entries given by;*

$$\begin{gathered} G_E^{\iota H}(j, X_{1,1}, X_{2,1})_{a,b} = \\ n_2^{-1}(\eta_{2,4} - \eta_{2,a})^{-3}(\eta_{2,4} - \eta_{2,b})^{-3} \times \sum_{r=1}^{n_2}(m_{2,a} \cdot X_{2,r}^a)(m_{2,b} \cdot X_{2,r}^b)(m_{2,4} \cdot X_{2,r}^4)^2 \\ + \\ n_1^{-1}(\eta_{1,4} - \eta_{1,a})^{-2}(\eta_{1,4} - \eta_{1,b})^{-2}\sum_{r=1}^{n_1}(m_{1,a} \cdot X_{1,r}^a)(m_{1,b} \cdot X_{1,r}^b)(m_{1,4} \cdot X_{1,r}^4)^2 \end{gathered} \tag{4.35}$$

where for $s = 1, 2$ and $\eta_{s,a},\ a = 1, .., 4$ are eigenvalues of $K_s = n_s^{-1}\sum_{r=1}^{n_s} X_{s,r}X_{s,r}^T$ in increasing order and $m_{s,a} = 1, ..., 4$, are corresponding linearly independent unit eigenvectors.

Proof. And for $\Sigma_{1,E}$ and $\Sigma_{2,E}^{\iota}$ are the extrinsic covariance matrices of $X_{1,1}$ and $X_{2,1}$ respectively. Without loss of generality, we now assume that $j([\overline{X}_{a,E}]) = P_j(\overline{j([X_{a,1}])})$ is a diagonal matrix, and lets take $\overline{j([X_{a,1}])} = D_a$ to be a diagonal matrix as well.

We then have the consistent estimators of $\Sigma_{2,E}^{\iota}$ and $\Sigma_{1,E}$ denoted $G_{2,E}^{\iota}(j, X_{2,1})$ and $G_{1,E}(j, X_{1,1})$ and with entries given by .

$$G_{2,E}^{\iota}(j, X_{2,1})_{a,b} = n_2^{-1}(\eta_{2,4} - \eta_{2,a})^{-2}(\eta_{2,4} - \eta_{2,b})^{-2} \times \sum_{r=1}^{n_2} X_{2,r}^a X_{2,r}^b (X_{2,r}^4)^2$$

$$G_{1,E}(j, X_{1,1})_{ab} = n_1^{-1}(\eta_{1,4} - \eta_{1,a})^{-1}(\eta_{1,4} - \eta_{1,b})^{-1}\sum_{r=1}^{n_1} X_{1,r}^a X_{1,r}^b (X_{1,r}^4)^2 \tag{4.36}$$

where for $s=1,2$ and $\eta_{s,a},\ a=1,..,4$ are eigenvalues of $K_s = n_s^{-1}\sum_{r=1}^{n_s} X_{s,r}X_{s,r}^T$ in increasing order and $m_{s,a}=1,...,4$, are corresponding linearly independent unit eigenvectors.

Now the extrinsic covariance matrix

$$\Sigma_E^{\iota H} = \frac{1}{\pi}\,(dH^{(1)})\,\Sigma_{2,E}^{\iota}(dH^{(1)})^T + \frac{1}{1-\pi}(dH^{(2)})\Sigma_{1,E}(dH^{(2)})^T \tag{4.37}$$

has the following consistent estimator

$$G_E^{\iota H}(j, X_{1,1}, X_{2,1}) = \frac{1}{n_2}\,(d\Gamma^{(1)})\,G_{2,E}^{\iota}(j, X_{2,1})(d\Gamma^{(1)})^T + \frac{1}{n_1}(d\Gamma^{(2)})G_{1,E}(j, X_{1,1})(d\Gamma^{(2)})^T \tag{4.38}$$

where $d\Gamma^{(1)}$ and $d\Gamma^{(2)}$ are matrices with entries given by

$$\begin{aligned} d\Gamma_{a,b}^{(1)} &= \left(d_{(D_2,D_1)}\hat{H}(\hat{e}_b)\cdot\tilde{e}_a(\hat{H}(D_2,D_1))\right) \\ d\Gamma_{a,b}^{(2)} &= \left(d_{D_2,D_1}\hat{H}(\hat{e}_{N_1+b})\cdot\tilde{e}_a(\hat{H}(D_2,D_1))\right),\ \ for\ a,b=1,2,3 \end{aligned}$$

where $\hat{D}=(D_2,D_1)$ and for $a=1,2$ $D_a\in\ S(4,\mathbb{R})$. Recall that $S(4,\mathbb{R})$ has the orthonormal basis $F_a^b,\ b\le a$, where, for $a<b$, the matrix F_a^b has all entries zeros except for those in the positions $(a,b),(b,a)$ that are equal to $2^{-1/2}$; also $F_a^a = j([e_a])$. We have that $\hat{D}\in S(4,\mathbb{R})\times S(4,\mathbb{R})$ and a convenient basis for such a manifold is $(F_{2,a}^b, 0_{4\times4})$ for $a,b=1,...4$ and $(0_{4\times4},\ F_{1,a}^b)$ For the entries of $d\Gamma^{(1)}$ we consider the following basis elements, $(F_{2,a}^b, 0_{4\times4})$ and the following element $d_{(D_2,D_1)}\hat{H}((F_{2,a}^b,0_{4\times4}))$ where,

$$\hat{H}((F_{2,a}^b, 0_{4\times4})) = j\circ H\circ(j^{-1})^{(2)}(\tilde{\iota}\circ P_j(F_{1,a}^b), P_j(0_{4\times4})) \tag{4.39}$$

We first look at the following derivatives

$$\begin{aligned} d_{(D_2,D_1)}\hat{H}((F_{2,4}^1, 0_{4\times4})) &= \left.\frac{d}{dt}\hat{H}(D_2+tF_{2,4}^1, D_1)\right|_{t=0} \\ \implies \left.\frac{d}{dt}\hat{H}(D_2+tF_{2,4}^1,D_1)\right|_{t=0} &= (\eta_{2,1}-\eta_{2,4})^{-1}\,d_{[e_4]}j(e_1) = (\eta_{2,1}-\eta_4)^{-1}\,e_1(P_j(\mu)) \end{aligned}$$

and

$$\begin{aligned} d_{(D_2,D_1)}\hat{H}((0_{4\times4}, F_{1,4}^1)) &= \left.\frac{d}{dt}\hat{H}(D_2, D_1+tF_{1,4}^1)\right|_{t=0} \\ &= (\eta_{1,4}-\eta_{1,1})^{-1}\,d_{[e_4]}j(\tilde{e}_1) = (\eta_{1,4}-\eta_{1,1})^{-1}\,e_a(\hat{H}(D_2,D_1)) \end{aligned} \tag{4.40}$$

$$d\Gamma^{(1)} = \begin{bmatrix} (\eta_{2,4}-\eta_{2,1})^{-1} & 0 & 0 \\ 0 & (\eta_{2,4}-\eta_{2,2})^{-1} & 0 \\ 0 & 0 & (\eta_{2,4}-\eta_{2,3})^{-1} \end{bmatrix}$$
$$d\Gamma^{(2)} = \begin{bmatrix} (\eta_{1,4}-\eta_{1,1})^{-1} & 0 & 0 \\ 0 & (\eta_{1,4}-\eta_{1,2})^{-1} & 0 \\ 0 & 0 & (\eta_{1,4}-\eta_{1,3})^{-1} \end{bmatrix}$$

$$\left[(d\Gamma^{(1)})\ G^{\iota}_{2,E}(j, X_{2,1})(d\Gamma^{(1)})^T\right]_{a,b} = n_2^{-1}(\eta_{2,4}-\eta_{2,a})^{-3}(\eta_{2,4}-\eta_{2,b})^{-3} \times \sum_{r=1}^{n_2} X^a_{2,r} X^b_{2,r} (X^4_{2,r})^2$$
$$\left[(d\Gamma^{(2)})\ G_{1,E}(j, X_{1,1})(d\Gamma^{(2)})^T\right]_{a,b} = n_1^{-1}(\eta_{1,4}-\eta_{1,a})^{-2}(\eta_{1,4}-\eta_{1,b})^{-2} \sum_{r=1}^{n_1} X^a_{1,r} X^b_{1,r} (X^4_{1,r})^2$$

$\square$

PROPOSITION 4.3.3. *For $a = 1, 2$, let $\{[X_{r_a}]\}_{r_a=1}^{n_a}$, $\|X_{r_a}\| = 1$, be independent random samples from j-nonfocal probability measures Q_a on $\mathbb{R}P^3$. Then the consistent estimator of Σ^{ι}_E is denoted $G^{\iota}_E(j, X_{1,1}, X_{2,1})$.*

(i)

$$n^{1/2}\ G^{\iota}_E(j, X_{1,1}, X_{2,1})^{-1/2} \tan_{j(\mu^{-1}_{2,E}\odot\mu_{1,E})} \left(j(\overline{X}^{-1}_{2,E} \odot \overline{X}_{1,E}) - j(\mu^{-1}_{2,E} \odot \mu_{1,E})\right) \to_d N_m(0_m, I_m) \tag{4.41}$$

so that

(ii)

$$n \left\| G^{\iota}_E(j, X_{1,1}, X_{2,1})^{-1/2} \tan_{j(\mu^{-1}_{2,E}\odot\mu_{1,E})} \left(j(\overline{X}^{-1}_{2,E} \odot \overline{X}_{1,E}) - j(\mu^{-1}_{2,E} \odot \mu_{1,E})\right) \right\|^2 \tag{4.42}$$

converges weakly to χ^2_m and the

CHAPTER 5

EXTRINSIC ANTI-MEAN

In this chapter Icontinue to focus on extrinsic analysis, which is the statistical analysis performed relative to ρ_j a chord distance on $\mathcal{M}$ induced by the Euclidean distance in $\mathbb{R}^N$ via an embedding $j : \mathcal{M} \to \mathbb{R}^N$, with an emphasis on compact object spaces. Most of the results in this section are due to the author of this dissertation, were presented at the second Conference of the International Society of Nonparametric Statistics, in Cadiz, Spain in 2015, and appeared in the peer reviewed publication [27]. Recall that the expected square distance from the random object X to an arbitrary point p defines what we call the Fréchet function associated with X and in extrinsic analysis it is given by;

$$\mathcal{F}(p) = \int_{\mathcal{M}} \|j(x) - j(p)\|_0^2 Q(dx), \tag{5.1}$$

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)[5]), 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, given $X_1, \ldots, X_n$ i.i.d random objects from Q, their *extrinsic sample mean (set)* is the extrinsic mean (set) of the empirical distribution $\hat{Q}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$. Recall that the existence of an extrinsic mean is tied to the existence of a unique projection of the mean μ of $j(Q)$ from the ambient space $\mathbb{R}^N$ onto the space $j(\mathcal{M}) \subset \mathbb{R}^N$. In the section 5.1 I introduce a new location parameter which is viewed as the (unique) maximizer of the Fréchet function given in (5.1) and is referred to as the *extrinsic anti-mean* (see Patrangenaru and Ellingson (2015)[21]) and I also express its corresponding *sample anti-mean* viewed as the maximizer of the Fréchet function associated with the empirical distribution $\hat{Q}_n$. In section 5.2 I give explicit formulas of the Veronesee-Whitney (VW) anti-mean on $\mathbb{R}P^m$. The following section involves inference problems for extrinsic means and anti-means on the 3-D projective shape space $(\mathbb{R}P^3)^q$. Section 5.4 using the results from the previous section, I perform a two sample test on a set of data consisting of digital images of flowers.

5.1 Geometric description of the extrinsic anti-mean

We 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 5.1.1. *The set of maximizers of the Fréchet function, is called the* extrinsic anti-mean set. *In case the extrinsic anti-mean set has one point only, that point is called* **extrinsic anti-mean** *of X, and is labeled $\alpha\mu_{j,E}(Q)$, or simply $\alpha\mu_E$, when j is known.*

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

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

where ρ_0 is the Euclidean distance in $\mathbb{R}^N$.

REMARK 5.1.1. *Recall that a point $y \in \mathbb{R}^N$ for which there is a unique point $p \in \mathcal{M}$ satisfying the equality,*

$$\rho_0(y, j(\mathcal{M})) = \inf_{x \in \mathcal{M}} \|y - j(x)\|_0 = \rho_0(y, j(p))$$

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 $\rho_0(y, j(\mathcal{M})) = \rho_0(y, j(p))$.

With this in mind we now have the following definition.

DEFINITION 5.1.2 (αj-nonfocal)**.** *(a) A point $y \in \mathbb{R}^N$ for which there is a unique point $p \in \mathcal{M}$ satisfying the equality,*

$$\sup_{x \in \mathcal{M}} \|y - j(x)\|_0 = \rho_0(y, j(p)) \tag{5.2}$$

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 farthest projection on $j(\mathcal{M})$ is the unique point $z = j(p) = P_{F,j}(y) \in j(\mathcal{M})$ with
$\sup_{x \in \mathcal{M}} \|y - j(x)\|_0 = \rho_0(y, j(p))$.

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 5.1.3. *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 anti-mean of distributions on a one dimensional topological manifold $\mathcal{M}$ where the distributions are j-nonfocal and also αj-nonfocal. Note that in the smooth case, given a family of distributions, for which the mean vector in the ambient space, slightly moves in a direction perpendicular on the tangent space $j(\mu_E)$, the extrinsic mean stays the same, while the extrinsic anti-mean may change; this shows that the extrinsic anti-mean is a new location parameter, that detects certain global aspects of a distribution, that are not captured by the extrinsic mean. On the second line of Figure 5.1, one displays the stickiness phenomenon in case of both the extrinsic mean and anti-mean. Recall that a sticky family of distributions is a family of distributions for which any small perturbation does not affect the location of the Fréchet mean; this phenomenon may occurs in case the Fréchet mean happens to be a singular point in both extrinsic analysis (see [9]) and intrinsic analysis (see [13]).

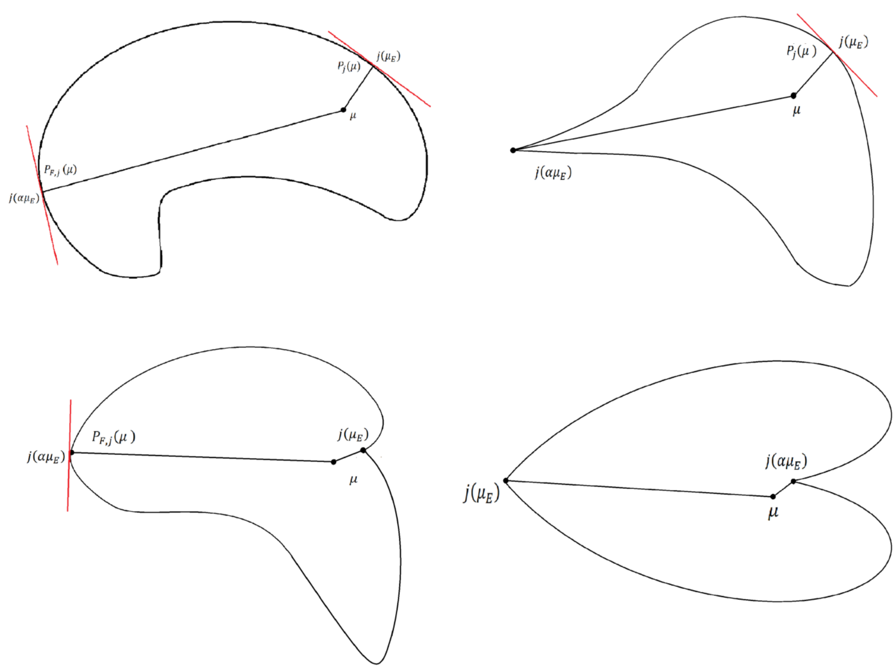


Figure 5.1: Extrinsic mean and extrinsic anti-mean on a 1-dimensional topological manifold (upper left: regular mean and anti-mean, upper right: regular mean and sticky anti-mean, lower left: sticky mean and regular anti-mean, lower right : sticky mean and anti-mean

THEOREM 5.1.1. *Let μ be the mean vector of $j(Q)$ in $\mathbb{R}^N$. Then the following hold true:*

(i) The extrinsic anti-mean set is the set of all points $x \in \mathcal{M}$ such that $\sup_{p\in\mathcal{M}} \|\mu - j(p)\|_0 = \rho_0(\mu, j(x))$.

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

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

$$\|j(p) - j(x)\|_0^2 = \|j(p) - \mu\|_0^2 - 2\,\langle\, j(p) - \mu,\ \mu - j(x)\,\rangle + \|\mu - j(x)\|_0^2 \tag{5.3}$$

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

$$\int_{\mathcal{M}} j(x)\, Q(dx) = \int_{\mathbb{R}^N} y j(Q)(dy) = \mu. \tag{5.4}$$

From equations (5.4), (5.3) it follows that

$$\mathcal{F}(p) = \|j(p) - \mu\|_0^2 + \int_{\mathbb{R}^N} \|\mu - y\|_0^2 j(Q)(dy) \tag{5.5}$$

Then, from (5.5),

$$\sup_{p\in\mathcal{M}} \mathcal{F}(p) = \sup_{p\in\mathcal{M}} \|j(p) - \mu\|_0^2 + \int_{\mathbb{R}^N} \|\mu - y\|_0^2\, j(Q)(dy) \tag{5.6}$$

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

$$\sup_{p\in\mathcal{M}} \|j(p) - \mu\|_0 = \|j(x) - \mu\|_0. \tag{5.7}$$

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 (5.7) holds true, which implies that μ is αj-nonfocal and $j(\alpha\mu_{j,E}(Q)) = P_{F,j}(\mu)$. □

DEFINITION 5.1.4. *Let $x_1, ..., x_n$ be random observations from a distribution Q on a compact metric space $(\mathcal{M},\ \rho)$, then their extrinsic sample anti-mean 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*

$$\hat{\mathcal{F}}_n(p) = \frac{1}{n}\sum_{i=1}^n \|j(x_i) - j(p)\|_0^2 \tag{5.8}$$

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

THEOREM 5.1.2. *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 anti-mean 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 anti-mean $a\bar{X}_{j,E}$ 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 5.1.1, 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 $\mathbb{P}(\|\overline{j(X)} - \mu\| > \delta) \leq \varepsilon, \forall n > n(\delta, \varepsilon)$. By taking a small enough $\delta > 0$, and using a continuity argument for $P_{F,j}$, the result follows. $\square$

REMARK 5.1.2. *A CLT for extrinsic sample anti-means is given in a paper I have coauthored (see Patrangenaru et. al.(2016)[22]).*

5.2 VW anti-means 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 $(\mathbb{R}P^m, \rho_j)$, where the distance ρ_j 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 anti-mean 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

$$j(T \cdot [x]) = T \circ j([x]), \ \ \forall\, T \in SO(m+1), \forall\, [x] \in \mathbb{R}P^m, \ \ where\, T \cdot [x] = [Tx]. \tag{5.9}$$

Such an embedding is said to be *equivariant* (see Kent (1992)[17], where the equivariance was used in the context of a VW embedding of a planar direct similarity shape space). For computational purposes, the equivariant embedding of $\mathbb{R}P^m$ that was used so far in the axial data analysis literature is the 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)[21] and references therein):

$$j([x]) = xx^T, \|x\| = 1, \tag{5.10}$$

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

$$T \circ A = TAT^T, \;\; \forall\, T \in SO(m+1), \forall\, A \in S_+(m+1,\mathbb{R}) \tag{5.11}$$

REMARK 5.2.1. *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 ρ_0 are given by respectively by $\|C\|_0^2 = \langle C,\ C\rangle_0$ and $\rho_0(A,B) = \|A-B\|_0,\ \forall C, A,\ B\ \in\ S(m+1,\mathbb{R})$.*

With the notation in Remark 5.2.1 we have

$$\mathcal{F}([p]) = \|j([p]) - \mu\|_0^2 + \int_{\mathcal{M}} \|\mu - j([x])\|_0^2\, Q(d[x]), \tag{5.12}$$

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)[21] 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 mm^T, where m is a unit eigenvector of A corresponding to its largest eigenvalue.

Furthermore, the VW mean of a random object $[X] \in \mathbb{R}P^m$, $\|X^TX\| = 1$ 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 (5.12), we get similar results in the case of an αj-nonfocal probability measure Q :

PROPOSITION 5.2.1. *(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 5.2.2. *Let Q be a distribution on $\mathbb{R}P^m$.*

(a) The VW-antimean set of a random object $[X], X^TX = 1$ 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 $[x_1], \dots, [x_n]$ be observations from a distribution Q on $\mathbb{R}P^m$, such that $\overline{j(X)}$ is αVW-nonfocal. Then the VW sample anti-mean of $[x_1], \dots, [x_n]$ 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 $J = \dfrac{1}{n}\sum_{i=1}^{n} x_i x_i^T$, where $x_i^T x_i = 1, \forall i = \overline{1,n}$.

5.3 Two-sample test for VW means and anti-means 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 $\pi = ([u_1], \dots, [u_5])$ 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)[23]). Recall from Section 2.5 that $\mathbb{R}P^3$ has a natural structure of Lie group. This multiplicative structure turns the $(\mathbb{R}P^3)^q$ into a product Lie group $(\mathcal{G}, \odot)$ where $\mathcal{G} = (\mathbb{R}P^3)^q$ (see Crane and Patrangenaru (2011)[7], Patrangenaru et. al. (2014)[25]). 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)[25]), is given by

$$j_q([x_1], \dots, [x_q]) = (j([x_1]), \dots, j([x_q])), \tag{5.13}$$

with $j : \mathbb{R}P^3 \to S_+(4, \mathbb{R})$ the VW embedding given in (6.19), 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 $Y = (Y^1, \dots, Y^q)$ on $(\mathbb{R}P^3)^q$ is given by

$$\mu_{j_q} = (\mu_{1,j}, \cdots, \mu_{q,j}), \tag{5.14}$$

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 r.o.'s 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-anti-means for these random objects. Note that the asymptotic results leading to nonparametric bootstrap confidence regions for VW-mean change are presented in Section 2.5. For VW anti-means we will simply use nonpivotal bootsrap computations, since for the sample VW-antimeans on $(\mathbb{R}P^3)^q$ for our data, involve sample covariance matrices that are degenerate.

5.3.1 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:

$$H_0 :\ \mu_{1,j_q} = \mu_{2,j_q} \text{ vs. } H_a :\ \mu_{1,j_q} \neq \mu_{2,j_q}, \tag{5.15}$$

which is equivalent to testing the following

$$H_0 :\ \mu_{2,j_q}^{-1} \odot \mu_{1,j_q} = 1_{(\mathbb{R}P^3)^q} \text{ vs. } H_a :\ \mu_{2,j_q}^{-1} \odot \mu_{1,j_q} \neq 1_{(\mathbb{R}P^3)^q} \tag{5.16}$$

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_q be the affine chart defined in a neighborhood of $1_{(\mathbb{R}P^3)^q}$ (see definition 3.1.1), with $\varphi_q(1_{(\mathbb{R}P^3)^q}) = 0$. Then, under H_0

$$n^{1/2}\ \varphi_q(\bar{Y}_{j_q,n_2}^{-1} \odot \bar{Y}_{j_q,n_1}) \to_d \mathcal{N}_{3q}(0_{3q}, \Sigma_{j_q}) \tag{5.17}$$

   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_q and $\varphi_q(Y_{a,1})$ has an absolutely continuous component and finite moment of sufficiently high order. Then the joint distribution

$$V = n^{\frac{1}{2}}\varphi_q(\bar{Y}_{j_q,n_2}^{-1} \odot \bar{Y}_{j_q,n_1}) \tag{5.18}$$

   can be approximated by the bootstrap joint distribution of

$$V^* = n^{1/2}\,\varphi_q(\bar{Y^*}^{-1}_{j_q,n_2} \odot \bar{Y}^*_{j_q,n_1})$$

From Patrangenaru et. al.(2010)[23], 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

$$\bar{Y}_{j_q,n} = ([g_1(m+1)], \ldots, [g_q(m+1)]). \tag{5.19}$$

REMARK 5.3.1. *Given the high dimensionality, the VW sample covariance matrix is often singular. Therefore, for nonparametric hypothesis testing, non-pivotal bootstrap is preferred. For details, on testing the existence of a mean change 3D projective shape, when sample sizes are not equal, using non-pivotal bootstrap, see Patrangenaru et al. (2014).*

5.3.2 Hypothesis testing for VW anti-means

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:

$$H_0 :\ \alpha\mu_{1,j_q} = \alpha\mu_{2,j_q} \text{ vs. } H_a :\ \alpha\mu_{1,j_q} \neq \alpha\mu_{2,j_q}, \tag{5.20}$$

which is equivalent to testing the following

$$H_0 :\ \alpha\mu_{2,j_q}^{-1} \odot \alpha\mu_{1,j_q} = 1_{(\mathbb{R}P^3)^q} \text{ vs. } H_a :\ \alpha\mu_{2,j_q}^{-1} \odot \alpha\mu_{1,j_q} \neq 1_{(\mathbb{R}P^3)^q} \tag{5.21}$$

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_q be the affine chart with $\varphi_q(1_{(\mathbb{R}P^3)^q}) = 0_{3q}$. Then, from Patrangenaru et al. (2016)[26], it follows that under H_0

$$n^{1/2}\,\varphi_q(a\bar{Y}^{-1}_{j_q,n_2} \odot a\bar{Y}_{j_q,n_1}) \to_d \mathcal{N}_{3q}(0_{3q}, \tilde{\Sigma}_{j_q}), \tag{5.22}$$

   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 anti-mean $\alpha\mu_{a,j_q}$ 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

$$aV = n^{\frac{1}{2}}\varphi_q(a\bar{Y}^{-1}_{j_q,n_2} \odot a\bar{Y}_{j_q,n_1}) \tag{5.23}$$

   can be approximated by the bootstrap joint distribution of

$$aV^* = n^{1/2}\,\varphi_q(a\bar{Y^*}^{-1}_{j_q,n_2} \odot a\bar{Y}^*_{j_q,n_1})$$

Now, from Proposition 5.2.2, we get the following result that is used for the computation of the VW sample anti-means.

PROPOSITION 5.3.1. *follows that given a random sample from a distribution Q on $\mathbb{R}P^m$, if $J_s, s = 1,\dots,q$ are the matrices $J_s = n^{-1}\sum_{r=1}^{n} X_r^s (X_r^s)^T$, and if for $a = 1,\dots,m+1, d_s(a)$ and $g_s(a)$ are the eigenvalues in increasing order and corresponding unit eigenvectors of J_s, then the VW sample anti-mean $a\bar{Y}_{j_q,n}$ is given by*

$$a\bar{Y}_{j_q,n} = ([g_1(1)],\dots,[g_q(1)]). \tag{5.24}$$

5.4 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 5.2: Flower 1 image sample

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

Figure 5.3: Flower 2 image sample

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 5.4: Landmarks for flower 1 and flower 2

For 3D reconstructions of k-ads we used the reconstruction algorithm in Ma et al (2005)[19]. 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)[7]). 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

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

Given φ, the Log chart on this Lie group, $\varphi_q(1_8) = 0_8$, we compute the bootstrap distribution

$$D_* = \varphi_q((\bar{Y}^*_{j_8,11})^{-1} \odot \bar{Y}^*_{j_8,8})$$

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.

5.4.1 Results for comparing the two flowers

We will fail to reject our null hypothesis

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

if all of our confidence intervals contain the value 0.

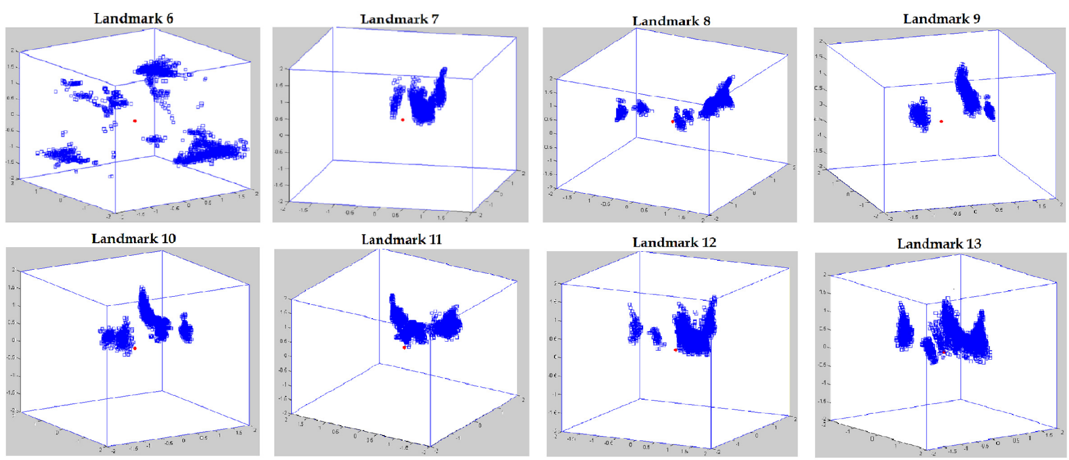


Figure 5.5: Bootstrap projective shape marginals for lily data

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 5.5. The bootstrap projective shape marginals for landmarks 11 and 12 we can also visually reinforce our choice of rejection of the null hypothesis.

5.4.2 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 $\alpha = 0.05$.

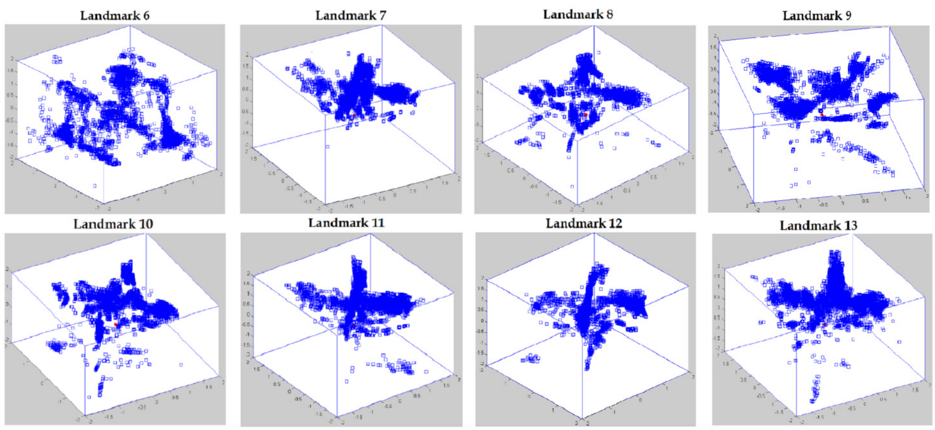


Figure 5.6: Bootstrap projective shape marginals for cross validation of lily flower

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	$(-1.092425, 1.795890)$	$(-1.222856, 1.963960)$	$(-1.128044, 1.762559)$	$(-1.035796, 2.227439)$

5.4.3 Comparing the sample anti-mean for the two lily flowers

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

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

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

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

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.

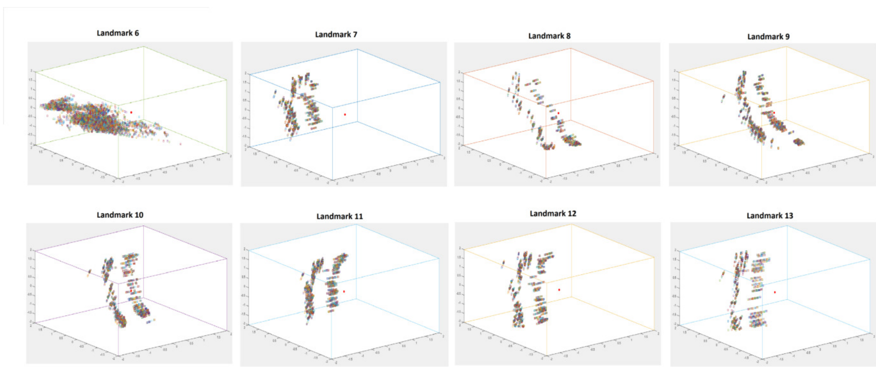


Figure 5.7: Eight bootstrap projective shape marginals for anti-mean of lily data

Highlighted in blue are the intervals not containing $0 \in \mathbb{R}$.
In conclusion there is significant anti-mean VW projective shape change between the two flowers, showing that the extrinsic anti-mean is a sensitive parameter for extrinsic analysis.

In this chapter we introduced a new population parameter, the extrinsic anti-mean. This new location parameter is based on a projection unlike the one in the extrinsic mean case, where we focus on projecting μ (the mean of $j(Q)$ in the ambient space) onto the closest (unique) point $j(\mu_E)$ on $j(\mathcal{M})$; we will instead project μ onto the farthest (unique) point $(j(\alpha\mu_E)$ on the embedded object space . Just as with the extrinsic mean,

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	$(-3.03, 1.91)$	$(-5.42, 1.98)$	$(-7.22, 2.41)$	$(-4.91, 2.62)$

the extrinsic anti-mean captures important features of a distribution on a compact object space. Certainly the definitions and results extend to the general case of arbitrary Fréchet anti-means.

CHAPTER 6

MANOVA ON MANIFOLDS

In this chapter I revisit MANOVA for comparing the mean vectors in g populations. I am extending such considerations to testing for the equality of extrinsic means from g populations on a manifold $\mathcal{M}$ embedded in an numerical space. In section 6.1 I introduce a new approach applied to various mean vectors. The main difference between this approach and classical MANOVA, is that we do not assume that all populations have a common covariance matrix Σ, and also we do not make any distributional assumption, except for the existence of sufficiently high order moments of the g populations. In section 6.2 I extend the work presented in the previous section to develop a hypothesis testing problem used to compare multiple means on smooth manifolds, and this test is performed on random samples of various sizes, collected from each of these g groups. This newly developed MANOVA test is then applied in section 6.3 to populations of 3D projective shapes.

6.1 Motivations for new MANOVA on manifolds

For $a = 1, ..., g$, suppose $X_{a,i} \sim N_p(\mu_a, \Sigma_a), i = 1, ..., n_a$ are p dimensional i.i.d random vectors. To test if the mean vectors of the g groups are the same, one considers the hypothesis testing problem

$$\begin{aligned} &H_0 : \ \mu_1 = \mu_2 = ... = \mu_g = \mu \qquad (6.1)\\ &H_a : \ at\ least\ one\ equation\ does\ not\ hold. \end{aligned}$$

Assuming that the covariance matrix Σ_a is invertible, by the Central Limit Theorem, for large sample sizes $n_a, a = 1, \ldots, g$, we have

$$\sqrt{n_a}\Sigma_a^{-\frac{1}{2}}(\bar{X}_a - \mu) \sim N_p(0_p, I_p), \qquad (6.2)$$

$$n_a(\bar{X}_a - \mu)^T \Sigma_a^{-1}(\bar{X}_a - \mu) \sim \chi^2_p. \qquad (6.3)$$

However, Σ_a is always unknown, so in practice, one has to use its unbiased estimator $S_a,\ a = 1, ..., g$.

$$n_a(\bar{X}_a - \mu)^T S_a^{-1}(\bar{X}_a - \mu) \sim \chi^2_p. \qquad (6.4)$$

Let us consider the pooled sample mean $\bar{X} = \frac{1}{n}(n_1\bar{X}_1 + ... + n_g\bar{X}_g),\ n = \sum_{a=1}^{g} n_a$.

LEMMA 6.1.1. *Under the null, $\bar{X}$ is a consistent estimator of μ, provided $\frac{n_a}{n} \to \lambda_a > 0,\ as\, n \to \infty,\ a = 1, ..., g$.*

Proof. Indeed, for any $a \in \{1, 2, ..., g\}$, since $\frac{n_a}{n} \to \lambda_a > 0,\ as\, n \to \infty$, and $\bar{X}_a$ is the consistent estimator of μ, therefore,

$$\bar{X} \to_p \lambda_1\mu + \lambda_2\mu + ... + \lambda_g\mu = \mu. \tag{6.5}$$

□

THEOREM 6.1.1. *The statistic for the hypothesis in* (6.1) *is*

$$\sum_{a=1}^{g} n_a(\bar{X}_a - \bar{X})^T S_a^{-1}(\bar{X}_a - \bar{X}) \sim \chi^2_{gp}. \tag{6.6}$$

So the rejection region at level c, for this test is

$$\sum_{a=1}^{g} n_a(\bar{X}_a - \bar{X})^T S_a^{-1}(\bar{X}_a - \bar{X}) > \chi^2_{gp}(c). \tag{6.7}$$

6.2 MANOVA on manifolds

In this section we will focus on the asymptotic behavior of statistics related to means on a manifold $\mathcal{M}$ based on samples of different sizes from different populations on $\mathcal{M}$. Now let's consider the set $X_{a,1}, \ldots, X_{a,n_a}$ ($a = 1, 2, ..., g$) of iid random objects on $\mathcal{M}$ with common probability measure Q_a. We denote the extrinsic mean of the j- nonfocal probability measure Q_a on $\mathcal{M}$ by $\mu_{a,E}$ for ease of notation and because there is no ambiguity about the embedding used. The corresponding extrinsic sample means are written $\bar{X}_{a,E}$ for $a = 1, \cdots, g$. From this point on, we will assume that all the distributions are j-nonfocal.

6.2.1 Hypothesis testing and T^2 statistic

Assume $X_{a,1}, \ldots, X_{a,n_a}$ are iid random objects on $\mathcal{M}$ a p-dimensional manifold, with probability measure Q_a with $a = 1, 2, ..., g$. We are interested in comparing multiple extrinsic means.

We would like to develop a test similar to (6.1) designed to test the difference between the g extrinsic means. One challenge that presents itself at the early stage is a proper definition of a pooled mean for random objects on a p-dimensional manifold $\mathcal{M}$. Linearity becomes an issue when dealing with extrinsic means. For a proper definition we will focus on the equalities tied to the assumption

$$A_0 : \mu_{1,E} = \cdots = \mu_{g,E}$$

DEFINITION 6.2.1. *Under the assumption A_0 and for any $a \in \{1, 2, ..., g\}$, with $\frac{n_a}{n} \to \lambda_a > 0,\ as\ n \to \infty$. We define*

(i) The **extrinsic pooled mean with weights** $\lambda = (\lambda_1, \dots, \lambda_g)$**,** *denoted $\mu_E(\lambda)$ as the value in $\mathcal{M}$ given by*

$$j(\mu_E) = P_j(\lambda_1 j(\mu_{1,E}) + \cdots + \lambda_g j(\mu_{g,E})) \tag{6.8}$$

Where $\mu_{a,E}$ is the extrinsic mean of the random object $X_{a,1}$ and $\Sigma_{a=1}^{g} \lambda_a = 1$

(ii) The **extrinsic pooled sample mean** *denoted $\bar{X}_E \in \mathcal{M}$ given by;*

$$j(\bar{X}_E) = P_j\left(\frac{n_1}{n} j(\bar{X}_{1,E}) + \cdots + \frac{n_g}{n} j(\bar{X}_{g,E})\right) \tag{6.9}$$

Where $\bar{X}_{a,E}$ is the extrinsic sample mean for $X_{a,1}$ and $n = \sum_{a=1}^{g} n_a$

Note that since A_0 implies $j(\mu_{1,E}) = \cdots = j(\mu_{g,E})$, and with our definition of the extrinsic pooled mean we get $j(\mu_E) = j(\mu_{a,E})$ for each $a = 1, \dots, g$. Furthermore, the linear combination $\lambda_1 j(\mu_{1,E}) + \cdots + \lambda_g j(\mu_{g,E}) \in j(\mathcal{M})$. Note that for $a = 1, \cdots, g$ $\bar{X}_{a,E}$ is a consistent estimator of $\mu_{a,E}$ and therefore we get that $j(\bar{X}_E) \to_p j(\mu_E)$. Since j is a homeomorphism from $\mathcal{M}$ to $j(\mathcal{M})$ we also have that $\bar{X}_E$ is a consistent estimator of μ_E the extrinsic pooled mean. With this definition at hand, we now express the following hypothesis test, designed to test the difference between extrinsic means and is given by;

$$\begin{aligned} &H_0: \ \mu_{1,E} = \mu_{2,E} = ... = \mu_{g,E} = \mu_E, \\ &H_a: \ \textit{at least one equality } \mu_{a,E} = \mu_{b,E}, 1 \leq a < b \leq g \textit{ does not hold.} \end{aligned} \tag{6.10}$$

And since the embedding $j : \mathcal{M} \to \mathbb{R}^N$ is one-to-one the hypothesis above can be interchangeably written

$$\begin{aligned} &H_0^j: \ j(\mu_{1,E}) = j(\mu_{2,E}) = ... = j(\mu_{g,E}) = j(\mu_E), \\ &H_a^j: \ \textit{at least one equality } \mu_{a,E} = \mu_{b,E}, 1 \leq a < b \leq g \textit{ does not hold.} \end{aligned} \tag{6.11}$$

In order to test hypothesis (6.10) we will use a T^2 like statistic. The theorem below, gives us the asymptotic behavior needed to establish such a statistic. For $a = 1, \dots, g$, we get, from Bhattacharya and Patrangenaru [6], the following:

(i) $S_{n_a} = (n_a)^{-1} \Sigma_{i=1}^{n_a} (j(X_{a,i}) - j(\bar{X}_E))(j(X_{a,i}) - j(\bar{X}_E))^T$ is a consistent estimator of Σ_a, and

(ii) $\tan_{j(\bar{X}_E)} \nu$ is a consistent estimator of $\tan_{P_j(\mu)} \nu$, where $\nu \in \mathbb{R}^N$.

It follows that $G_{\bar{X}}(j, X_a)$, given by

$$G_{\bar{X}}(j,X_a) = \left[\left[\sum_{a=1}^{m} d_{\overline{j^{(p)}(X)}} P_j(e_b) \cdot e_i(j(\bar{X}_E))\, e_i(j(\bar{X}_E))\right]_{i=1,\ldots,p}\right] \cdot\ S_{n_a}$$
$$\left[\left[\sum_{a=1}^{m} d_{\overline{j^{(p)}(X)}} P_j(e_b) \cdot e_i(j(\bar{X}_E)) e_i(j(\bar{X}_E))\right]_{i=1,\ldots,p}\right]^T$$

where for $\overline{j^{(p)}(X)} = \frac{n_1}{n} j(\bar{X}_{1,E}) + \cdots + \frac{n_g}{n} j(\bar{X}_{g,E})$ and is a consistent estimator of μ such that $P_j(\mu) = j(\mu_E)$. One must note that the extrinsic sample covariance matrix $G(j, X_a)$ is expressed in terms of $d_{\overline{j^{(p)}(X)}} P_j(e_b) \in T_{j(\bar{X}_E)} j(\mathcal{M})$ and not in term of $d_{\overline{j(Xa,1)}} P_j(e_b) \in T_{j(\bar{X}_{a,E})} j(\mathcal{M})$.

THEOREM 6.2.1. *Assume $j : \mathcal{M} \to \mathbb{R}^N$ is a closed embedding of $\mathcal{M}$. Let $\{X_{a,i}\}_{i=1}^{n_a}$ for $a = 1, ..., g$ be random samples from the j-nonfocal distributions $\mathcal{Q}_a$. Let $\mu_a = E(j(X_{a,1}))$ and assume $j(X_{a,1})$'s have finite second-order moments and the extrinsic covariance matrices $\Sigma_{a,E}$ of $X_{a,1}$ are nonsingular. We also let $(e_1(p), ..., e_N(p))$, for $p \in \mathcal{M}$ be an orthonormal frame field adapted to j.*
Furthermore, let $\frac{n_a}{n} \to \lambda_a > 0$, as $n \to \infty$, with $n = \Sigma_{a=1}^{g} n_a$, and $\Sigma_{a=1}^{g} \lambda_a = 1$.Then we have the following asymptotic behavior;

$$\sum_{a=1}^{g} n_a\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E))^T \Sigma_{a,E}^{-1}\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E)) \to_d \chi^2_{gp}.$$

It follows that the statistics for hypothesis (6.10) *have the following behaviors;*

(a) the statistic

$$\sum_{a=1}^{g} n_a\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E))^T G_{\bar{X}}(j, X_a)^{-1}\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E)) \to_d \chi^2_{gp}.$$

(b) the statistic

$$\sum_{a=1}^{g} n_a\ \tan_{j(\bar{X}_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E))^T G_{\bar{X}}(j, X_a)^{-1}\ \tan_{j(\bar{X}_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E)) \to_d \chi^2_{gp}.$$

Proof. recall that from Bhattacharya and Patrangenaru (2005) [6] we have

$$\sqrt{n_a}\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E)) \to_d N(0_p, \Sigma_{a,E}), \quad for\ a = 1, 2, ..., g$$

where

$$\Sigma_{a,E} = \left[\left[\sum d_{\mu} P_j(e_b) \cdot e_k(P_j(\mu))\right]_{k=1,\ldots,p}\right]\ \Sigma_a\ \left[\left[\sum d_{\mu} P_j(e_b) \cdot e_k(P_j(\mu))\right]^T_{k=1,\ldots,p}\right]$$

where $\mu = \lambda_1 j(\mu_{1,E}) + \cdots + \lambda_g j(\mu_{g,E})$ and the Σ_a's are the covariance matrices of the $j(X_{a,1})$'s with respect to the canonical basis $e_1, ..., e_N$. And under the null, from 6.10, the matrices $\Sigma_{a,E}$ are defined with respect to the basis $f_1(\mu_E), ..., f_p(\mu_E)$ of local frame fields. We then have for each $a = 1, ..., g$

$$n_a \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E))^T \Sigma_{a,E}^{-1} \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E)) \to_d \chi^2_p.$$

and since the random samples are independent we have,

$$\sum_{a=1}^{g} n_a \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E))^T \Sigma_{a,E}^{-1}\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\mu_E)) \to_d \chi^2_{gp}. \tag{6.12}$$

$\bar{X}_E$ is the consistent estimator of μ_E, then the pooled sample mean

$$j(\bar{X}_E) = P_j\left(\frac{1}{n}\sum_{a=1}^{g} n_a j(\bar{X}_{a,E})\right) \to_p j(\mu_E) \quad \text{(by lemma 6.1.1)} \tag{6.13}$$

And since $G_{\bar{X}}(j, X_a)$ consistently estimate Σ_a and $\tan_{j(\bar{X}_E)}$ is a consistent estimator of $\tan_{j(\mu_E)}$, we have the following

$$\sum_{a=1}^{g} n_a \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E))^T G_{\bar{X}}(j, X_a)^{-1}\ \tan_{j(\mu_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E)) \to_d \chi^2_{gp}.$$
$$\sum_{a=1}^{g} n_a\ \tan_{j(\bar{X}_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E))^T G_{\bar{X}}(j, X_a)^{-1}\ \tan_{j(\bar{X}_E)}(j(\bar{X}_{a,E}) - j(\bar{X}_E)) \to_d \chi^2_{gp}.$$

□

6.2.2 Nonparametric bootstrap confidence regions

From Corollary 3.2 in [6] under the hypothesis

$$\begin{cases} H_0 & : \mu_{1,E} = \mu_{2,E} = ... = \mu_{g,E} = \mu_E, \\ H_a & :\ni (i,j) 1 \leq i < j < g, \text{s.t. } \mu_{i,E} \neq \mu_{j,E}, \end{cases}$$

we have:

COROLLARY 6.2.1. *Under the assumptions of Theorem* (6.2.1) *, a confidence regions for μ_E of asymptotic level $1-c$ is given by $C^{(g)}_{n,c}$ and $D^{(g)}_{n,c}$ which are defined below*

(a) $C^{(g)}_{n,c} = j^{-1}(U_{n,c})$ *where*
$U_{n,c} = \{j(\nu) \in j(\mathcal{M}) :\ n\ \sum_{a=1}^{g} n_a \left\|G_{\bar{X}}(j, X_a)^{-1/2}\ \tan_{j(\nu)}(j(\overline{X}_{a,E}) - j(\nu))\right\|^2 \leq \chi^2_{gp,1-c}\}$

(b) $D^{(g)}_{n,c} = j^{-1}(V_{n,c})$ *where*

$V_{n,c} = \{j(\nu) \in j(\mathcal{M}) : \sum_{a=1}^{g} n_a \left\| G_{\bar{X}}(j, X_a)^{-1/2}\ \tan_{j(\bar{X}_E)}(j(\overline{X}_{a,E}) - j(\nu))\right\|^2 \leq \chi^2_{gp,1-c}\}$

where $\bar{X}_E$ *is the extrinsic pooled sample mean defined in Definition 6.2.1* (ii)

Most of the data we will be focusing on will have value of n relatively small. We will need to use re sampling, in particular bootstrap methods. For $a = 1, ..., g$, let $\{X_{a,i}\}_{i=1}^{n_a}$ be i.i.d.r.o's from the j-nonfocal distributions $\mathcal{Q}_a$. Let $\{X^*_{a,r}\}_{r=1,...,n_a}$ be random re samples with repetition from the empirical $\hat{Q}_{n_a}$ conditionally given $\{X_{a,i}\}_{i=1}^{n_a}$. The confidence regions $C^{(g)}_{n,c}$ and $D^{(g)}_{n,c}$ described above have the corresponding bootstrap analogue $C^{*(g)}_{n,c}$ and $D^{*(g)}_{n,c}$ which are defined in the corollary below.

COROLLARY 6.2.2. *The* $(1-c)100\%$ *bootstrap confidence regions for* μ_E *with* $d = gp$ *are given by*

(a) $C^{*(g)}_{n,c} = j^{-1}(U^*_{n,c})$ *and*

$$U^*_{n,c} = \{j(\nu) \in j(\mathcal{M}) : \sum_{a=1}^{g} n_a \left\| G_{\bar{X}}(j, X_a)^{-1/2}\ \tan_{j(\nu)}(j(\overline{X}_{a,E}) - j(\nu))\right\|^2 \leq c^{*(g)}_{1-c}\} \qquad (6.14)$$

where $c^{*(g)}_{1-c}$ *is the upper* $100(1-c)\%$ *point of the values*

$$\sum_{a=1}^{g} n_a \left\| G_{\bar{X}^*}(j, {X^*}_a)^{-1/2}\ \tan_{j(\bar{X}_E)}(j(\overline{X^*}_{a,E}) - j(\bar{X}_E))\right\|^2 \qquad (6.15)$$

among the bootstrap re samples.

(b) $D^{*(g)}_{n,c} = j^{-1}({V^*}_{n,c})$ *and*

$${V^*}_{n,c} = \{j(\nu) \in j(\mathcal{M}) : \sum_{a=1}^{g} n_a \left\| G_{\bar{X}}(j, X_a)^{-1/2}\ \tan_{j(\bar{X}_E)}(j(\overline{X}_{a,E}) - j(\nu))\right\|^2 \leq d^{*(g)}_{1-c}\} \qquad (6.16)$$

where $d^{*(g)}_{1-c}$ *is the upper* $100(1-c)\%$ *point of the values*

$$\sum_{a=1}^{g} n_a \left\| G_{\bar{X}^*}(j, X^*_a)^{-1/2}\ \tan_{j(\bar{X}^*_E)}(j(\overline{X^*}_{a,E}) - j(\bar{X}_E))\right\|^2 \qquad (6.17)$$

where $\bar{X}^*_E$ *is the extrinsic pooled re sampled mean given by*

$$j(\bar{X}^*_E) = P_j\left(\frac{n_1}{n} j(\bar{X}^*_{1,E}) + \cdots + \frac{n_g}{n} j(\bar{X}^*_{g,E})\right) \qquad (6.18)$$

among the bootstrap re samples. Both of the regions given by (6.16) *and* (6.14) *have coverage erro* $O_p(n^{-2})$.

Note that $G_{\bar{X}^*}(j, X_a^*)$

$$G_{\bar{X}^*}(j, X_a^*) = \left[\left[\sum_{a=1}^{m} d_{\overline{j^{(p)}(X^*)}} P_j(e_b) \cdot e_i(j(\bar{X}_E^*))\, e_i(j(\bar{X}_E^*))\right]_{i=1,\dots,p}\right] \cdot\ S_{n_a}^* \left[\left[\sum_{a=1}^{m} d_{\overline{j^{(p)}(X^*)}} P_j(e_b) \cdot e_i(j(\bar{X}_E^*)) e_i(j(\bar{X}_E^*))\right]_{i=1,\dots,p}\right]^T$$

where $S_{n_a}^* = (n_a)^{-1}\Sigma_{i=1}^{n_a}(j(X_{a,i}^*) - j(\bar{X}_E^*))(j(X_{a,i}^*) - j(\bar{X}_E^*))^T$.

We now express the following test statistics that will be used in our analysis and are tied to the confidence regions mentioned above.

PROPOSITION 6.2.1. *Let $\{X_{a,i}\}_{i=1}^{n_a}$ for $a = 1, ..., g$ be random samples from the j-nonfocal distributions $\mathcal{Q}_a$. Let $\mu_a = E(j(X_{a,1}))$ and assume $j(X_{a,1})$'s have finite second-order moments and the extrinsic covariance matrices $\Sigma_{a,E}$ of $X_{a,1}$ are nonsingular.*

(a) Then the distribution of $T_c(X^{(g)}, \hat{Q}^{(g)}) = \sum_{a=1}^{g} n_a \left\| G_\mu(j, X_a)^{-1/2}\ \tan_{j(\mu_E)}(j(\overline{X}_{a,E}) - j(\mu_E)) \right\|^2$ can be approximated by the bootstrap distribution function of
$T_c(X^{*(g)}, \hat{Q}^{(g)}) = \sum_{a=1}^{g} n_a \left\| G_{\bar{X}}(j, X_a^*)^{-1/2}\ \tan_{j(\bar{X}_E)}(j(\overline{X}_{a,E}^*) - j(\bar{X}_E)) \right\|^2$

(b) Similarly, the distribution of $T_d(X^{(g)}, \hat{Q}^{(g)}) = \sum_{a=1}^{g} n_a \left\| G(j, X_a)^{-1/2}\ \tan_{j(\bar{X}_E)}(j(\overline{X}_{a,E}) - j(\mu_E)) \right\|^2$ can be approximated by the bootstrap distribution function of
$T_d(X^{*(g)}, \hat{Q}^{*(g)}) = \sum_{a=1}^{g} n_a \left\| G_{\bar{X}^*}(j, X_a^*)^{-1/2}\ \tan_{j(\bar{X}_E^*)}(j(\overline{X^*}_{a,E}) - j(\bar{X}_E)) \right\|^2$

with coverage error $O_p(n^{-2})$.

Note that $T(X^{*(g)}, \hat{Q}^{(g)})$ is obtained from $T(X^{(g)}, \hat{Q}^{(g)})$ by substituting $X_1^{(g)} = (X_{1,1}, \cdots, X_{g,1})^T$ with re samples $X_1^{*(g)} = (X_{1,1}^*, \cdots, X_{g,1}^*)^T$.

Using the bootstrap analogue in the previous Proposition 6.2.1 yields simpler method for finding $100(1 - c)\%$ confidence regions. We will utilize the tests statistics expressed above to conduct our analysis with confidence regions $C_{n,c}^*$ and $D_{n,c}^*$ as shown in the Corollary 6.2.2.

6.3 MANOVA on $(\mathbb{R}P^3)^q$

We start with the 3-dimensional real projective space $\mathbb{R}P^3$. It is a space of 1-dimensional linear subspaces of $\mathbb{R}^4$ and is also a 3-dimensional manifold. A projective point $p = [x] \in \mathbb{R}P^3$, is an equivalence class of $x = (x^1, \cdots, x^4) \in \mathbb{R}^4$ and can also be represented by $p = [x^1 : x^2 : x^3 : x^4]$ (homogeneous coordinates notation). We will identify $\mathcal{M} = \mathbb{R}P^3$ with the sphere S^3 with the antipodal points identified, $[x] = \{x, -x\} \in \mathbb{R}P^3, x \in \mathbb{R}^4, \ \|x\| = 1$. We will often refer to this identification as the *spherical representation* of the real projective space. $\mathbb{R}P^3$ is an embedded manifold with the embedding

$$\begin{aligned} j : \mathbb{R}P^3 &\to \mathcal{S}(4,\ \mathbb{R}) \\ j([x]) &= xx^T \end{aligned} \tag{6.19}$$

And for $[X]$ a random object on j-nonfocal probability measure Q on $\mathbb{R}P^3$ the projection $P_j : S_+(4, \mathbb{R})\backslash\mathcal{F} \to j(\mathbb{R}P^3)$ assigns to each nonnegative definite symmetric matrix A with highest eigenvalue of multiplicity 1, the matrix $j([\gamma])$, where γ is a unit eigenvector of A corresponding to its largest eigenvalue(see Bhattacharya and Patrangenaru [6]).

Our analysis will be conducted on $P\Sigma_3^k$, the projective shape space of 3D k-ads in $\mathbb{R}P^m$ for which $\pi = ([u_1], \ldots, [u_5])$ is a projective frame in $\mathbb{R}P^3$. $P\Sigma_3^k$ is homeomorphic to the manifold $\left(\mathbb{R}P^3\right)^{k-5}$ with $k-5 = q$ (see Patrangenaru et. al (2010)). The embedding on this space is the VW (Veronese-Whitney) embedding given by

$$\begin{aligned} j_k : \left(\mathbb{R}P^3\right)^q &\to (S(4, \mathbb{R}))^q \\ j_k([x_1], \ldots, [x_q]) &= (j([x_1]), \ldots, j([x_q])), \end{aligned} \tag{6.20}$$

with $j : \mathbb{R}P^3 \to S_+(4, \mathbb{R})$ the embedding given in (6.19). Additionally j_k is an equivariant embedding w.r.t. the group $(S_+(4, \mathbb{R}))^q$ and has the corresponding projection

$$\begin{aligned} P_{j_k} : (S_+(4, \mathbb{R}))^q \backslash \mathcal{F}_q &\to j_k \left(\mathbb{R}P^3\right)^q \\ P_{j_k}(A_1, \ldots, A_q) &= (j([m_1]), \ldots, j[m_q])) \end{aligned} \tag{6.21}$$

where $m_1, \ldots, m_q$ are unit eigenvectors of $A_1, \ldots, A_q$ (respectively) corresponding to the respective highest eigenvalues of those nonnegative definite symmetric matrices. Let Y be be a random object from a VW distribution Q on $(\mathbb{R}P^3)^q$, where $Y = (Y^1, \ldots, Y^q)$, and $Y^s = [X^s] \in \mathbb{R}P^3$ for all $s = \overline{1, q}$. The VW mean is given by

$$\mu_{j_k} = ([\gamma_1(4)], \cdots, [\gamma_q(4)]), \tag{6.22}$$

where, for $s = \overline{1, q}, \ \lambda_s(r)$ and $\gamma_s(r), r = 1, \ldots, 4$ are the eigenvalues in increasing order and the corresponding eigenvectors of $E\left[X^s(X^s)^T\right]$.

In case of a random object $[X]$ on $\mathbb{R}^3$, let us assume that $\mu_{E,j} = [\nu_4]$, where η_r and $\nu_r,\ r = 1,2,3,4$, are eigenvalues in increasing order and corresponding unit eigenvectors of $\mu = E[XX^T]$ corresponding to eigenvalues in their increasing order. The corresponding extrinsic sample mean, for a sample of size n, is given by $\overline{X}_{E,j} = [g(4)]$, where $d(r)$ and $g(r) \in \mathbb{R}^4,\ r = 1,2,3,4$, are eigenvalues in increasing order and corresponding unit eigenvectors of $J = \frac{1}{n}\sum_{i=1}^n X_i X_i^T$.

We now recall the result from Theorem 4.1 in Bhattacharya and Patrangenaru (2005) [6] well as represent the statistics

$$T([X],Q) = n\|S(j,X)^{-1/2}\tan_{j(\mu_{E,j})}\left(j(\overline{X}_{E,j}) - j(\mu_{E,j})\right)\|^2$$

We have for $T([X],Q) = T([X],[\nu_4])$

$$T([X],[\nu_4]) = n\ g(4)^T\ [(\nu_r)]_{r=1,2,3}S(j,X)^{-1}[(\nu_r)]_{r=1,2,3}^T\ g(4) \tag{6.23}$$

This results extends to the statistics

$$T([X],\hat{Q}) = T([X],[g(4)]) = \|S(j,X)^{-1/2}\tan_{j(\overline{X}_{E,j})}\left(j(\overline{X}_{E,j}) - j(\mu_{E,j})\right)\|^2$$

$$T([X],[g(4)]) = n\ \nu_4^T\ [g(r)]_{r=1,2,3}S(j,X)^{-1}[g(r)]_{r=1,2,3}^T\ \nu_4, \tag{6.24}$$

where

$$S(j,X)_{ab} = n^{-1}(d(4)-d(a))^{-1}(d(4)-d(b))^{-1} \times \sum_{i=1}^n (g(a)\cdot X_i)(g(b)\cdot X_i)(g(4)\cdot X_i)^2$$

and, asymptotically $T([X],[\nu_4])$ and $T([X],[g(4)])$ both have a χ^2_3 distribution.(see Bhattacharya and Patrangenaru (2005) [6])

Before we express our statistics of interest, it will be important to note another result from Crane and Patrangenaru (2011) [7] concerning the statistics

$$T(Y,\mu_{E,j_k}) = n\|S_{\bar{Y}}(j_k,Y)^{-1/2}\tan_{j(\overline{Y}_{E,j_k})}\left(j(\overline{Y}_{E,j_k}) - j(\mu_{E,j_k})\right)\|^2$$

And this Hotelling T^2 type statistic is given by

$$T(Y,([\gamma_1(4)],\cdots,[\gamma_q(4)])) = n\ \left(\gamma_1(4)^T D_1 \ldots \gamma_q(4)^T D_q\right)\ \ S_{\bar{Y}}(j_k,Y)^{-1}\left(\gamma_1(4)^T D_1 \ldots \gamma_q(4)^T D_q\right)^T \tag{6.25}$$

where for $s = 1,\ldots,q$ we have $D_s = (g_s(1)\ g_s(2)\ g_s(3)) \in \mathcal{M}(4,3,\mathbb{R})$ and for a pair of indices (s,a), $s = 1,\ldots,q$ and $a = 1,2,3$ in their lexicographic order we have

$$S_{\bar{Y}}(j_k,Y)_{(s,a),(t,b)} = n^{-1}(d_s(4)-d_s(a))^{-1}(d_t(4)-d_t(b))^{-1}\times\sum_{i=1}^n (g_s(a)\cdot X_i^s)(g_t(b)\cdot X_i^t)(g_s(4)\cdot X_i^s)(g_t(4)\cdot X_i^t) \tag{6.26}$$

In the next theorem we will take advantage of these results.

$$
\begin{aligned}
&H_0 : \ \mu_{1,E} = \mu_{2,E} = ... = \mu_{g,E} = \mu_E, \\
&H_a : \ \text{at least one equality } \mu_{a,E} = \mu_{b,E}, 1 \leq a < b \leq g \text{ does not hold.}
\end{aligned}
\tag{6.27}
$$

We aim to have an explicit representation of the expressions,

$$
T_c\left(Y^{(g)}, \mu_E^{(p)}\right) = n_a \sum_{a=1}^{g} \left\| S_{\bar{Y}}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\mu_E^{(p)}\right)} \left(j_k(\overline{Y}_{a,E}) - j_k\left(\mu_E^{(p)}\right)\right)\right\|^2 \tag{6.28}
$$

$$
T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right) = n_a \sum_{a=1}^{g} \left\| S_{\bar{Y}}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_{a,E}) - j_k\left(\mu_E^{(p)}\right)\right)\right\|^2 \tag{6.29}
$$

where $\mu_{a,E} = ([\nu_1^a(4)], \ldots, [\nu_q^a(4)])$ are the VW mean from distribution Q_a (of Y_{r_a}) and $(\eta_s^a(r), \nu_s^a(r))$, $r = 1, \ldots, 4$, are eigenvalues and corresponding unit eigenvectors of $E(X_{a,1}^s (X_{a,1}^s)^T]$. The corresponding VW sample mean is given by$\overline{Y}_{a,E} = ([g_1^a(4), \ldots, [g_q^a(4)])$ and for each $s = 1, \ldots, q$ we have for $r = 1, \ldots, 4$, $(d_s^a(r), g_s^a(r))$ which are eigenvalues in increasing order and corresponding unit eigenvectors of $J_s^a = \frac{1}{n_a} \sum_{i=1}^{n_a} X_{a,i}^s (X_{a,i}^s)^T$. Also $\mu_E^{(p)}$ is the VW pooled mean given by

$$
j_k\left(\mu_E^{(p)}\right) = P_{j_k}\left(\sum_{a=1}^{g} \frac{\lambda_a}{\lambda} j_k(\mu_{a,E})\right) \tag{6.30}
$$

$$
\mu_E^{(p)} = ([\gamma_1^{(p)}(4)], \ldots, [\gamma_q^{(p)}(4)]) \tag{6.31}
$$

and $\overline{Y}_E^{(p)}$ is the corresponding pooled mean, given by

$$
j_k\left(\overline{Y}_E^{(p)}\right) = P_{j_k}\left(\sum_{a=1}^{g} \frac{n_a}{n} j_k(\overline{Y}_{a,E})\right) \tag{6.32}
$$

$$
\overline{Y}_E^{(p)} = ([\mathbf{g}_1^{(p)}(4)], \ldots, [\mathbf{g}_q^{(p)}(4)]) \tag{6.33}
$$

where for $s = 1, \ldots, q$, $\mathbf{d}_s^{(p)}(r)$ and $\mathbf{g}_s^{(p)}(r) \in \mathbb{R}^4, \ r = 1, 2, 3, 4$, are eigenvalues in increasing order and corresponding unit eigenvectors of the matrix $J^{(p)} = \sum_{a=1}^{g} \frac{n_a}{n} j_k(\overline{Y}_{a,E})$.

We now express the following matrices

$$
\mathbf{C}_s = (\gamma_s^{(p)}(1)\ \gamma_s^{(p)}(2)\ \gamma_s^{(p)}(3)) \in \mathcal{M}(4, 3 : \mathbb{R}) \tag{6.34}
$$

$$
\mathbf{D}_s = (\mathbf{g}_s^{(p)}(1)\ \mathbf{g}_s^{(p)}(2)\ \mathbf{g}_s^{(p)}(3)) \in \mathcal{M}(4, 3 : \mathbb{R}) \tag{6.35}
$$

COROLLARY 6.3.1. *Assume j_k is the VW embedding of $(\mathbb{R}P^3)^q$ and $\{Y_{a,r_a}\}_{r_a=1,\dots,n_a}$, $a = 1,\dots,g$ are independent random samples from j_k-nonfocal probability measures Q_a on $(\mathbb{R}P^m)^q$ that have non degenerate j_k-extrinsic covariance matrices. Then the statistics*

(i) $T_c\left(Y^{(g)}, \mu_E^{(p)}\right) = \sum_{a=1}^g n_a \ \left((g_1^a(4))^T\mathbf{C}_1 \dots (g_s^a(4))^T\mathbf{C}_q\right) \ S_{\bar{Y}_a}(j_k, Y_a)^{-1} \left(g_1^a(4)^T\mathbf{C}_1 \dots g_q^a(4)^T\mathbf{C}_q\right)^T$

(ii) $$T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right) = \sum_{a=1}^g n_a \ \left[(\gamma_1^{(p)}(4) - g_1^a(4))^T\mathbf{D}_1 \dots (\gamma_q^{(p)}(4) - g_q^a(4))^T\mathbf{D}_q\right] S_{\bar{Y}_a}(j_k, Y_a)^{-1} \left[(\gamma_1^{(p)}(4) - g_1^a(4))^T\mathbf{D}_1 \dots (\gamma_q^{(p)}(4) - g_q^a(4))^T\mathbf{D}_q\right]^T.$$

where

$$S_{\bar{Y}_a}(j_k, Y_a)_{(s,c)(t,b)} = n_a^{-1}(\mathbf{d}_s^{(p)}(4) - \mathbf{d}_s^{(p)}(c))^{-1}(\mathbf{d}_t^{(p)}(4) - \mathbf{d}_t^{(p)}(b))^{-1} \times \sum_i (\mathbf{g}_s^{(p)}(c)\cdot X_{a,i}^s)(\mathbf{g}_t^{(p)}(b)\cdot X_{a,i}^t)(\mathbf{g}_s^{(p)}(4)\cdot X_{a,i}^s)(\mathbf{g}_t^{(p)}(4)\cdot X_{a,i}^t)$$

and $s, t = 1,\dots,q$ and $c, b = 1,\dots,m$. Both $T_c\left(Y^{(g)}, \mu_E^{(p)}\right)$ and $T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right)$ have, asymptotically a χ^2_{3q} distribution.

Proof. For part (i) we note that for each $a = 1,\dots g$ we get a natural extension of the result in theorem 4.1 Bhattacharya and Patrangenaru (2005) [6] as shown in 6.23.For part (ii) recall that

$$T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right) = n_a \sum_{a=1}^g \left\| S_{\bar{Y}_a}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_{a,E}) - j_k\left(\mu_E^{(p)}\right)\right)\right\|^2$$

we start by rewriting the expression above and we have

$$T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right) = n_a \sum_{a=1}^g \left\| S_{\bar{Y}_a}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_E^{(p)}) - j_k\left(\mu_E^{(p)}\right)\right) - S_{\bar{Y}_a}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_E^{(p)}) - j_k\left(\overline{Y}_{a,E}\right)\right)\right\|^2$$

$$T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right) = \sum_{a=1}^g n_a \left\| S_{\bar{Y}_a}(j_k, Y_a)^{-1/2} \left[(\gamma_1^{(p)}(4))^T\mathbf{D}_1 \dots (\gamma_q^{(p)}(4))^T\mathbf{D}_q\right]^T - S_{\bar{Y}_a}(j_k, Y_a)^{-1/2}\left[(g_1^a(4))^T\mathbf{D}_1 \dots (g_q^a(4))^T\mathbf{D}_q\right]^T\right\|^2 \tag{6.36}$$

□

If $Y_r a$ are j_k-nonfocal populations on $(\mathbb{R}P^3)^q$ we can construct an Edgeworth expansion up to order $O_p(n^{-2})$ of the pivotal statistics $T_c\left(Y^{(g)}, \mu_E^{(p)}\right)$ and $T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}\right)$. under the hypothesis

$$\begin{cases} H_0 & : \mu_{1,E} = \mu_{2,E} = ... = \mu_{g,E} = \mu_E^{(p)}, \\ H_a & :\ni (i,j) 1 \leq i < j < g, \text{s.t. } \mu_{i,E} \neq \mu_{j,E}. \end{cases}$$

COROLLARY 6.3.2. *The* $(1-c)100\%$ *bootstrap confidence regions for* μ_E *with* $d = gp$ *are given by*

(a) $C^*{}^{(g)}_{n,c} = j^{-1}(U^*_{n,c})$ *and* $U^*_{n,c} = \{j_k(\nu) \in j_k((\mathbb{R}P^3)^q) : T_c\left(Y^{(g)}, \nu\right) \leq c^*{}^{(g)}_{1-c}\}$ *where* $c^*{}^{(g)}_{1-c}$ *is the upper* $100(1-c)\%$ *point of the values*

$$T_c\left(Y^{*(g)}, \overline{Y}_E^{(p)}\right) = \sum_{a=1}^{g} n_a \ \left((g^*{}^a_1(4))^T \mathbf{D}_1 \ldots (g^*{}^a_s(4))^T \mathbf{D}_q\right) \ S_{\bar{Y}_a^*}(j_k, Y_a^*)^{-1} \left(g^*{}^a_1(4)^T \mathbf{D}_1 \ldots g^*{}^a_q(4)^T \mathbf{D}_q\right)^T \tag{6.37}$$

among the bootstrap re samples.

(b) $D^*{}^{(g)}_{n,c} = j^{-1}(V^*{}_{n,c})$ *and* $V^*{}_{n,c} = \{j_k(\nu) \in j_k((\mathbb{R}P^3)^q) : T_c\left(Y^{(g)}, \overline{Y}_E^{(p)}, \nu\right) \leq d^*{}^{(g)}_{1-c}\}$ *where* $T_d\left(Y^{(g)}, \overline{Y}_E^{(p)}, \nu\right) = n_a \sum_{a=1}^{g} \left\| S_{\bar{Y}_a}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_{a,E}) - j_k(\nu)\right) \right\|^2$ *where* $d^*{}^{(g)}_{1-c}$ *is the upper* $100(1-c)\%$ *point of the values*

$$T_d\left(Y^{*(g)}, \overline{Y^*}_E^{(p)}, \overline{Y}_E^{(p)}\right) = \sum_{a=1}^{g} n_a \ \left\| S_{\bar{Y}_a^*}(j_k, Y_a^*)^{-1/2} \ \tan_{j_k\left(\overline{Y}_E^{*(p)}\right)} \left(j_k(\overline{Y}^*_{a,E}) - j_k(\overline{Y}_E^{(p)})\right) \right\|^2 \tag{6.38}$$

among the bootstrap re samples. Both of the regions given by (6.16) *and* (6.14) *have coverage error* $O_p(n^{-2})$.

Note that here

$$S_{\bar{Y}_a^*}(j_k, Y_a^*)_{(s,c)(t,b)} = n_a^{-1} (\mathbf{d}_s^{*(p)}(4) - \mathbf{d}_s^{*(p)}(c))^{-1} (\mathbf{d}_t^{*(p)}(4) - \mathbf{d}_t^{*(p)}(b))^{-1}$$
$$\times$$
$$\sum_i (\mathbf{g}_s^{*(p)}(c) \cdot X_{a,i}^{*s})(\mathbf{g}_t^{*(p)}(b) \cdot X_{a,i}^{*t})(\mathbf{g}_s^{*(p)}(4) \cdot X_{a,i}^{*s})(\mathbf{g}_t^{*(p)}(4) \cdot X_{a,i}^{*t}), b, c = 1, 2, 3.$$

6.4 Application to face data

We will now test for the existence of 3D mean projective shape change to differentiate between three faces which are represented in Fig 6.4

Our analysis will be conducted on $g = 3$ individuals. The 3D reconstruction was done using the AGISOFT software. The images in Fig 6.4 represent 19 facial reconstructions. Each of those reconstruction was created

Figure 6.1: Faces used in MANOVA analysis

Figure 6.2: Sample of facial reconstructions

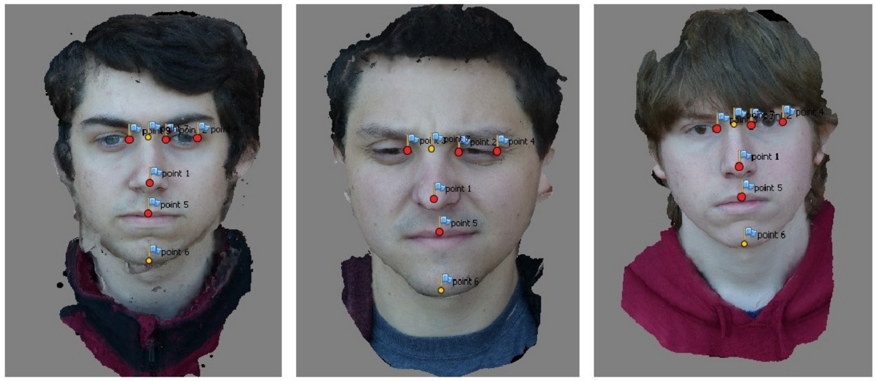


Figure 6.3: Projective frame shown in red

using mostly 4 to 5 digital camera images of a given individual. We are also able to place and recover 7 landmarks which are shown in figure 6.4.

Five of those landmarks (colored in red) will be used to construct a projective frame and the resulting two projective coordinate will determine our 3D projective shapes. We will compare these faces by conducting a MANOVA on manifold to compare $g = 3$ VW-means on $P\Sigma_3^7 = (\mathbb{R}P^3)^2$. For $n = \sum_{a=1}^{3} n_a = 19$ where $n_1 = 6$, $n_2 = 6$ and $n_3 = 7$ our hypothesis problem will be

$$H_0 : \ \mu_{1,E} = \mu_{2,E} = \mu_{3,E} = \mu_E,$$
$$H_a : \ \textit{at least one equation does not hold.}$$

Since the true pulled mean is unknown and our data set is relatively small we will reject the null hypothesis if

$T_d\left(Y^{(3)}, \overline{Y}_E^{(p)}\right) = \sum_{a=1}^{3} n_a \left\| S_{\bar{Y}_a}(j_k, Y_a)^{-1/2} \ \tan_{j_k\left(\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_{a,E}) - j_k(\overline{Y}_E^{(p)})\right)\right\|^2$ does not belong to $V^*{}_{n,c} = \{j_k(\nu) \in j_k((\mathbb{R}P^3)^2) : T_c\left(Y^{(3)}, \overline{Y}_E^{(p)}, \nu\right) \leq d^*{}_{1-c}^{(3)}\}$, where $d^*{}_{1-c}^{(3)}$ is the $(1-c)100\%$ cutoff of the corresponding bootstrap distribution.

Using 46800 resamples we obtain a value for $T_d\left(Y^{(3)}, \overline{Y}_E^{(p)}\right) = 757260$ and for the $d^*{}_{0.95}^{(3)} = 355420$ and we therefore reject the null hypothesis. And we conclude that there exist a statistically significant VW-mean projective shape face difference between at least two of the individuals.

CHAPTER 7

FUTURE WORK

In this chapter we explore some of the possible directions for extrinsic data analysis.

7.1 New test statistics for data on $(\mathbb{R}P^3)^q$ and MANOVA for anti-means

7.1.1 MANOVA cross validation

Although I was able to conclude effectively that there is a statistically significant VW-mean projective shape difference between at least two of the individuals, this test involved only $g = 3$. I would like to significantly increase the number g of samples to be compared in order to find the numerical limits of this particular method.

I would also like to use the data collected to conduct a cross-validation test. It will mean that I will compare g samples of the same face in order to verify that this method can in fact be used to properly differentiate between objects (faces, flours, etc...).

7.2 Anti-mean and MANOVA on manifolds

The results about the asymptotic of the anti-means are part of a joint paper with my colleague Ruite Guo and professor Patrangenaru (see Patrangenaru et all (2016b) [22]). I include this under future work, as more credit for this paper should be attributed to Ruite.

7.2.1 CLT for the sample anti-means

Assume j is an embedding of a d-dimensional manifold $\mathcal{M}$ such that $j(\mathcal{M})$ is closed in $\mathbb{R}^k$, and Q is a αj-nonfocal probability measure on M such that $j(Q)$ has finite moments of order 2. Let μ and Σ be the mean and covariance matrix of $j(Q)$ regarded as a probability measure on $\mathbb{R}^k$. Let $\mathcal{F}$ be the set of αj-focal points of $j(\mathcal{M})$, and let $P_{F,j} : \mathcal{F}^c \to j(\mathcal{M})$ be the projection on $j(\mathcal{M})$. $P_{F,j}$ is differentiable at μ and has the differentiability class of $j(\mathcal{M})$ around any αj-nonfocal point.

Assume $x \to (f_1(x), \ldots, f_d(x))$ is a local frame field on an open subset of M such that for each $x \in M$, $(d_x j(f_1(x)), \ldots, d_x j(f_d(x)))$ are orthonormal vector in $\mathbb{R}^k$. A local frame field $p \to (e_1(p), e_2(p), \ldots, e_k(p))$

defined on an open neighborhood $U \subseteq \mathbb{R}^k$ is adapted to the embedding j if it is an orthonormal frame field and $\forall x \in j^{-1}(U), e_r(j(x)) = d_x j(f_r(x)), r = 1, \ldots, d$.

Let $e_1, e_2, \ldots, e_k$ be the canonical basis of $\mathbb{R}^k$ and assume $(e_1(p), e_2(p), \ldots, e_k(p))$ is an adapted frame field around $P_{F,j}(\mu) = j(\mu_{\alpha E})$. Then $d_\mu P_{F,j}(e_b) \in T_{P_{F,j}(\mu)} j(\mathcal{M})$ is a linear combination of $e_1(P_{F,j}(\mu)), e_2(P_{F,j}(\mu)), \ldots, e_d(P_{F,j}(\mu))$:

$$d_\mu P_{F,j}(e_b) = \sum_{a=1}^{d} (d_\mu P_{F,j}(e_b)) \cdot e_a(P_{F,j}(\mu)) e_a(P_{F,j}(\mu)). \tag{7.1}$$

By the delta method, $n^{1/2}(P_{F,j}(\overline{j(X)}) - P_{F,j}(\mu))$ converges weakly to $N_k(0_k, \alpha\Sigma_\mu)$, where $\overline{j(X)} = \frac{1}{n}\sum_{i=1}^{n} j(X_i)$ and

$$\begin{aligned} \alpha\Sigma_\mu = &[\sum_{a=1}^{d} d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu)) e_a(P_{F,j}(\mu))]_{b=1,\ldots,k} \\ &\times \Sigma[\sum_{a=1}^{d} d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu)) e_a(P_{F,j}(\mu))]_{b=1,\ldots,k}^T \end{aligned} \tag{7.2}$$

Here Σ is the covariance matrix of $j(X_1)$ w.r.t the canonical basis $e_1, e_2, \ldots, e_k$.

The asymptotic distribution $N_k(0_k, \alpha\Sigma_\mu)$ is degenerate and the support of this distribution is on $T_{P_{F,j}} j(\mathcal{M})$, since the range of $d_\mu P_{F,j}$ is $T_{P_{F,j}(\mu)} j(\mathcal{M})$. Note that $d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu)) = 0$ for $a = d+1, \ldots, k$. we obtain the following asymptotic result, our CLT for extrinsic anti-mean, on the tangent space of $j(\mathcal{M})$ at $P_{F,j}(\mu) = j(\alpha\mu_E)$.

$$tan_{P_{F,j}(\mu)} \left(P_{F,j}(\overline{(j(X))}) - P_{F,j}(\mu) \right) \to_d N(0, \alpha\Sigma_{j,E}) \tag{7.3}$$

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

$$\begin{aligned} \alpha\Sigma_{j,E} &= e_a(P_{F,j}(\mu))^t \alpha\Sigma_\mu e_b(P_{F,j}(\mu))_{1 \le a,b \le d} \\ &= [\Sigma d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu))]_{a=1,\ldots,d} \Sigma \\ &\quad \times [\Sigma d_\mu P_{F,j}(e_b) \cdot e_a(P_{F,j}(\mu))]_{a=1,\ldots,d}^T \end{aligned} \tag{7.4}$$

The matrix $\alpha\Sigma_{j,E}$ given above is the extrinsic anti-covariance matrix of the αj -nonfocal distribution Q(of X_1) w.r.t. the basis $f_1(\mu_{\alpha E}), \ldots, f_d(\mu_{\alpha E})$.

7.2.2 MANOVA for anti-means

I will start by considering the following extension to my MANOVA on manifolds method, from Chapter 6.

DEFINITION 7.2.1. *Under the assumption $\alpha A_0 : \alpha\mu_{1,E} = \cdots = \alpha\mu_{g,E}$ and for any $a \in \{1, 2, ..., g\}$, with $\frac{n_a}{n} \to \lambda_a > 0,\ as\ n \to \infty$. We define*

(i) The **extrinsic pooled anti-mean with weights** $\lambda = (\lambda_1, \ldots, \lambda_g)$*, denoted $\alpha\mu_E(\lambda)$ as the value in $\mathcal{M}$ given by*

$$j(\alpha\mu_E) = P_{F,j}(\lambda_1 j(\alpha\mu_{1,E}) + \cdots + \lambda_g j(\alpha\mu_{g,E})) \tag{7.5}$$

Where $\alpha\mu_{a,E}$ is the extrinsic anti-mean of the random object $X_{a,1}$ and $\Sigma_{a=1}^{g}\lambda_a = 1$

(ii) The **extrinsic sample pooled anti-mean** *denoted $a\bar{X}_E \in \mathcal{M}$ given by;*

$$j(a\bar{X}_E) = P_{F,j}\left(\frac{n_1}{n} j(a\bar{X}_{1,E}) + \cdots + \frac{n_g}{n} j(a\bar{X}_{g,E})\right), \tag{7.6}$$

where $a\bar{X}_{a,E}$ is the extrinsic sample anti-mean for $X_{a,1}$ and $n = \sum_{a=1}^{g} n_a$

With this definition at hand, I can now express the following hypothesis test, designed to test the difference between extrinsic anti-means and is given by;

$$H_0:\ \alpha\mu_{1,E} = \alpha\mu_{2,E} = ... = \alpha\mu_{g,E} = \alpha\mu_E, \tag{7.7}$$

$$H_a:\ at\ least\ one\ equality\ \alpha\mu_{a,E} = \alpha\mu_{b,E}, 1 \le a < b \le g\ does\ not\ hold.$$

The results in chapter 6 can be adapted to extrinsic anti-means and pooled anti-means as well and I will take advantage of these results. After some effort I will be able to have an explicit representation of the expressions,

$$\alpha T_c\left(Y^{(g)}, \alpha\mu_E^{(p)}\right) = \sum_{a=1}^{g} \left\| aS_{\bar{Y}}(j_k, Y_a)^{-1/2} \tan_{j_k\left(\alpha\mu_E^{(p)}\right)} \left(j_k(a\overline{Y}_{a,E}) - j_k\left(\alpha\mu_E^{(p)}\right)\right)\right\|^2 \tag{7.8}$$

$$\alpha T_d\left(Y^{(g)}, a\overline{Y}_E^{(p)}\right) = \sum_{a=1}^{g} \left\| aS_{\bar{Y}}(j_k, Y_a)^{-1/2} \tan_{j_k\left(a\overline{Y}_E^{(p)}\right)} \left(j_k(\overline{Y}_{a,E}) - j_k\left(\alpha\mu_E^{(p)}\right)\right)\right\|^2, \tag{7.9}$$

where $\alpha\mu_{a,E} = ([\nu_{1,}^a(1)], \ldots, [\nu_q^a(1)])$ are the VW anti-mean from distribution Q_a (of Y_{r_a}) and $(\eta_s^a(r), \nu_{s,}^a(r))$ are eigenvalues and corresponding unit eigenvectors of $E(X_{a,1}^s(X_{a,1}^s)^T]$. The corresponding VW sample anti-mean is given by$a\overline{Y}_{a,E} = ([g_1^a(1)], \ldots, [g_q^a(1)])$ and for each $s = 1, \ldots, q$ we have for $r = 1, \ldots, 4$, $(d_s^a(r), g_s^a(r))$ which are eigenvalues in increasing order and corresponding unit eigenvectors of

$J_s^a = \frac{1}{n_a}\sum_{i=1}^{n_a} X_{a,i}^s (X_{a,i}^s)^T$. Also $\alpha\mu_E^{(p)}$ is the VW pooled mean given by

$$j_k\left(\alpha\mu_E^{(p)}\right) = P_{F,j_k}\left(\sum_{a=1}^{g}\frac{n_a}{n} j_k(\alpha\mu_{a,E})\right) \tag{7.10}$$

$$\alpha\mu_E^{(p)} = ([\gamma_1^{(p)}(1)], \ldots, [\gamma_q^{(p)}(1)]) \tag{7.11}$$

and $a\overline{Y}_E^{(p)}$ is the corresponding pooled sample anti-mean, given by

$$j_k\left(a\overline{Y}_E^{(p)}\right) = P_{F,j_k}\left(\sum_{a=1}^{g}\frac{n_a}{n} j_k(a\overline{Y}_{a,E})\right) \tag{7.12}$$

$$a\overline{Y}_E^{(p)} = ([\mathbf{g}_1^{(p)}(1)], \ldots, [\mathbf{g}_q^{(p)}(1)]), \tag{7.13}$$

where for $s = 1, \ldots, q$, $\mathbf{d}_s^{(p)}(r)$ and $\mathbf{g}_s^{(p)}(r) \in \mathbb{R}^4,\ r = 1, 2, 3, 4$, are eigenvalues in increasing order and corresponding unit eigenvectors of the matrix $J^{(p)} = \sum_{a=1}^{g}\frac{n_a}{n} j_k(\mu_{a,E})$.

I will then be able to construct confidence regions for $\alpha\mu_E^{(p)}$ of asymptotic level $1-c$ much like in the case of VW means, and when our sample size is relatively small we will be able to build a $(1-c)100\%$ confidence regions for $\alpha\mu_E^{(p)}$ using nonparametric bootstrap. These confidence regions will be the tool I will use to differentiate between different objects.

7.3 Dependence on embedded manifolds

We are interested in determining the dependence between the random objects, X on $\mathbb{S}^2$ and Y a random variable. And for that we start by observing the dependence structure between $\iota(X)$ a random vector in $\mathbb{R}^3$ and Y a random variable. We will call upon copula functions to start this process. At this point it is important to note that copula functions have been widely used to model the dependence structure between random vectors which is of importance in the computation of certain financial products such as VAR (Value At Risk). And the copula framework offers a wide variety of copulas, such as the Gaussian, student t copula, Frank's copula, Archimedes family of copula and so on. We will focus on only one type of copula, the Gaussian copula. We first define a two dimensional copula function.

DEFINITION 7.3.1. *The copula function C is a copula for the random vector (X, Y) with $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^k$, if it is the joint distribution of the random vector (U, V) where $U = F_1(X)$, and $V = F_2(Y)$ and $F_a,\ a = 1, 2,$ are the marginal distribution functions of X and Y respectively. This implies that*

$$H(x, y) = C(F_1(x), F_2(y)) = C(u, v) \tag{7.14}$$

Where H is the joint distribution function of (X, Y). If F_1 and F_2 are continuous the copula C is unique.

Note that

$$P(X \leq x, Y \leq y) = P(F_1(X) \leq F_1(x), F_2(Y) \leq F_2(y)) = C(F_1(x), F_2(y))$$

The results of the Sklar Theorem (see Rockinger and Jondeau (2001) [29]) show that we may link any group of univariate distributions, of any type with any copula and we will have defined a valid multivariate distribution.

DEFINITION 7.3.2. *[Gaussian Copula] This copula is given by*

$$C_{Gaussian}(u, v) = P(\Phi(X) \leq u, \Phi(Y) \leq v) = \Phi_{\Sigma}(\Phi^{-1}(u), \Phi^{-1}(v)) \tag{7.15}$$

where Φ is the standard normal cdf and Φ_{Σ} is the joint distribution function of a standard Gaussian random vector $\mathbf{Z} = (X, Y)^T \sim N_2(0, \Sigma)$. Note that Σ can also be viewed as a correlation matrix of $\mathbf{Z}$. And in two dimensions we have

$$C_{Gaussian}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{ \frac{-(s_1^2 - 2\rho s_1 s_2 + s_2^2)}{2(1-\rho^2)} \right\} ds_1 ds_2 \tag{7.16}$$

(see [28].)

REMARK 7.3.1. *It is important to note that U and V are independent if and only if the correlation matrix Σ is the identity. Recall that in the case of Gaussian random vector this result holds and $C_{Gaussian}(u, v) = uv$.*

PROPOSITION 7.3.1. *Let X and Y be random vectors on $\mathbb{R}^m$ and $\mathbb{R}^k$ respectively then X and Y are independent if and only if $U = F_1(X)$ and $V = F_2(Y)$ (viewed as random variables) are independent.*

Proof. Note that X and Y independent implies $H(x, y) = P(X \leq x)P(Y \leq y) = F_1(x)F_2(y) = uv = C(u, v)$ and we conclude that U and V are independent (recall the cdf of a uniform $U(0, 1)$ is $F(u|(0, 1)) = u$). The other direction follows from the same set of equalities. For the direction from left to right please see [1]. $\square$

We will now use the proposition above along with the useful property of the Gaussian copula correlation matrix to design an independence test.

7.3.1 Test for independence

Now back to our data set made up of X a random object on $\mathbb{S}^2$ and Y a random variable on $\mathbb{R}$. We will first use the proposition and Gaussian copula to test for independence between the embedded variable $\iota(X)$ (random vector on $\mathbb{R}^3$) and Y a random variable on $\mathbb{R}$. We will also assume that F_1 and F_2 are, respectively, the cdf's of $\iota(X)$ and Y. We can now do the following

1. Define $U = F_1(\iota(X))$ and $V = F_2(Y)$
2. Find the Gaussian Copula that fit our random vectors U and V. This process is done using Matlab and the function called **copulafit(...,)**
3. After fitting, the resulting correlation matrix is used to conclude dependence between U and V
4. Once the dependence is established we draw the necessary conclusion about $\iota(X)$ and Y, by relying on proposition 7.3.1

PROPOSITION 7.3.2. *The random object X and the random variable Y are independent if and only if $U = F_1(\iota(X))$ and $V = F_2(Y)$ are independent random variables.*

Proof. From the proposition 7.3.1 we have that $\iota(X)$ and Y are independent iff U and V are independent. And since ι is one-to-one we have our desired result. (see [28]) $\square$

Step one above, requires knowledge of the cdf's of the marginal distributions of $\iota(X)$ and Y which may not be known at the time. Now assume that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are i.i.d random objects from a joint distribution on $(\mathbb{S}^2, \mathbb{R})$ with marginal cdf's F_1 and F_2 respectively. We can use the corresponding empirical cdf's $\hat{F}_1$ and $\hat{F}_2$. We can then use the following steps,

1. Define $\hat{U} = \hat{F}_1(\iota(X))$ and $\hat{V} = \hat{F}_2(Y)$
2. Find the Gaussian Copula that fit our random vectors $\hat{U}$ and $\hat{V}$. This process is done using Matlab and the function called **copulafit(...,)**
3. After fitting, the resulting correlation matrix is used to conclude dependence between U and V
4. Once the dependence is established we draw the necessary conclusion about $\iota(X)$ and Y, by relying on proposition 7.3.2

BIBLIOGRAPHY

[1] R. B. Ash and C. A. Doléans-Dale. *Probability & Measure Theory*. Harcourt Academic Press, USA, 2000.

[2] V. Balan, M. Crane, V. Patrangenaru, and X. Liu. Projective shape manifolds and coplanarity of landmark configurations. a nonparametric approach. *Balkan Journal of Geometry and Its Applications*, 14:1– 10, 2009.

[3] Rabi N. Bhattacharya, Marius Buibas, Ian L. Dryden, Leif A. Ellingson, David Groisser, Harrie Hendriks, Stephan Huckemann, Huiling Le, Xiuwen Liu, James S. Marron, Daniel E. Osborne, Vic Patrângenaru, Armin Schwartzman, Hilary W. Thompson, and Andrew T. A.Wood. Extrinsic data analysis on sample spaces with a manifold stratification. *Advances in Mathematics, Invited Contributions at the Seventh Congress of Romanian Mathematicians, Brasov, 2011.*, pages 241–252, 2013.

[4] R.N. Bhattacharya, L. Lin, and V. Patrangenaru. *A Course in Mathematical Statistics and Large Sample Theory*. Statistics Series. Springer, New York, USA, 2016.

[5] R.N. Bhattacharya and V. Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds-part one. *The Annals of Statistics*, 31:1–29, 2003.

[6] R.N. Bhattacharya and V. Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds-part two. *The Annals of Statistics*, 33:1211– 1245, 2005.

[7] M. Crane and V. Patrangenaru. Random change on a lie group and mean glaucomatous projective shape change detection from stereo pair images. *Journal of Multivariate Analysis*, 102:225–237, 2011.

[8] Bradley Efron. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7:1–26, 1979.

[9] L. Ellingson, V. Patrangenaru, H. Hendriks, and P. S. Valentin. *Topics in Nonparametric Statistics. Editors: M.G. Akritas, S.N. Lahiri and D. N. Politis*, chapter CLT on Low Dimensional Stratified Spaces, pages 227–240. Springer, 2014.

[10] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? *Proc. European Conference on Computer Vision, LNCS 588*, pages 563–578, 1992.

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

[12] R. I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 761 – 764, 1992.

[13] Thomas Hotz, Stephan Huckemann, Huiling Le, James S. Marron, Jonathan C. Mattingly, Ezra Miller, James Nolen, Megan Owen, Vic Patrangenaru, and Sean Skwerer. Sticky central limit theorems on open books. *Annals of Applied Probability*, 23:2238–2258, 2013.

[14] Stephan Huckemann. Formation of stress fibres in adult stem cells. *https://www.fields.utoronto.ca/programs/scientific/13-14/modelmethods/slides/Huckemann.pdf*, 2014.

[15] R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, USA, 2008.

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

[17] J. T. Kent. New directions in shape analysis. *The Art of Statistical Science, A Tribute to G. S. Watson. Published in Wiley Ser. Probab. Math. Statist. Probab. Math. Statist.*, pages 115–127, 1992.

[18] J. M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, USA, 2002.

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

[20] K. V. Mardia and V. Patrangenaru. Directions and projective shapes. *The Annals of Statistics*, 33:1666–1699, 2005.

[21] V. Patrangenaru and L. A. Ellingson. *Nonparametric Statistics on Manifolds and their Applications*. Chapman & Hall/CRC Monographs on Statistics & Applied Probab. CRC Press, FL, USA, 2015.

[22] V. Patrangenaru, Ruite Guo, and K. D. Yao. Nonparametric inference for location parameters via féchet functions. *Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, Beer Sheva, Israel*, pages 254–262, 2016.

[23] V. Patrangenaru, X. Liu, and S. Sugathadasa. Nonparametric 3D projective shape estimation from pairs of 2d images - i. *Journal of Multivariate Analysis*, 101:11–31, 2010.

[24] V. Patrangenaru, R. Paige, K. Yao, M. Qiu, and D. Lester. Projective shape analysis of contours and finite 3D configurations from digital camera images. *Statistical Papers*, 57:1017–1040, 2016.

[25] V. Patrangenaru, M. Qiu, and M. Buibas. Two sample tests for mean 3D projective shapes from digital camera images. *Methodology and Computing in Applied Probability*, 16:485–506, 2014.

[26] V. Patrangenaru, K. D. Yao, and V. Balan. 3D face analysis from digital camera images. *BSG Proceedings. The International Conference Differential Geometry - Dynamical Systems DGDS-2015*, 23:43–55, 2016.

[27] V. Patrangenaru, K. D. Yao, and Ruite Guo. Extrinsic means and antimeans. *2nd ISNPS, Cadiz, June 2014 Editors: Cao, Ricardo; Gonzalez Manteiga, Wenceslao; Romo, Juan*, 175:161–178, 2016.

[28] E. Platen and N. Bruti-Liberati. *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, volume 64 of *Stochastic Modeling and Applied Probability*. Springer, New York, USA, 2010.

[29] M. Rockinger and E. Jondeau. Conditional dependency of finanical series: An application of copulas. *NER (Notes d' Étude et de Recherche*, 82, 2001.

BIOGRAPHICAL SKETCH

The author was born in Abidjan, the economic capital city of Ivory Coast (officially named Côte d' Ivoire). He spent most of his younger life between his country of birth and Dakar, the capital city of Senegal. His education prior to coming to the United States took place in Dakar. Upon graduation from high school, he continued his education in the state of Arkansas where he obtained a Bachelor and a Masters in sciences in mathematics. After spending some time as a high school teaching faculty he decided to pursue a doctorate degree in Mathematics at Florida State University. He is the proud father of a son and is happily married to his lovely wife. The author's current research interests revolve around data analysis on manifold more specifically on spherical data and on 3D projective shapes.