Copyright Notice

Huckemann, S.; Hotz, T.; Munk, A.: *Intrinsic MANOVA for Riemannian Manifolds with an Application to Kendall's Space of Planar Shapes,* IEEE Transactions on Pattern Analysis and Machine Intelligence

---

Intrinsic MANOVA for Riemannian Manifolds with an Application to Kendall's Space of Planar Shapes

Stephan Huckemann, Thomas Hotz, and Axel Munk

Abstract—We propose an intrinsic multi-factorial model for data on Riemannian manifolds as typically occur in the statistical analysis of shape. Due to the lack of a linear structure, linear models cannot be defined in general; to date only one-way MANOVA is available. For a general multi-factorial model, we assume that variation not explained by the model is concentrated near elements defining the effects. By determining the asymptotic distributions of respective sample covariances under parallel transport, we show that they can be compared by standard MANOVA. Often in applications, manifolds are only implicitly given as quotients where the bottom space parallel transport can be expressed through a differential equation. For Kendall's space of planar shapes, we provide an explicit solution. We illustrate our method by an intrinsic two-way MANOVA for a set of leaf shapes. While biologists can identify genotype effects by sight, we can detect height effects that are otherwise not identifiable.

Index Terms—Shape Analysis, Non-Linear Multivariate Analysis of Variance, Riemannian Manifolds, Orbifolds, Orbit Spaces, Geodesics, Lie Group Actions, Non-Linear Multivariate Statistics, Covariance, Inference, Test, Intrinsic Mean, Forest Biometry

1 Introduction

In high level image analysis and understanding, e.g. for security or morphology related issues in biometry, reliable pattern recognition is of particular interest. One may either model biometric feature expression in a way directly reflecting an underlying specific image understanding (e.g. [20]), or in a rather generic setting using statistics to obtain a "low dimensional" feature expression. Other potential applications of interest may include geophysical measurements or measurements of the shape of molecules as they occur in structural biology and drug design. In this work our aim is to generalize statistical testing methodology. In the context of image analysis research on this is rather sparse, for some references see [5] or [6].

We undertake this by generalizing "linear models" to data on Riemannian manifolds. In classical multivariate analysis (e.g. [37], [3] and [8]), linear models such as the two-factorial model

$$Z_{i,j,n} \;=\; \mu_{i,j} + \epsilon_{i,j,n} \tag{1}$$

serve as a powerful tool to identify and discriminate between multiple effects. Here, i, j denote the levels of two factors and the replications are numbered by n. Due to the linear formulation, *multivariate analysis of variance* (MANOVA) can be employed to test corresponding hypotheses. Although classical MANOVA is developed under the assumption of normality, it is known that MANOVA is to some extent robust to nonnormality if fourth moments are finite (cf. [36, p. 378-9]). In this paper we are concerned with an extension of this method to data on more general, non-linear spaces. Usually, values of random variables on non-linear spaces permit no additive coupling thus rendering the very notion of "linear models" inappropriate. In some applications (cf. e.g. [27] as well as [13]) a non-commutative multiplication may be available making the definition of a non-commutative multiplicative model possible. In our work, however, we aim at more general spaces as occur for example in the statistical analysis of shape. In MANOVA it is important for interpretation to decompose the effects $\mu_{i,j} = \alpha_i + \beta_j + \gamma_{i,j}$ into main and interaction effects. We stress that for our purposes additive models like

$$Z_{i,j,n} \;=\; \alpha_i + \beta_j + \gamma_{i,j} + \epsilon_{i,j,n}$$

can in general not be formulated and verified by a decomposition of variance. The difficulty there is in general two-fold: firstly, because of the lack of a commutative multiplication, the model would depend on the order of the effects; and secondly, there is no multiplication in general, i.e. the very notion of "effects" (in particular, acting distinctly corresponding to $\gamma_{i,j} = 0$) is not at all obvious and an issue of separate research ([19], e.g. what does it mean that two different shapes are deformed "in the same way"?). Rather one may assume different *intrinsic means* $\mu_{i,j}$ (equivalently an "expected value" ([42]) or a "center of gravity" w.r.t. an intrinsic non-Euclidean metric ([30, p.109]) for each combination of levels resulting in a test of a hypothesis

H_0: there are μ_i such that $\mu_{i,j} = \mu_i$ for all levels j

versus the alternative

• *S. Huckemann, T. Hotz and A. Munk are with the Institute of Mathematical Stochastics, Georgia Augusta University, Göttingen, Germany. E-mail: huckeman@math.uni-goettingen.de*

Manuscript received October 15, 2008.

H_1: there is a pair $(i,j) \neq (i,j')$ with $\mu_{i,j} \neq \mu_{i,j'}$.

Suppose that random variables $P_{i,j,n}$ are distributed on a Riemannian manifold. Effects explained by the model result in distributions around different intrinsic means for each level. These distributions carry the remaining variation which, under the model hypothesis, is confined to a local neighborhood of such a mean,

$$X'_{i,j,n} = \exp^{-1}_{\mu_{i,j}}(P_{i,j,n}) \,. \tag{2}$$

Here, the Euclidean data $Z_{i,j,n}$ of (1) corresponds to data $P_{i,j,n}$ on the manifold in (2), the Euclidean errors $\epsilon_{i,j,n}$ correspond to tangent space valued errors $X'_{i,j,n}$ under the inverse *Riemann exponential*, the tangent space located at the mean $\mu_{i,j}$.

In order to compare error covariances in different tangent spaces with one-another, some *connection* between these tangent spaces is necessary. If the error distributions are anisotropic then the desired connection should respect anisotropy as well. As such a connection, the *Levi-Civitá connection*, also called *parallel transport* of the corresponding Riemannian structure qualifies. In fact in a Euclidean space, parallel transport guarantees the "linearity" of the corresponding linear model. In general, parallel transport from one point to another is unique only if there is a unique geodesic segment of minimal length joining the two. Hence, we assume that all random shape variables are supported by a subset of the Riemannian manifold in which any point can be connected by a unique minimizing geodesic to a prespecified offset. For most realistic data this will be the case with probability one.

Within this framework, intrinsic MANOVA can be performed based on classical MANOVA thus allowing for testing of different models as we show in Sections 2 and 3. We note that testing a model with one global effect versus a single factorial model with several levels does not require parallel transport: since the means do not differ under the null hypothesis, the test can be performed in the overall mean's tangent space. For the intrinsic mean, such tests have been proposed by [4]. Using extrinsic means rather than intrinsic means, corresponding tests are available from [10, Chapter 7], [17] as well as [4]. Testing models with at least one factor influencing the outcome, however, requires a connection of different tangent spaces which is naturally provided by parallel transport.

For spheres, parallel transport is easily accessable, cf. Example A.2. Often in applications however, the Riemannian manifold in question is given only as a quotient and thus parallel transport is not explictly available. In this case a formula due to [40] can be used to determine parallel transport. We develop the necessary details in the appendix and explicitly determine in a short calculation parallel transport on Kendall's space of planar shapes (which are essentially complex projective spaces) in Appendix A.2. We note that [33] computed parallel transport differently, based on which, [31] extended spline-fitting for spherical curves by [24] to shape curves. In another method [12] extended polynomial regression for intrinsic curve fitting.

We conclude our work with an application to forest biometry in Section 4 that, in fact, initiated this research. Two-dimensional shapes of poplar leafs are modeled by genotype and height levels. To the biologist, the first factor's influence is identifiable by sight, the second factor's is not. Intrinsic MANOVA for Kendall's spaces of planar shapes, however, identifies height effects as well.

2 Multi-Factorial Models for Manifolds

In this section we consider random points on manifolds. First we review the concept of intrinsic means. Then we formulate the *multi-factorial model for manifolds* and reduce the corresponding test to classical MANOVA applied to nonnormal distributions. In conclusion we discuss the robustness of classical MANOVA in case of nonnormality.

The following terminology for Riemannian manifolds can be found in any introductory textbook, e.g. [29]. The concept of *parallel transport* is introduced in detail in Appendix A, in particular parallel transport for Kendall's spaces of planar shapes is computed in Appendix A.2.

Throughout this section, M is a complete Riemannian D-dimensional manifold, with induced distance $d : M \times M \to [0,\infty)$. $\exp_p : T_pM \to M$ denotes the *Riemannian exponential map* from the tangent space at $p \in M$ to the manifold. Its inverse $\log_p : U \to T_pM$ is well defined in a neighborhood U around p. In particular, every $p' \in U$ can be reached from p by a unique geodesic of minimal length $d(p,p')$. For $p' \in U$,

$$\theta_{p,p'} : T_pM \to T_{p'}M$$

denotes the unique *parallel transport* of tangent spaces along geodesics of minimal length induced by the *Riemannian connection*, cf. Definition A.1. Example A.2 gives parallel transport for spheres which is used to compute parallel transport on Kendall's space of planar shapes in Theorem A.6.

A neighborhood U is called *convex* if for every $p', p'' \in U$ the connecting geodesic segment of minimal length is fully contained in U. In particular, $\log_{p'}$ is well defined on a convex neighborhood U for all $p' \in U$. Finally denote by

$$B_r(p) := \{\exp_p(v) : v \in T_pM, \|v\| < r\}$$

the *geodesic ball of radius* r around p. It is well-known that geodesic balls are convex for sufficiently small $r > 0$, e.g. [29, p.166].

2.1 Intrinsic Means on Manifolds

Suppose that P is a random element on M. Any minimizer

$$\mu \in \operatorname{argmin}_{p\in M} E\big(d(P,p)^2\big)$$

is called an *intrinsic population mean* of P. The intrinsic mean is unique under the following condition, cf. [25, p. 510-1],

$$\left.\begin{array}{l}\text{there are } p \in M \text{ and } r > 0 \text{ with } d(P,p) < r \text{ a.s.,}\\ B_r(p) \text{ is convex with sectional curvatures } \leq \kappa,\\ \text{if } \kappa > 0 \text{ then } 4r < \frac{\pi}{\sqrt{\kappa}}.\end{array}\right\} \quad (3)$$

For example, if M is a sphere, and P is contained in a proper quarter-sphere then the intrinsic mean is unique.

For $P_1, \ldots, P_n$ i.i.d. as P, any minimizer

$$\overline{P}_n \quad \in \quad \mathrm{argmin}_{p \in M} \sum_{i=1}^{n} d(P_i, p)^2 \quad (4)$$

is an *intrinsic sample mean*. [42] established a strong law of large numbers (SLLN) for quasi-metrical spaces. In particular under (3) we have that

$$\overline{P}_n \to \mu \text{ a.s. if } E\big(d(P,p)^2\big) < \infty \text{ for at least one } p \in P. \quad (5)$$

If P has an intrinsic mean μ and if P assumes values only within a neighborhood U of μ in which the inverse Riemannian exponential $\log_\mu$ is well defined then

$$E(\log_\mu \circ P) \quad = \quad 0 \quad (6)$$

([30, p.110-111] and [32]) with the usual expectation E of the multivariate real random variable $X = \log_\mu \circ P$. In fact, (6) characterizes intrinsic means, based on which [32] developed an algorithm for computing intrinsic means and discussed its convergence. An alternate algorithmic method in [23] is based on Lagrange multipliers.

2.2 Two-Factorial Models for Manifolds

In this section we formulate *multi-factorial models* for data on a D-dimensional manifold M. In fact, it suffices to formulate a *two-factorial model* since any larger number of factors can be viewed as two factors with more levels, as all interaction factors are assumed to be present.

Hence, we suppose that we have random elements $P_{i,j,n}$ on M due to the effects of levels

$$i = \{1, \ldots, I\}, \ j = \{1, \ldots, J\}$$

of two factors, n denoting replications. In particular we assume that the $P_{i,j,n}$ have a "common" distribution around certain effects $\mu_{i,j}$. In a Euclidean scenario this is saying that the distributions of the $P_{i,j,n}$ agree modulo linear translations, in particular, involving no rotations. On a general Riemannian manifold this concept naturally generalizes to the requirement that the distributions of the $P_{i,j,n}$ agree modulo *parallel transport* to a specific location $\nu \in M$, i.e. that the

$$X_{i,j,n} \quad := \quad \theta_{\mu_{i,j},\nu} \circ \log_{\mu_{i,j}} \circ P_{i,j,n} \ \in \ \mathbb{R}^D \quad (7)$$

are identically distributed. Here, $\theta_{\mu_{i,j},\nu}$ denotes the parallel transport as in Definition A.1 from $\mu_{i,j}$ to ν sufficiently close to $\mu_{i,j}$ such that the parallel transport is well defined. Note that slightly different from (2) we included parallel transport here in order to obtain values in a common tangent space at ν.

Definition 2.1. *Let M be a complete Riemannian manifold, $I, J \in \mathbb{N}$. In the* two-factorial model for manifolds *we assume for $1 \leq i, i' \leq I$, $1 \leq j, j' \leq J$, $n = 1, \ldots, N_{i,j}$ and a specific $\nu \in M$ that*

(M1) *the random elements $P_{i,j,n}$ satisfy condition (3), in particular they are a.s. contained in a convex geodesic ball $U_{i,j}$ around the unique intrinsic mean $\mu_{i,j}$;*

(M2) *$P_{i,j,n}$ is independent of $P_{i',j',n'}$ for $(i,j,n) \neq (i',j',n')$;*

(M3) *all $\mu_{i,j}$ are contained in a neighborhood U around ν in which $\log_\nu$ is well defined, in particular the $X_{i,j,n}$ in (7) are well defined;*

(M4) *the distributions of $X_{i,j,n}$ and $X_{i',j',n'}$ agree.*

Note that (M2) and (M4) can be replaced by

(M4') *the $X_{i,j,n}$ are i.i.d.;*

this corresponds to the $\epsilon_{i,j,n}$ being i.i.d. in the classical setting (1). A generic random variable with the distribution of $X_{i,j,n}$ will be denoted by X. Since all $P_{i,j,n}$ have compact support, all moments of X exist; its covariance matrix will be denoted by Σ.

In the single-factorial model for manifolds *we assume additionally that for all $1 \leq i \leq I$ there are $\mu_i \in M$ such that*

$$\mu_{i,1} = \ldots, \mu_{i,J} = \mu_i$$

i.e. the second factor has no influence.

For the single factorial model for manifolds in case of only two levels $I = 2$, the specific location ν can be chosen arbitrarily on the geodesic segment connecting μ_1 with μ_2. For more levels the intrinsic mean of the entire population may be chosen.

In order to obtain a statistic for a test of hypothesis H_0 versus H_1 from Section 1 we investigate empirical covariances under H_0, i.e. under the single-factorial model.

Condition (M1) and hence (3) guarantees in case of the single-factorial model that the empirical group means $\overline{P}_i$ of the $P_{i,1,l}, \ldots, P_{i,1,N_{i1}}, P_{i,2,1}, \ldots, P_{i,J,N_{iJ}}$ are unique $(1 \leq i \leq I)$. Thus

$$Y_{i,j,n} \quad := \quad \theta_{\overline{P}_i,\nu} \circ \log_{\overline{P}_i} \circ P_{i,j,n} \quad (8)$$

are well defined under the single-factorial model. Note that $E(X_{i,j,n}) = 0 = \sum_{j=1}^{J} \sum_{n=1}^{N_{i,j}} Y_{i,j,n}$ as a consequence of (6). Letting

$$\overline{X}_{i,j,\cdot} := \frac{1}{N_{i,j}} \sum_{n=1}^{N_{i,j}} X_{i,j,n}, \quad \overline{Y}_{i,j,\cdot} := \frac{1}{N_{i,j}} \sum_{n=1}^{N_{i,j}} Y_{i,j,n}\,.$$

(here the index "${}_{i,j,\cdot}$" indicates the usual Euclidean mean) we let

$$\widehat{\Sigma}(Y)_{ij} \quad := \quad \frac{1}{N_{i,j}} \sum_{n=1}^{N_{i,j}} (Y_{i,j,n} - \overline{Y}_{i,j,\cdot})(Y_{i,j,n} - \overline{Y}_{i,j,\cdot})^T \text{ and}$$

$$\widehat{\Sigma}(X)_{ij} \quad := \quad \frac{1}{N_{i,j}} \sum_{n=1}^{N_{i,j}} (X_{i,j,n} - \overline{X}_{i,j,\cdot})(X_{i,j,n}^T - \overline{X}_{i,j,\cdot})^T$$

denote the corresponding sample covariance matrices.

Lemma 2.2. *Under the single-factorial model for manifolds, for every* $1 \le i \le I,\ 1 \le j \le J$, *and fixed* n, $1 \le n \le N_{i,j}$,

$$Y_{i,j,n} \to X_{i,j,n} \text{ a.s. and } \widehat{\Sigma}(Y)_{ij} \to \Sigma \text{ a.s.}$$

as $N_{i,j} \to \infty$.

Proof: By hypothesis, all second moments of components exist. Hence the SLLN (5) is applicable yielding $\overline{P}_i \to \mu_i = \mu_{i,1} = \ldots = \mu_{i,J}$ a.s. By continuity for fixed (i, j, n)

$$\begin{aligned} & Y_{i,j,n} - X_{i,j,n} \\ = \ & \theta_{\overline{P}_i,\nu}\big(\log_{\overline{P}_i}(P_{i,j,n})\big) - \theta_{\mu_i,\nu}\big(\log_{\mu_i}(P_{i,j,n})\big) \\ \to \ & 0 \text{ a.s.} \end{aligned}$$

yielding the first assertion. To see the second assertion consider for $v = \log_{\mu_i}(\overline{P}_i) \in \mathbb{R}^D$, $w = \log_\nu \overline{P}_i \in \mathbb{R}^D$ and $w_n = \log_{\mu_i}(P_{i,j,n}) \in \mathbb{R}^D$ the smooth function

$$\begin{aligned} f(v, w, w_n) &:= Y_{i,j,n} - X_{i,j,n} \\ &= a(w, w_n)v + O(\|v\|^2) \end{aligned}$$

with a suitable smooth function a (by Definition A.1, parallel transport is smooth). Under condition (M1), $\|w_n\| = d(\mu_i, P_{i,j,n})$ is a.s. uniformly bounded for all $1 \le n \le N_{i,j}$ and so is $\|w\|$. Hence there is a constant M independent of $N_{i,j}$ such that $\|X_{i,j,n}X^T_{i,j,n} - Y_{i,j,n}Y^T_{i,j,n}\| \le M\|v\|^2$ a.s. for all $1 \le n \le N_{i,j}$. Thus, with another constant $M' > 0$

$$\begin{aligned} \|\widehat{\Sigma}(Y)_{ij} - \widehat{\Sigma}(X)_{ij}\| &\le M'\|v\| \\ &= M'd(\overline{P}_{i,j}, \mu_{i,j}) \ \to \ 0 \ \text{ a.s.}. \end{aligned}$$

Further $\widehat{\Sigma}(X)_{ij} \to \Sigma$ a.s. This yields the assertion. □

Theorem 2.3. *Under the two-factorial model for manifolds we have for every* $1 \le i \le I,\ 1 \le j \le J$, *that*

$$\sqrt{N_{i,j}}(\Sigma^{-1/2}\widehat{\Sigma}(Y)_{ij}\Sigma^{-1/2} - I_D) \to \mathcal{H} \text{ in distribution}$$

as $N_{i,j} \to \infty$. *Here* I_D *denotes the* $D \times D$ *unit matrix and* $\mathcal{H}$ *denotes a Gaussian matrix with zero mean and covariance*

$$\mathrm{COV}(R^rR^s, R^tR^u), \quad R = (R^1, \ldots, R^D)^T := \Sigma^{-1/2}X_{i,j,n}$$

of the (r, s)*-th and the* (t, u)*-th entry.*

Proof: The asserted convergence for $\widehat{\Sigma}(X)_{ij}$ in place of $\widehat{\Sigma}(Y)_{ij}$ follows from the classical multivariate Central-Limit Theorem with the asymptotic covariance given in B.1 in the Appendix. The assertion follows for $\widehat{\Sigma}(Y)_{ij}$ as well, as a consequence of Lemma 2.2. □

3 MANOVA

In the preceeding section we devised a method how to map manifold data under a two-factorial model to data within a common Euclidean space. For the covariance of this Euclidean data we derived a common asymptotic distribution. Under the assumption that all $X_{i,j,n}$ from (7) are multivariate normally distributed and, even more boldly that all $Y_{i,j,n}$ from (8) are multivariate normally distributed, classical MANOVA could be applied to the $Y_{i,j,n}$. Obviously this assumption is far-fetched for our applications on manifolds. It is well-known, however, that the methods of classical MANOVA extend to non-normal data as well in a large number of cases (e.g. [36, p. 465]), one condition being a common asymptotic covariance distribution as in Theorem 2.3. For this reason we review below classical MANOVA for normal data, discuss its robustness to nonnormality and apply the method to Euclidean data obtained from manifold data as above.

3.1 Classical MANOVA Revisited

As a standard reference for MANOVA we refer to [37], [3] and [8]). Here the *Wishart distribution* $W_D(\Sigma, N)$ with N degrees of freedom and $D \times D$ covariance matrix Σ plays a prominent role. In case of $D = 1$ it is the well-known χ^2_N distribution.

Suppose that $Z_1, \ldots, Z_N$ are independently multivariate normal $N_D(0, \Sigma)$-distributed and let $W_n := Z_n - \overline{Z}.$ with $\overline{Z}. = \frac{1}{N}\sum_{n=1}^n Z_n$. Then

$$\begin{aligned} N\widehat{\Sigma}(Z) &:= \sum_{i=1}^n Z_nZ_n^T \ \sim \ W_D(\Sigma, N) \ \text{ and} \\ N\widehat{\Sigma}(W) &:= \sum_{n=1}^N W_nW_n^T \ \sim \ W_D(\Sigma, N-1) \end{aligned}$$

In particular if Σ is of rank D then

$$\sqrt{N}\left(\Sigma^{-1/2}\widehat{\Sigma}(W)\Sigma^{-1/2} - I_D\right) \ \to \ \mathcal{G} \tag{9}$$

in distribution, where I_D is the D-dimensional unit matrix and $\mathcal{G}$ is a Gaussian matrix with independent entries of variance 2 on the diagonal and of variance 1 elsewhere (cf. Theorem 2.3.).

Under the single-factorial model for manifolds, if all $Y_{i,j,n}$ from (8) are multivariate normally distributed ($1 \le i \le I,\ 1 \le j \le J$) then

$$\sum_{i=1}^I\sum_{j=1}^J N_{i,j}\widehat{\Sigma}(Y)_{ij} \ \sim \ W_D\left(\Sigma, \sum_{i=1}^I\sum_{j=1}^J(N_{i,j}-1)\right)$$

is independent of

$$\sum_{i=1}^I N_i\widehat{\Sigma}(Y)_i - \sum_{i=1}^I\sum_{j=1}^J N_{i,j}\widehat{\Sigma}(Y)_{ij} \ \sim \ W_D\Big(\Sigma, I(J-1)\Big).$$

Here, $N_i = \sum_{j=1}^J N_{i,j}$ and

$$\begin{aligned} \widehat{\Sigma}(Y)_i &:= \sum_{j=1}^J\sum_{n=1}^{N_{i,j}}(Y_{i,j,l} - \overline{Y}_{i,\cdot,\cdot})(Y_{i,j,l} - \overline{Y}_{i,\cdot,\cdot})^T \ \text{ with} \\ \overline{Y}_{i,\cdot,\cdot} &= \frac{1}{\sum_{j=1}^J N_{i,j}}\sum_{j=1}^J\sum_{n=1}^{N_{i,j}} Y_{i,j,l}\,. \end{aligned}$$

In consequence, several one-dimensional test statistics involving independent Wishart-distributed matrices have been proposed. In fact, there are essentially two different types of tests, one involving largest eigenvectors and one involving determinants i.e. geometric means of eigenvectors. We prefer the latter as it is more robust to nonnormality (cf. [36, 465]). With *Wilks' Lambda* distribution,

$$\frac{\det\left(\sum_{i=1}^{I}\sum_{j=1}^{J} N_{i,j}\widehat{\Sigma}(Y)_{ij}\right)}{\det\left(\sum_{i=1}^{I} N_i\widehat{\Sigma}(Y)_i\right)} \sim \Lambda\left(D, \sum_{i=1}^{I}\sum_{j=1}^{J}(N_{i,j}-1), I(J-1)\right) \quad (10)$$

for $I, J = 2$, the identity

$$G_{\Lambda(D,N,2)}(\lambda) = 1 - G_{F_{2D,2(N-D+1)}}\left(\frac{N-D+1}{D}\,\frac{1-\sqrt{\lambda}}{\sqrt{\lambda}}\right) \quad (11)$$

can be employed (here $F_{n,m}$ denotes the F-distribution with n and m degrees of freedom), and in the general case for sufficiently large N, *Bartlett's approximation*

$$G_{\Lambda(D,N,K)}(\lambda) \approx 1 - G_{\chi^2_{NK}}\left(-\left(N - \frac{1}{2}(D-K+1)\right)\log\lambda\right) \quad (12)$$

can be used. In both formulas, G denotes the respective cumulative distribution function (cf. [37, p. 83-4]).

3.2 MANOVA on Manifolds

Let us first briefly discuss effects of nonnormality for the benefit of readers less versed in asymptotic statistics. For nonnormal deviates $Y_{i,j,n}$ as is realistic in the one- and two-factorial models for manifolds – note that we have not made any distributional assumption in (M4) – the covariances of the limit distributions of Theorem 2.3 and (9) disagree in general. For this reason, $\widehat{\Sigma}(Y)_{ij}$ and $\widehat{\Sigma}(Y)_i$ are not well approximated by the corresponding Wishart distribution. This is the same phenomenon as occurs in the one-dimensional χ^2-test under nonnormality. If the respective covariances as in Theorem 2.3 agree with one another then the ratio of the respective determinants in (10), however, is usually well approximated by the corresponding Λ statistics, even for relatively small values of D, $N_{i,j}$, I and J. Note that this corresponds to a univariate F-test with higher degrees of freedom in case of two levels for each factor. Contrary to the χ^2-statistics, F-statistics are again robust under nonnormality. The left image of Figure 1 illustrates that the Λ-statistics is indeed almost uniformly distributed in the one-effect model for relatively small dimensions D and group sizes $N_{i,j}$.

Hence for general data $Y_{i,j,n}$ under the two-factorial model for manifolds we may use (10) in approximation under the single-factorial hypothesis. In case of $I(J-1) = 2$ and large $N = \sum_{i,j} N_{i,j} - I$ the exact distribution (11) can be used, for $I(J-1) > 2$ Bartlett's approximation (12) will do fine. Indeed, the right image of Figure 1 illustrates that Bartlett's approximation is valid already for comparatively small group sizes.

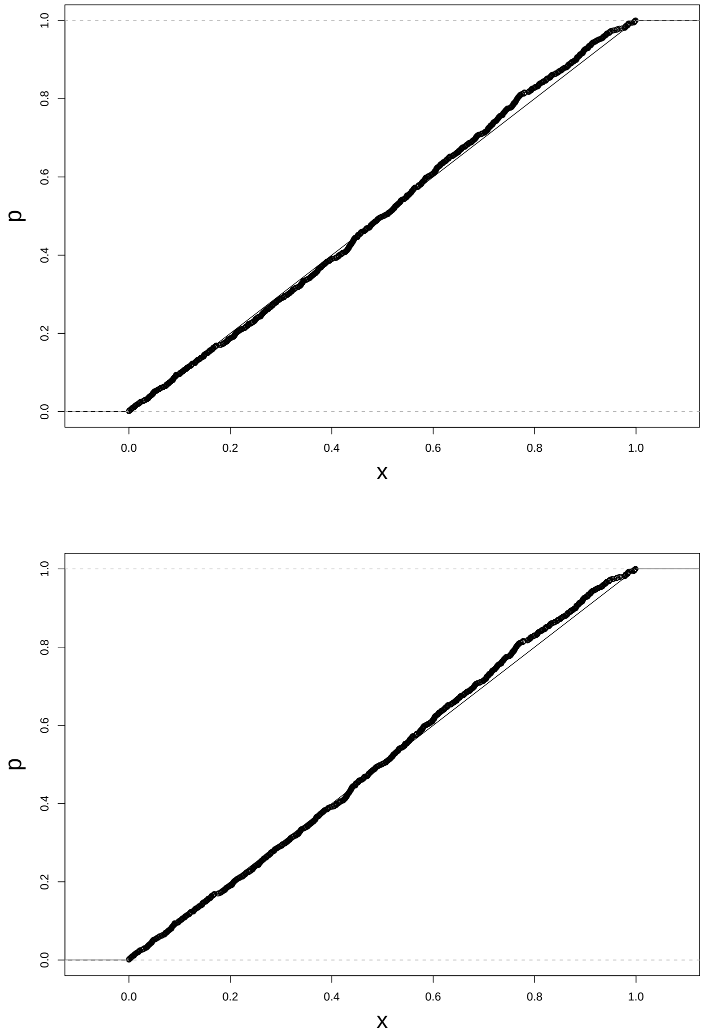


Fig. 1. *Robustness of Wilks' Lambda-test for 2 way MANOVA under nonnormality: empirical test statistics (10) generated with $1'000$ repetitions for exponential deviates under the hypothesis of only one single effect, $D = 4$ and $N_{i,j} = 5$ for $1 \le i, j \le 2 = I = J$. Top: the distribution of the exact statistics (11). Bottom: the distribution of Bartlett's approximation (12).*

4 Application: Intrinsic Two-Way MANOVA

In collaboration with the Institute for Forest Biometry and Informatics at the University of Göttingen, the influence on the shape of tree leaves of their vertical position is studied for different tree genotypes. The dataset studied here is based on digitized 2D-contours of Canadian black poplar leaves taken from a recent plantation of three clones ($i = 1$) and a reference tree ($i = 2$) at two different heights. Since the trees are young and of little height, level $j = 1$ corresponds to approx. 1 meter and level $j = 2$ to approx. 1.5 meters. The left image of Figure 2 shows some typical original leaves.

From sight, experts in forestry can discriminate rather well between shapes of leaves of the different genotypes. Shape variation due to different height levels, however, can hardly be visually detected. As is the case for other plants, there are grounds to believe that height levels affect leaf shapes as well, e.g. [11]. In fact as we see in the following, such effects can be identified by intrinsic MANOVA using Kendall's space Σ_2^4 of planar shapes with a "minimal" number of 4 landmarks. Kendall's shape spaces are introduced in Appendix A.2.

For each leaf, two anatomical landmarks have been placed at the bottom and top of the main leaf vain and two more mathematical landmarks at the largest extent of the leaf, orthogonal to the dominating direction of the main leaf vain. The right image of Figure 2 shows a typical digitized contour and its quadrangular configuration. For the two-factorial model, we thus obtain a sample $P_{i,j,n}$ from four groups in Σ_2^4, $1 \leq i, j \leq 2$, $1 \leq n \leq N_{i,j}$ with

$$N_{1,1} = 22, \quad N_{1,2} = 51, \quad N_{2,1} = 13, \quad N_{2,2} = 18,$$
$$\sum_{1 \leq i,j \leq 2} N_{i,j} = 104\,.$$

As introduced above, j refers to genotype and i to height level.

Indeed, geodesic principal component analysis (as has been introduced for Kendall's shape spaces of planar configurations in [21]) as well as a Hotelling T^2-test based on [10, Chapter 7] for the data mapped to the tangent space at their common intrinsic sample mean under the inverse exponential, both endorse the finding that leaf shapes vary between different genotypes with high significance (p-value of 0.002). Figure 3 depicts the geodesic scores by projection of the entire data to the first two principal component geodesics. They explain $83\ \%$ of data variation. Obviously, height levels cannot easily be discriminated in Figure 3.

In order to see whether different height levels affect leaf shape as well, intrinsic two-way MANOVA has been employed as introduced in Section 2. Note (using Remark A.7) that the parallel transport θ is well defined on all Σ_2^k except for a set of measure zero. Thus, for our experimental situation parallel transport is well defined. Conforming to the notation of Section 2.2 we chose $\nu = \overline{P}_1$ and considered the data

$$Y_{i,j,n} := \theta_{\overline{P}_i,\nu}\left(\log_{\overline{P}_i} P_{i,j,n}\right)$$

with $\overline{P}_i$ the intrinsic sample mean of the $P_{i,j,l}$ with $1 \leq n \leq N_{i,j}$ and $1 \leq j \leq 2$ for $1 \leq i \leq 2$. For numerical feasibility we lifted the data to the respective horizontal space of the pre-shape sphere S_2^4. Then the explicit formula for the horizontal lift of the bottom space parallel transport θ from Theorem A.6 can be used. We note that the likelihood ratio test of [15] for equality of the respective four covariance matrices under normality assumptions lead to a p-value of 0.146 thus giving no evidence for unequal covariance matrices.

Computing the test-statistics (10) via (11) and in approximation (12) we obtained a p-value of 0.022 in both cases. Hence, the null hypothesis that there are only genotype effects but no level effects can be rejected with high significance.

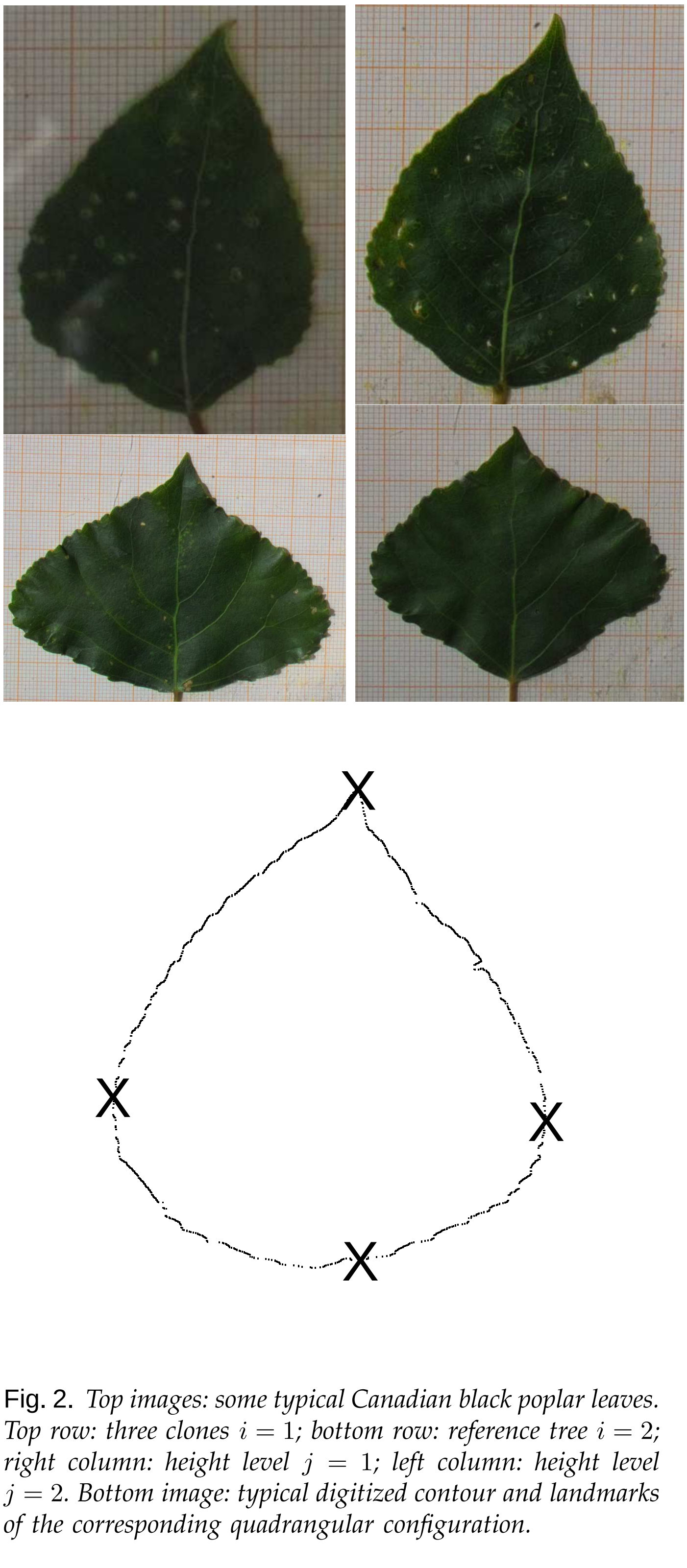


Fig. 2. *Top images: some typical Canadian black poplar leaves. Top row: three clones $i = 1$; bottom row: reference tree $i = 2$; right column: height level $j = 1$; left column: height level $j = 2$. Bottom image: typical digitized contour and landmarks of the corresponding quadrangular configuration.*

5 Conclusion

In this paper we proposed an extension of classical MANOVA by a generalization of certain linear models to

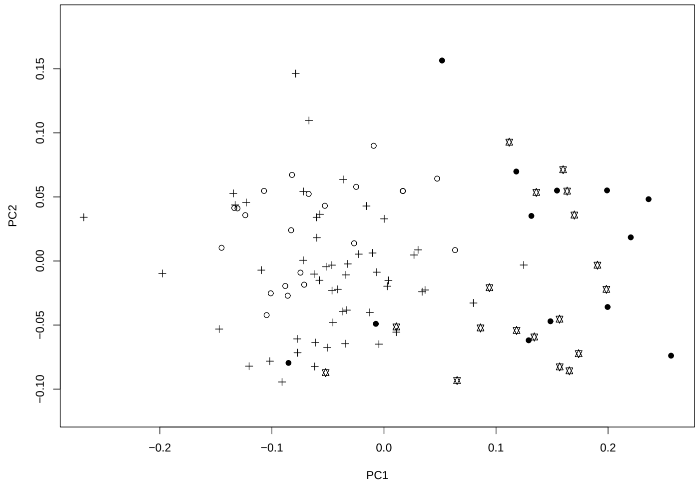


Fig. 3. *Geodesic scores of quadrangular shapes of poplar leaves. Level $j = 1$ for clones $i = 1$ as circles, level $j = 2$ for clones $i = 1$ as crosses. Level $j = 1$ for reference tree $i = 2$ as filled circles, level $j = 2$ for reference tree $i = 2$ as stars.*

non-linear Riemannian manifolds. To date only a single-factorial model with a corresponding Hotelling T^2-test or Goodall's F-test has been available. Cf. [10, Chapter 7] as well as [17] for a tangent space approximation at a single extrinsic mean and [4] for using a single extrinsic or intrinsic mean. We note that we also weakened the assumption of isotropy in Goodall's F-test for one-way MANOVA as suggested in [10, Chapter 7]. Nonetheless, the assumption of equal covariances of the groups still is debatable and calls for further research. For our data at hand, neither normality nor unequality could be rejected. In general under nonnormality, tests based on [16] could be employed.

Our newly introduced models allow to test for the effects of multiple factors by comparing images in distinct tangent spaces centered at different intrinsic means under parallel transport. In case the Riemannian manifold is only implicitly given as is the case for Kendall's similarity shape spaces, we provided for a method to pull back parallel transport of the bottom space to the top space in Corollary A.4. This we computed explictly for Kendall's spaces of planar shapes. We illustrated the use of this new method by an intrinsic 2-way MANOVA for objects of forest biometry with two factors. Effects of genotype that are visible to the trained eye of the scientist can be identified with existing methodology. Effects of the second factor, otherwise not accessible, were identified by intrinsic MANOVA.

Recall that due to the non-linearity of the underlying manifold, a decomposition of effects acting separately can neither be modelled nor expected. It seems that for this very reason, biologists cannot geometrically identify the second effect.

At this point we note a complication for models involving more than 2 levels arising from the fact that parallel transport is not transitive if curvature is present. One may work-around by choosing the specific location ν as a population mean. Alternatively, one may choose a specifically distinguished shape, for example the shape of a regular polygon. The hypothesis of equally distributed parallelly mapped tangent space residuals, however, will in general not be valid in both tangent spaces. However, if only two levels are involved, it is comforting to know that the results obtained are independent of ν if chosen arbitrarily on the geodesic between the two respective means.

For highly concentrated data in low curvature regions of shape space, one might approximate by mapping all data into a single tangent space and perform classical MANOVA there. E.g. this might be achieved by flattening the space using the Riemann exponential, cf. [41]. For high curvature present or larger data spread, however, results obtained by extrinsic analysis may deviate considerably from results obtained by intrinsic analysis, cf. Examples 1 and 2 of [22] with an extended discussion of the validity of extrinsic approximations. In the specific data example considered, a considerable data spread is visible in Figure 3: the two tree-means are further apart than $10\,\%$ of the maximal distance of $\pi/2$ in the shape space Σ_2^m.

In principle, our method is applicable to all shape spaces of the form $\Sigma = \mathcal{S}/G$ where G is a Lie group of shape invariants acting on a Riemannian *pre-shape space* $\mathcal{S}$ immersed in a Euclidean space. We note that this quotient is locally trivial in the statistical analysis of *projective shapes* if modelled as direct products of real projective spaces, cf. e. g. [39]. Hence with the parallel transport on spheres as in Example A.2, our method of intrinsic MANOVA is available also for projective shape analysis. It can also be applied directly to the *spherical shape space* of star-shaped pre-aligned configurations, cf. [9] and [18].

With more effort, this method could be applied to Grassmanian manifolds as are used in the statistical analysis of *affine shapes*, cf. [2] as well as [38]; or to *medial axes shape spaces* as have been introduced by [7] and are currently of high interest, e.g. [13] as well as [14].

In general, it may be difficult to compute bottom space parallel transport. For example in case of Kendall's space of 3 and higher-dimensional shapes, no vertical geodesics are available in general (cf. [22, Example 5.1]) thus making the evaluation of the l.h.s. of (17) difficult. Also for infinite dimensional shape spaces such as the *shape spaces of closed contours* (cf. [28]), if the vertical spaces of the respective submersions are finite dimensional, the differential equation (17) could be solved numerically thus making intrinsic MANOVA available.

In conclusion we note that parallel transport of covariances is a mathematically natural way to extend MANOVA to manifolds. Parallel transport of covariances seems in particular natural for *locally symmetric spaces* – those are spaces that allow local isometries that reverse geodesics. Then, sectional curvature is invariant under parallel transport, cf. [29, pp.300-304]. The spaces considered in this work, namely spheres and Kendall's space of planar shapes are locally symmetric ([29, pp.275]). Also for our application in forest biometry, the assumption of invariance of the empirical covariances under parallel transport appears reasonable. We note that [14] use parallel transport of covariance matrices as well, in order to define a Mahalanobis distance on shape manifolds. Clearly, a discussion of this assumption for applications in general is beyond the scope of this paper. Recall the "Geodesic Hypothesis" by [34] (cf. also [18]) which states that natural biological growth tends to occur along geodesics in shape space: this hypothesis as well as the assumption of parallel transport of covariances in effect link a specific Riemannian geometry to biological shape. This topic certainly deserves future research.

APPENDIX A
PARALLEL TRANSPORT

In this section we review basic concepts of Riemannian geometry found in any standard textbook (specifically [35] is very appropriate for the following), in particular formulae relating covariant derivatives of Riemannian immersions and submersion. These provide a differential equation lifting the parallel transport on shape space to Euclidean space. In fact, the projection of ambient Euclidean space to a sphere, as well as the quotient mapping from the *pre-shape sphere* to *Kendall's space of planar shapes* are an example of a Riemannian immersion followed by a Riemannian submersion. It is the aim of this section to lift parallel transport from Kendall's space to the pre-shape sphere and from there to Euclidean space.

For a Riemannian D-dimensional manifold M denote by $\langle V_p, W_p\rangle^M$ the *Riemannian metric* of tangent spaces, by $d^M : M \times M \to [0,\infty)$ the induced distance on M and by $\nabla^M_V W$ the *covariant derivative* of vector-fields. Here $V, W \in T(M)$ denote vector-fields with values V_p, W_p in the tangent space T_pM of M at $p \in M$. A vector-field $W \in T(M)$ is *parallel* along a smooth curve $t \to \gamma(t)$ on M if

$$\nabla_{\dot\gamma} W \quad = \quad 0\,. \tag{13}$$

Subsequent arguments exploit the fact that the *parallel equation* (13) written in local coordinates is a system of D first order differential equations for W. It is well-known that there is locally a unique solution along γ for a given initial value. In Euclidean space the left hand side of the system has the simple form (14). In particular, *geodesics* are characterized by the fact that their velocity is parallel: $\nabla_{\dot\gamma}\dot\gamma = 0$.

The covariant derivative is also called the *Levi-Civitá connection*. Indeed, if two offsets $p, p' \in M$ can be joined by a unique geodesic segment of minimal length (in particular, this is the case if p' is sufficiently close to p), their respective tangent spaces can be *connected* via *parallel transport*.

Definition A.1. *$w' \in T_{p'}M$ is the* parallel transplant *$\theta_{p,p'}(w)$ of $w \in T_pM$ if there are*
1) *a unique unit speed geodesic $t \to \gamma(t)$ connecting $p = \gamma(0)$ with $p' = \gamma\big(d^M(p,p')\big)$, and*
2) *a vector field $W \in T(M)$ parallel along γ with $W_p = w, W_{p'} = w'$.*

The Euclidean spaces $\mathbb{R}^n$ can be identified with all of their tangent spaces, i.e. $\langle v, w\rangle := \langle v, w\rangle^{\mathbb{R}^n} = \sum_{i=1}^n v^i w^i$, $\|x\|^2 := \langle x, x\rangle$ and the covariant derivative is just the usual multivariate derivative by components,

$$\nabla^{\mathbb{R}^n}_{(v^1,\dots,v^n)}(w^1,\dots,w^n) = \sum_{i=1}^n v^i \left(\frac{\partial w^1}{\partial x^i},\dots,\frac{\partial w^n}{\partial x^i}\right).$$

In particular, if $v = \dot x(t)$, i.e. $\frac{dx^i}{dt} = v^i$ we have that

$$\nabla^{\mathbb{R}^n}_{\dot x(t)} W \quad = \quad \frac{d}{dt} W_{x(t)}\,. \tag{14}$$

Thus as desired, parallel transport on Euclidean spaces is given by affine translations.

For short we write $W(t)$ for the value $W_\gamma(t)$ along a selfunderstood smooth curve $t \to \gamma(t)$ and

$$\dot W(t) \;:=\; \frac{d}{dt} W_{\gamma(t)}$$

in the Euclidean case.

A.1 Riemannian Immersions and Submersions

A smooth mapping $\Phi : M \to N$ of Riemannian manifolds M and N induces a differential mapping $d\Phi_p : T_pM \to T_{\Phi(p)}N$ of tangent spaces. Φ is called

an *immersion* if both Φ and $d\Phi$ are injective,
a *submersion* if both Φ and $d\Phi$ are surjective,
an *isometry* if $\langle V_p, W_p\rangle^M = \langle d\Phi_p V_p, d\Phi_p W_p\rangle^N$, $\forall p \in M$ and $V, W \in T(M)$.

An isometric immersion (submersion) is a *Riemannian immersion (submersion)* respectively. As a consequence of the implicit function theorem, every Riemannian immersion admits locally an orthogonal projection $\Psi : U \cap N \to V \cap M$ which is in general a non-Riemannian submersion. Here, $U \subset M$ and $V \subset N$ are suitable open sets. In consequence for a Riemannian immersion $\Phi : M \to N$, $X, Y \in T(M)$ and arbitrary local extensions $\widetilde X, \widetilde Y \in T(N \cap U)$ of $d\Phi X, d\Phi Y \in T(\Phi(M \cap V)$ we have

$$d\Psi\left(\nabla^N_{\widetilde X}\widetilde Y\right) \quad = \quad \nabla^M_X Y\,. \tag{15}$$

Example A.2 (Parallel Transport on Spheres)**.** *The immersion of the unit-hypersphere $M = S^{n-1}$ in the Euclidean space $N = \mathbb{R}^n \setminus \{0\}$ can be viewed as a global projection $\Psi : \mathbb{R}^n \to S^{n-1}, x \mapsto \Psi(x) = \frac{x}{\|x\|}$. Obviously, this submersion is not isometric. Identifying T_xS^{n-1} with $\{v \in \mathbb{R}^n : \langle x, v\rangle = 0\}$ we have that $\widetilde W = W$ for all $W \in T(S^{n-1})$. If $t \to \gamma(t)$ is* *a smooth curve on S^{n-1}, then (15) yields with (14) at once that*

$$\nabla^{S^{n-1}}_{\dot\gamma(t)} W(t) \;=\; d\Psi_{\gamma(t)}\big(\dot W(t)\big) \;=\; \dot W(t) - \langle \dot W(t), \gamma(t)\rangle\, \gamma(t)\,.$$

Solving the parallel equation (13) for $W(t) = \dot\gamma(t)$ yields that every unit speed geodesic on S^{n-1} is a great circle of form $\gamma_{x,v} : t \to x\cos t + v \sin t$, $x, v \in S^{n-1}$, $\langle x, v\rangle = 0$. Solving it for arbitrary $W(t) \in T(S^{n-1})$ with initial condition $W(0) = w \perp x$ along the geodesic $\gamma = \gamma_{x,v}$ we obtain that

$$W(t) \;=\; w - \langle w, v\rangle\, v + \langle w, v\rangle\, \dot\gamma(t)\,.$$

Geometrically speaking, the part of w orthogonal to the geodesic $\gamma_{x,v}$ is mapped under affine translation and its orthogonal complement is transported just as velocity (which is of course parallel). In conclusion, we have: $\log_x$ is well defined on all $S^{n-1} \setminus \{-x\}$ and for $x' \neq \pm x$,

$$w' \;=\; w - \langle w, v\rangle \Big((1 - \langle x, x'\rangle)\, v - \langle v, x'\rangle\, x \Big)$$

is the parallel transplant $\theta_{x,x'}(w)$ of $w \in T_x S^{n-1}$ to $T_{x'}S^{n-1}$ where

$$v \;=\; \frac{x' - \langle x, x'\rangle\, x}{\|x' - \langle x, x'\rangle\, x\|}\,.$$

Next, consider a Riemannian submersion $\Phi : M \to N$. M is called the *top space* and N the *bottom space*. Every fiber $\Phi^{-1}(q)$, $q \in N$ is a submanifold of M that is locally a topological embedding. With the *vertical space* $T_p\Phi^{-1}\big(\Phi(p)\big)$ along the submanifold we have the orthogonal tangent space decomposition

$$T_pM = T_p\Phi^{-1}\big(\Phi(p)\big) \oplus H_pM\,.$$

The orthogonal complement H_pM is the *horizontal space*. Since $H_pM \cong T_{\Phi(p)}N$, every $V \in T(N)$ has a unique horizontal lift $\widetilde V \in H_pM$ characterized by $d\Phi\widetilde V = V$. For arbitrary $W \in T(M)$ denote by $\cdot^\perp : p \to W_p^\perp$ the orthogonal projection to the vertical space.

The following Theorem due to [40] allows to lift bottom space parallel transport to the top space. In addition to (15) this provides the vertical part as well which in general is non-zero for submersions.

Theorem A.3. *Let $\Phi : M \to N$ be a Riemannian submersion and let $X, Y \in T(N)$. Then we have with the* Lie bracket *$[\cdot,\cdot]$ on M that*

$$\nabla^M_{\widetilde X}\widetilde Y \;=\; \widetilde{\nabla^N_X Y} + \frac{1}{2}\,[\widetilde X, \widetilde Y]^\perp\,. \tag{16}$$

As an immediate consequence of Theorem A.3 and the parallel equation (13) we have the following Corollary.

Corollary A.4. *Let $\Phi : M \to N$ be a Riemannian submersion, $t \to \gamma(t)$ a horizontal geodesic on M and $W(t)$ a horizontal vector-field along γ. Then $d\Phi W$ is parallel along $\delta = \Phi \circ \gamma$ if and only if*

$$\nabla^M_{\dot\gamma(t)} W(t) \;=\; \frac{1}{2}\big[\dot\gamma(t), W(t)\big]^\perp\,. \tag{17}$$

Using (17) with the derivative $W(t)$ of a geodesic in bottom space, observe that this geodesic lifts to a horizontal geodesic in top space. In fact every top space geodesic with horizontal initial velocity has horizontal velocity throughout its course. Such geodesics are called *horizontal geodesics*. Every bottom space geodesic is a projection of a horizontal top space geodesic.

A.2 Parallel Transport on Kendall's Space of Planar Shapes

Kendall's landmark based similarity shape analysis is based on *configurations* consisting of $k \geq m+1$ labelled vertices in $\mathbb{R}^m$ called *landmarks* that do not all coincide. A configuration

$$x = (x^1, \ldots, x^k) = (x^{ij})_{1 \leq i \leq m, 1 \leq j \leq k}$$

is thus an element of the space $M(m,k)$ of matrices with k columns, each an m-dimensional landmark vector. Disregarding center and size, these configurations are mapped to the *pre-shape sphere*

$$M = S^k_m := \{p \in M(m, k-1) \colon \|p\| = 1\}\,,$$

where $\|p\|^2 = \langle p, p\rangle$ and $\langle p, v\rangle := \mathrm{tr}(pv^T)$ is the standard Euclidean product. This can be done by, say, multiplying by a sub-Helmert matrix, cf. [10] for a detailed discussion of this and other normalization methods. The Euclidean metric of $M(m, k-1)$ and the spherical metric of S^k_m are related as in Example A.2.

In order to filter out rotation information define on S^k_m a smooth action of the *special orthogonal group $SO(m)$* by the usual matrix multiplication $S^k_m \xrightarrow{g} S^k_m : p \mapsto gp$ for $g \in SO(m)$. The orbit $\pi(p) = \{gp\colon g \in SO(m)\}$ is the Kendall *shape* of $p \in S^k_m$ and the quotient

$$\pi : S^k_m \;\to\; \Sigma^k_m \;:=\; S^k_m / SO(m) \tag{18}$$

is called Kendall's *shape space*. In case of $m = 2$, the action of $SO(2)$ on $M(2, k-1)$ may be identified with the scalar action of

$$S^1 := \{e^{i\alpha} : \alpha \in [0, 2\pi)\}$$

on $\mathbb{C}^{k-1}$ and we have the embedding

$$S^k_2 \cong S^{2(k-1)-1} \;:=\; \{z = (z_1, \ldots, z_{k-1}) \in \mathbb{C}^{k-1} : \|z\|^2 = 1\}.$$

Hence, for $m = 2$, (18) defines a Riemannian submersion which is equivalent to the well-known *Hopf fibration*, mapping a unit sphere to *complex projective space*

$$S^k_2 \to \Sigma^k_2 \;\cong\; S^{2(k-1)-1}/S^1 \;=\; P\mathbb{C}^{k-2}\,.$$

For ease of notation, we use complex notation henceforth exclusively.

Contrary to Example A.2, complex projective space is not explicitly available. Hence, for computational feasibility, parallel transport in the bottom space $P\mathbb{C}^{k-2}$ needs to be pulled back to the top space $S^{2(k-1)-1}$. For this task we review additional background (cf. [21]).

For $z, z' \in S^k_2$, we say that $e^{i\alpha}z'$ is z' rotated into *optimal position* to z if

$$d^{S^m_2}(z, e^{i\alpha}z') = d^{\Sigma^k_2}\big(\pi(z), \pi(z')\big)\,.$$

Since S^1 is compact, every point can be rotated into optimal position to a given point. We note that the relation "in optimal position" is reflexive but not in general transitive (cf. [42, p. 602]). Optimal positioning is unique if $d^{\Sigma_2^k}\big(\pi(z),\pi(z')\big) < \frac{\pi}{2}$, cf. [26, p.121]. If $z \neq z'$ are in optimal position (in particular, $z' \neq -z$) then

$$v = \frac{z' - \langle z', z\rangle\, z}{\|z' - \langle z', z\rangle\, z\|} \tag{19}$$

is horizontal at z and the (thus horizontal) geodesic $\gamma_{z,v} : t \to z\cos t + v\sin t$ is the unique unit speed geodesic joining $z = \gamma_{z,v}(0)$ with $z' = \gamma_{z,v}\big(d^{S_2^m}(z,z')\big)$. Moreover, iz spans the vertical space at z, in particular $i\gamma_{z,v}$ is the *vertical field* along $\gamma_{z,v}$. In real coordinates we have

$$i\gamma_{z,v} = \sum_{j=1}^{k-1}\big(\gamma^{1,\alpha}\partial_{2,\alpha} - \gamma^{2,j}\partial_{1,j}\big)$$

which yields

$$d\omega = 2\sum_{j=1}^{k-1} d^{1,j}\wedge d^{2,j}$$

for the exterior derivative of the one-form ω dual to the vertical field. As a consequence of the general property

$$\begin{aligned} &d\omega(X,Y)\\ &= 2\sum_{j=1}^{k-1} d^{1,j}\wedge d^{2,j}\Big(\sum_{j'=1}^{k-1}\big(X^{1,j'}\partial_{1,j'} + X^{2,j'}\partial_{2,j'}\big),\\ &\qquad \sum_{j''=1}^{k-1}\big(Y^{1,j''}\partial_{1,j''} + Y^{2,j''}\partial_{2,j''}\big)\Big)\\ &= \sum_{j=1}^{k-1}\big(X^{1,j}Y^{2,j} - X^{2,j}Y^{1,j}\big)\\ &= \langle X, iY\rangle \end{aligned}$$

we have the following Lemma:

Lemma A.5. *$d\omega(X,X) = 0 = d(X,Y)$ and $d\omega(X,iX) = \|X\|^2$ for arbitrary $X, Y \in T(S_2^k)$ with $X \perp iY$.*

Theorem A.6. *Let $p, p' \in \Sigma_2^k$, $0 < d^{\Sigma_2^k}(p,p') < \frac{\pi}{2}$. Then with the notation below,*

$$w' = d\pi_{z'}\big(W(\arccos\langle z,z'\rangle)\big)$$

is the well defined parallel transplant of $w \in T_p\Sigma_2^k$ to $T_{p'}\Sigma_2^k$. In particular, there are $z, z' \in S_2^k$ in optimal position, $\pi(z) = p$, $\pi(z') = p'$ and a unique unit-speed geodesic $\gamma_{z,v} : t \to z\cos t + v\sin t$ connecting $z = \gamma_{z,v}(0)$ with $z' = \gamma_{z,v}\big(d^{\Sigma_2^k}(p,p')\big)$ which is horizontal and

$$\begin{aligned} &W(t) =\\ &\widetilde{w} - \Big(\langle\widetilde{w},v\rangle\, v + \langle\widetilde{w},iv\rangle\, iv\Big) + \langle\widetilde{w},v\rangle\,\dot\gamma_{z,v}(t) + \langle\widetilde{w},iv\rangle\, i\dot\gamma_{z,v}(t) \end{aligned}$$

is the horizontal lift at $\gamma_{z,v}(t)$ of the parallel transplant of w along the unique geodesic $\pi\circ\gamma_{z,v}$ of minimal length connecting p with p'. Here, v is from (19) and $\widetilde{w}$ denotes the horizontal lift of w at z.

Proof: With the preceeding, all we need to show is that $W(t)$ defined above satisfies equation (17). This is a consequence of the following three identities:

$$\begin{aligned} \nabla^{S_2^k}_{\dot\gamma_{z,v}(t)} W(t) &= \dot W(t) - \big\langle \dot W(t), \gamma_{z,v}(t)\big\rangle\, \gamma_{z,v}(t)\\ &= -\langle\widetilde{w}, iv\rangle\, i\gamma_{z,v}(t) \end{aligned}$$

from Example A.2,

$$\big[\dot\gamma_{z,v}(t), W(t)\big]^{\perp} = \Big\langle i\gamma_{z,v}(t), \big[\dot\gamma_{z,v}(t), W(t)\big]\Big\rangle\, i\gamma_{z,v}(t)$$

with the vertical field, and

$$\begin{aligned} &\Big\langle i\gamma_{z,v}(t), \big[\dot\gamma_{z,v}(t), W(t)\big]\Big\rangle\\ &= \dot\gamma_{z,v}(t)\Big(\underbrace{\big\langle W(t), i\gamma_{z,v}(t)\big\rangle}_{=0}\Big)\\ &\qquad - W(t)\Big(\underbrace{\big\langle i\gamma_{z,v}(t), \dot\gamma_{z,v}(t)\big\rangle}_{=0}\Big) - 2d\omega\big(\dot\gamma_{z,v}(t), W(t)\big)\\ &= -2\langle\widetilde{w}, iv\rangle \end{aligned}$$

from Lemma A.5. Here we used the well-known (e.g. [29, p.36])

$$\omega[X,Y] = X\big(\omega(Y)\big) - Y\big(\omega(X)\big) - 2d\omega(X,Y)\,.$$

□

Geometrically speaking, the horizontal lift of bottom space parallel transport along a horizontal geodesic is obtained by top space parallel transport of the horizontal part (cf. Example A.2), and by mapping the vertical part along the vertical field with constant modulus. A transformation giving negative parallel transport for the velocity of the geodesic and the identity on the orthogonal complement is given in Lemma 3 of [1].

Remark A.7. *Since the largest possible distance between shapes on Σ_2^k is $\pi/2$, parallel transport is well defined on Σ_2^k except for a set of measure zero.*

Appendix B
Computation of Covariance

For random real variates $X_1,\dots,X_n,Y_1,\dots,Y_n$ let as usual $\overline{X}. = \frac{1}{n}\sum_{j=1}^n X_j$ and

$$S_{X,Y} = \frac{1}{(n-1)}\sum_{j=1}^{n}(X_i - \overline{X}.)(Y_i - \overline{Y}.)\,.$$

The proof of the following Lemma is tedious but straightforward.

Lemma B.1. *Let $(X_1,Y_1,W_1,Z_1),\dots,(X_n,Y_n,W_n,Z_n)$ i.i.d. $\sim (X,Y,W,Z)$ and suppose that all fourth moments $E(X^{\alpha_X}Y^{\alpha_Y}W^{\alpha_W}Z^{\alpha_Z})$ exist for $\alpha_X+\alpha_Y+\alpha_W+\alpha_Z = 4$. Then*

$$\begin{aligned} &\mathrm{COV}(S_{X,Y}, S_{W,Z}) =\\ &\frac{1}{n}\,\mathrm{COV}\Big((X-EX)(Y-EY),(W-EW)(Z-EZ)\Big)\\ &+\frac{1}{n(n-1)}\Big(\mathrm{COV}(X,W)\,\mathrm{COV}(Y,Z)\\ &\qquad + \mathrm{COV}(X,Z)\,\mathrm{COV}(Y,W)\Big)\,. \end{aligned}$$

ACKNOWLEDGMENTS

The authors would like to thank their colleagues Branislav Sloboda and Michael Henke from the Institute for Forest Biometry and Informatics at the University of Göttingen who provided us with the data set and whose interest in leaf shapes prompted this research. Also many thanks go to Vic Patrangenaru and Alexander Lytchak for discussing differential geometric aspects. The authors appreciate the editor's and reviewers' comments which helped in further improving this paper. The first author gratefully acknowledges support by DFG grant MU 1230/10-1. The second author gratefully acknowledges support by the German Federal Ministry of Education and Research, Grant 03MUPAH6 and SFB 803 and the third author for FOR 916. Also, the authors would like to thank the three anonymous referees for their helpful comments and suggestions.

REFERENCES

[1] G.J.A. Amaral, I.L. Dryden, A.T.A. Wood,*Pivotal Bootstrap Methods for k-Sample Problems in Directional Statistics and Shape Analysis*, Journal of the American Statistical Association (102) 478, 695–707, 2007

[2] R.V. Ambartzumian, *Factorization, Calculus and Geometric Probability*, Cambridge University Press, 1990

[3] T.W. Anderson, *An Introduction to Multivariate Analysis*, second ed., Wiley, New York, 1984

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

[5] N. Bissantz, L. Dümbgen, L., H. Holzmann and A. Munk, *Nonparametric confidence bands in deconvolution density estimation*, J. Royal Statist. Society Ser. B. (69), 483–506, 2007

[6] N. Bissantz, H. Holzmann and M. Pawlak, *Testing for Image Symmetries – with Application to Confocal Microscopy*, IEEE Transactions on Information Theory, to appear, 2009

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

[8] T. F. Cox, *An Introduction to Multivariate Data Analysis*, Hodder Arnold, London, 2005

[9] I.L. Dryden, *Statistical analysis on high-dimensional spheres and shape spaces*, Ann. Statist (33), 1643–1665, 2005

[10] I.L. Dryden and K.V. Mardia, *Statistical Shape Analysis*, Wiley, 1998

[11] J.R. England and P.M. Attiwill, *Changes in leaf morphology and anatomy with tree age and height in the broadleaved evergreen species, Eucalyptus regnans F. Muell*, Journal Trees – Structure and Function (20) 1, 79–90, 2006

[12] K. Evans, I.L. Dryden and H. Le, *Shape Curves and Geodesic Modelling*, manuscript, `http://www.maths.nottingham.ac.uk/personal/ild/papers/curves2.pdf`

[13] P.T. Fletcher and S.C. Joshi, *Principal Geodesic Analysis on Symmetric Spaces: Statistics of Diffusion Tensors*, ECCV Workshops CVAMIA and MMBIA, 87–98, 2004

[14] M. Fuchs and O. Scherzer, *Regularized Reconstruction of Shapes with Statistical A Priori Knowledge*, International Journal of Computer Vision (79) 2, 119–135, 2008

[15] A.K. Gupta and J. Tang, *Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate Gaussian models*, Biometrika (71) 3, 555–559, 1984

[16] M. Hallin and D. Paindaveine, *Optimal tests for homogeneity of covariance, scale, and shape*, Journal of Multivariate Analysis (100), 422–444, 2009

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

[18] T. Hotz, S. Huckemann, D. Gaffrey, A. Munk and B. Sloboda, *Shape spaces for pre-alingend star-shaped objects – modelling the temporal evolution of tree stems*, in revision, 2007

[19] S. Huckemann, *Parallel Shape Deformation*, submitted, 2009

[20] S. Huckemann, T. Hotz and A. Munk, *Global Models for the Orientation Field of Fingerprints: An Approach Based on Quadratic Differentials*, IEEE Transactions on Pattern Analysis and Machine Intelligence (30) 9, 1507–1519, 2008

[21] S. Huckemann and T. Hotz, *Principal Components Geodesics for Planar Shape Spaces*, Journal of Multivariate Analysis (100), 699–714, 2009

[22] S. Huckemann, T. Hotz and A. Munk, *Intrinsic Shape Analysis: Geodesic Principal Component Analysis for Riemannian Manifolds Modulo Lie Group Actions. Discussion paper*, Statistica Sinica, to appear, 2009

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

[24] P.E. Jupp and J.T. Kent, *Fitting Smooth Paths to Spherical Data*, Appl. Statist. (36) 1, 34–46, 1987

[25] H. Karcher, *Riemannian Center of Mass and Mollifier Smoothing*, Communications on Pure and Applied Mathematics XXX, 509–541, 1977

[26] D.G. Kendall, D. Barden, T.K. Carne and H. Le, *Shape and Shape Theory*, Wiley, 1999

[27] P.T. Kim and D.St.P. Richards, *Deconvolution density estimation on compact Lie groups*, Contemporary Mathematics (287), 155–171, 2001

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

[29] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, vol I, Wiley, 1963

[30] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, vol II, Wiley, 1969

[31] H. Le, I. Dryden and A. Kume, *Shape-space smoothing splines for planar landmark data*, Biometrika (94), 3, 513–528, 2007

[32] H. Le, *Locating Fréchet Means with an Application to Shape Spaces*, Adv. Appl. Prob. (SGSA) 33, 324–338, 2001

[33] H. Le, *Unrolling Shape Curves*, J. London Math. Soc. (2) 68, 511–526, 2003

[34] H. Le and A. Kume, *Detection of Shape Changes in Biological Features*, Journal of Microscopy (200), 140–147, 2000

[35] J.M. Lee, *Riemannian Manifolds : An Introduction to Curvature*, Springer, 1997

[36] E.L. Lehmann, *Testing Statistical Hypotheses (Springer Texts in Statistics)*, Springer, 1997

[37] K.V. Mardia, J.T. Kent and J.M. Bibby, *Mulitvariate Analysis*, Series in Probability and Mathematical Statistics, Academic Press, 1979

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

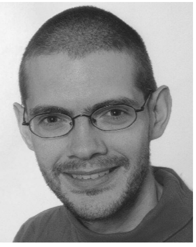
[39] K.V. Mardia and V. Patrangenaru, *Directions and projective shapes*, Ann. Statist. (33), 1666–1699, 2005

[40] B. O'Neill, *The fundamental equations of a submersion*, Michigan Math. J. (13) 4, 459–469, 1966

[41] V. Patrangenaru, *New Large Sample and Bootstrap Methods on Shape Spaces in High Level Analysis of Natural Images*, Commun. Statist.-Theory Meth. (30) 8&9, 1675-1693, 2001

[42] H. Ziezold, *Expected Figures and a Strong Law of Large Numbers for Random Elements in Quasi-Metric Spaces*, Trans. 7th Prague Conf. Inf. Theory, Stat. Dec. Func., Random Processes (A), Reidel Dordrecht, 591–602, 1977

Stephan Huckemann (Dr. rer. nat.) received his degree in mathematics from the University of Giessen, Germany in 1987. He was a visting lecturer and scholar at the University of Michigan, Ann Arbor (1987 - 1989) and a postdoctoral research fellow at the ETH Zürich, Switzerland (1989 - 1990). He then worked as a commercial software developer and returned to academia as a contributor to the computer algebra system MuPAD at the University of Paderborn, Germany (2001 - 2003). Currently he holds a DFG research position at the Institute for Mathematical Stochastics, University of Göttingen, Germany. His research interests include pattern recognition, statistical shape analysis, statistical inverse problems and fingerprint modelling.



Thomas Hotz (Dr. rer. nat.) received his degree in mathematics from the University of Heidelberg, Germany in 2002. He then worked as a research associate at the Department of Epidemiology and Public Health, University of Leicester, UK. Afterwards he held a position as Associate Statistician at the United Nations Statistics Division, New York, before he returned to Germany. In 2007 he received a PhD from the Institute for Mathematical Stochastics, University of Göttingen, Germany where he currently works as a postdoctoral research fellow. His research interests comprise mathematical and statistical methods in image analysis, as well as for the analysis of shapes; he combines both when studying fingerprints.

Axel Munk (Prof. Dr. rer. nat.) is director of the Institute for Mathematical Stochastics, University of Göttingen, Germany. He received his degree in mathematics and is PhD from the University in Göttingen in 1992 and 1994, respectively. His current research interests include non-parametric statistics, statistical methods in pattern recognition, medical statistics, and statistical inverse problems.

In 1992 Axel Munk received the Gustav Adolf Lienert Award of the International Biometric Society, German Section. From 1992 to 1994 he was supported by the "Studienstiftung des Deutschen Volkes", and afterwards by the German Research Foundation (DFG) until 1995. He recently (2005/06) was awarded a research semester on "Inverse Statistical Problems under Qualitative Prior Information" by the German Research Foundation (DFG). Axel Munk has been guest editor of several journals on statistics and medical statistics, and is currently associate editor of three statistical journals.