CENTRE FOR **STOCHASTIC GEOMETRY**
AND ADVANCED **BIOIMAGING**

Aasa Feragen and Andrea Fuster

Geometries and interpolations for symmetric positive definite matrices

No. 14, November 2016

Geometries and interpolations for symmetric positive definite matrices

Aasa Feragen[1] and Andrea Fuster[2]

[1]Department of Computer Science, University of Copenhagen, Denmark, aasa@di.ku.dk
[2]Department of Mathematics and Computer Science, Eindhoven Technical University, The Netherlands, a.fuster@tue.nl

Abstract

In this survey we review classical and recently proposed Riemannian metrics and interpolation schemes on the space of symmetric positive definite (SPD) matrices. We perform simulations that illustrate the problem of tensor fattening not only in the usually avoided Frobenius metric, but also in other classical metrics on SPD matrices such as the Wasserstein metric, the affine invariant / Fisher Rao metric, and the log Euclidean metric. For comparison, we perform the same simulations on several recently proposed frameworks for SPD matrices that decompose tensors into shape and orientation. In light of the simulation results, we discuss the mathematical and qualitative properties of these new metrics in comparison with the classical ones. Finally, we explore the nonlinear variation of properties such as shape and scale throughout principal geodesics in different metrics, which affects the visualization of scale and shape variation in tensorial data. With the paper, we will release a software package with Matlab scripts for computing the interpolations and statistics used for the experiments in the paper.[1]

1 Introduction

The space of symmetric positive definite (SPD) matrices, denoted $\mathrm{Sym}^+(n)$, is defined as

$$\mathrm{Sym}^+(n) = \left\{ T \in \mathbb{R}^{n\times n} \mid T_{ij} = T_{ji} \ \forall\ i,j = 1,\dots,n,\ x^T T x > 0 \ \forall\ x \in \mathbb{R}^n \setminus 0 \right\},$$

where the first property enforces symmetry and the second property enforces positive definiteness. An equivalent definition of positive definiteness is that all eigenvalues of T are positive.

SPD matrices play an important role in many data science applications. They coincide with covariance matrices of multivariate normal distributions, and therefore appear both in information geometry [3] and through covariance descriptors

[1]The code will be made available upon publication at `https://sites.google.com/site/aasaferagen/home/software`.

in computer vision [34]. They also represent second order tensors which e.g. model diffusion in diffusion tensor imaging (DTI) [5]. These applications have led to a rich theory of statistics and geometry in *spaces of SPD matrices*, which has interest both from a practical and theoretical point of view. Throughout the paper, we will use the terms "SPD matrix" and "tensor" interchangingly.

In this paper we survey classical and modern geometries defined on the space $\mathrm{Sym}^+(3)$ of SPD 3×3 matrices. All of these geometries can be extended to $\mathrm{Sym}^+(n)$, most of them directly, but for some this is technically more challenging. We perform simulations that illustrate the properties of the different approaches, in particular concerning the preservation information in tensor interpolation and statistics. Finally, we discuss properties of the different metrics regarding the visualization of statistical properties of datasets in $\mathrm{Sym}^+(3)$, in the context of recent work in visualization [37].

2 The positive definite cone and the Frobenius metric

It is easy to see that the space $\mathrm{Sym}^+(3)$ is a convex subset of the Euclidean space $\mathbb{R}^{3\times 3}$ of 3×3 matrices. Let $T_1, T_2 \in \mathrm{Sym}^+(3)$ and $t \in [0,1]$; now $T_1 + t(T_2 - T_1) \in \mathrm{Sym}^+(3)$ because it is clearly still symmetric, and

$$x^T(T_1 + t(T_2 - T_1))x = \underbrace{(1-t)}_{\geq 0}\underbrace{x^T T_1 x}_{>0} + \underbrace{t}_{\geq 0}\underbrace{x^T T_2 x}_{>0} > 0 \qquad \text{for all } x \in \mathbb{R}^3 \setminus 0,$$

where the inequality holds because only one of the factors $(1-t)$ and t can be 0. More precisely, the elements of $\mathrm{Sym}^+(3)$ reside on the convex *positive definite cone* in $\mathbb{R}^{3\times 3}$, which is the interior of the more frequently used positive semidefinite cone defined by $x^T T x \geq 0$.

Since $\mathrm{Sym}^+(3)$ is convex, we can perform linear interpolation between elements of $\mathrm{Sym}^+(3)$ while remaining within the manifold. These straight lines are geodesics on $\mathrm{Sym}^+(3)$ with respect to the so-called *Frobenius* metric on $\mathrm{Sym}^+(3)$, which is just the inherited Euclidean metric from $\mathbb{R}^{3\times 3}$. In particular, distances with respect to the Frobenius metric are given by the ambient space Euclidean distances:

$$d_F(T_1, T_2) = \|T_1 - T_2\|_2 = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{3}((T_1)_{i,j} - (T_2)_{i,j})^2}.$$

The Frobenius geodesic $\gamma_F \colon [0,1] \to \mathrm{Sym}^+(3)$ from T_1 to T_2 is trivially computed through Euclidean coordinate-wise interpolation:

$$\gamma_F(t) = T_1 + t(T_2 - T_1).$$

While the Frobenius metric is simple and efficient to work with, it comes with a significant cost when used for statistics on $\mathrm{Sym}^+(3)$. Frobenius interpolation between tensors in $\mathrm{Sym}^+(3)$ leads to a significant *swelling effect*, illustrated in Figure 1.

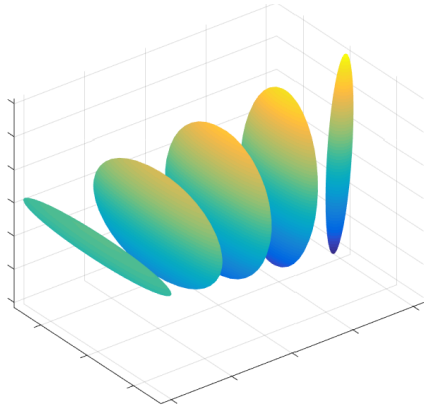


Figure 1: Samples from a geodesic interpolation between two identical line-like tensors at an angle of 85 degrees in the Frobenius metric. Note that the samples at the middle of the geodesic are very disc-like, thus exhibiting a very different shape from the two endpoints, and thus containing very little orientation information. This is called the swelling effect.

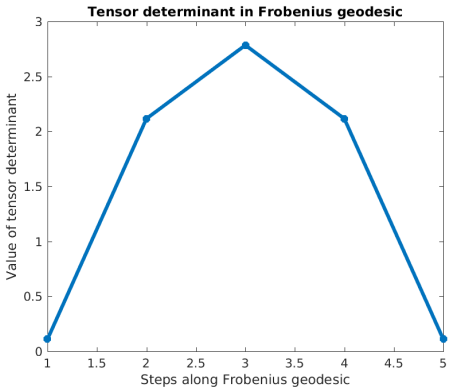


Figure 2: Values of the tensor determinant along the geodesic illustrated in Figure 1. The non-constant behavior of the determinant is what, in the literature, is referred to as the *swelling* effect of the Frobenius metric.

Here, two tensors are interpolated which have identical, ellipsoidal shape but which have 85 degrees difference in orientation. We observe that the tensors in the middle of the geodesic are much rounder than the endpoint tensors. The swelling effect is problematic for a number of applications in DTI. For example, when tensor interpolation is used for upsampling, the swelling effect leads to smoothing in the upsampled tensor field. The problem becomes more serious when statistics in $\mathrm{Sym}^+(3)$ are used e.g. for voxel-based morphometry. The mean of two tensors is the midpoint of the geodesic conencting them, which carries less shape information than the original two tensors. The mean of multiple tensors will typially exhibit even stronger smoothing of information, leading to reduced predictive or discriminative power.

The swelling effect is defined analytically as an increase in the determinant of the elements in $\mathrm{Sym}^+(3)$ as one interpolates two identically shaped but differently oriented ellipsoidal tensors. Figure 2 shows a plot of the determinant of each tensor depicted along the geodesic in Figure 1. As seen from the plot, the determinant is non-constant, and this is directly related to the tensor swelling throughout the geodesic connecting the two identically shaped, but differently oriented tensors in Figure 1.

In this paper, we will first survey classical Riemannian metrics on $\mathrm{Sym}^+(3)$ and illustrate that while most of these avoid the swelling effect, they still exhibit a strong and unwanted *fattening* effect, as is is also remarked in [22]. Next, we proceed to reviewing and exploring some more recently proposed metrics and interpolation schemes that aim to avoid the fattening effect by decoupling tensor shape from tensor orientation. While these have modelling advantages, they do come at a price: losing the statistics that come with a well-defined and computationally efficient geometric framework.

2.1 Acknowledgement

This paper was largely motivated by discussions of whether nonlinear geometries on $\mathrm{Sym}^+(3)$ might improve statistics and visualization for populations of tensors at the Dagstuhl seminar "Multidisciplinary Approaches to Multivalued Data: Modeling, Visualization, Analysis" (16142). We return to remark on this question in Section 5.5.

3 Classical Riemannian metrics on $\mathrm{Sym}^+(3)$

Riemannian metrics and other interpolation schemes for SPD matrices became an active area of research with the advent of diffusion tensor imaging (DTI), starting a quest for metrics that avoid the swelling effect while being computationally efficient. We start out by surveying the classics: The *Wasserstein* metric, the *affine-invariant* metric known in other contexts as the Fisher-Rao metric, and the *Log-Euclidean* metric.

3.1 The Wasserstein metric

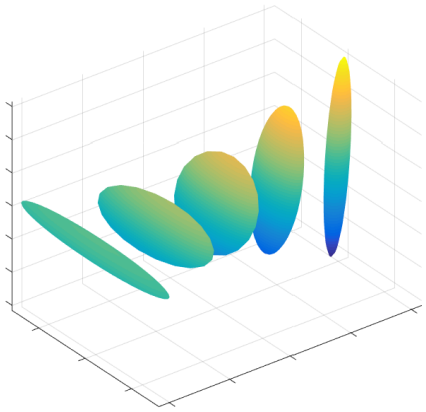


Figure 3: We observe a swelling of the tensor as we move throughout the geodesic between equally shaped, ellipsoidal tensors.

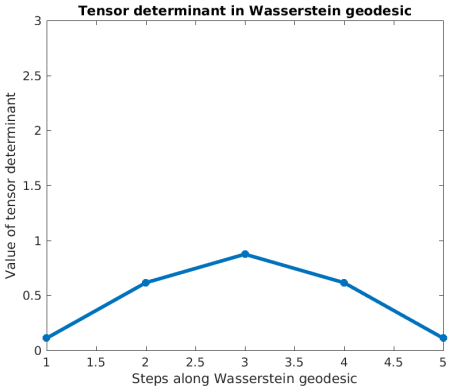


Figure 4: The tensor determinant increases as we move through the geodesic in Figure 3.

The *Wasserstein metric*, also known as the *earth mover's distance*, defines a general distance metric between arbitrary probability distributions on general metric spaces, which intuitively measures the amount of mass needed to transport one distribution into the other. Given a metric space (X, d) which is also a Radon space [4], and given $p \geq 1$, denote by $P_p(X)$ the set of probability measures μ on X such that

$$\int_X d^p(x, x_0) d\mu(x) < \infty.$$

The p^{th} Wasserstein distance between two probability measures $\mu, \eta \in P_p(X)$ is given by

$$W_p(\mu, \eta) = \left(\inf_{\gamma \in \Gamma(\mu,\eta)} \int_{X \times X} d^p(x, y) d\gamma(x, y) \right)^{\frac{1}{p}}, \qquad (x, y) \in X \times X,$$

where $\Gamma(\mu, \eta)$ is the set of measures on $X \times X$ whose marginals are μ and η, respectively. We shall focus on the case $p = 2$ for multivariate normal distributions $\mathcal{N}(\mathbf{0}, \Sigma)$ centered at the origin in the metric space $X = \mathbb{R}^3$ [30]. Any such multivariate normal distribution is described entirely by its covariance matrix Σ, and the set of such covariance matrices Σ is exactly the space of SPD matrices $\mathrm{Sym}^+(3)$. For general metric spaces and distributions, the Wasserstein distance is difficult to compute, but for centered normal distributions in $\mathbb{R}^n$, it has an analytical expression. We therefore obtain a Riemannian metric on $\mathrm{Sym}^+(3)$ by representing any SPD matrix $T \in \mathrm{Sym}^+(3)$ as a multivariate normal distribution with zero mean and covariance $\Sigma = T$. This defines the Riemannian Wasserstein metric on $\mathrm{Sym}^+(3)$ as a pull-back from the corresponding metric on the space of normal distributions.

Given this Riemannian metric, we can compute geodesics and geodesic distances, where the geodesic distance $d_W(T_1, T_2)$ between tensors $T_1, T_2 \in \mathrm{Sym}^+(3)$ is the length of the (shortest) geodesic $\gamma_W \colon [0, 1] \to \mathrm{Sym}^+(3)$ joining T_1 and T_2. In the space of normal distributions, both geodesics and geodesic distances have analytical solutions. The geodesic $\gamma_W \colon [0, 1] \to \mathrm{Sym}^+(3)$ connecting $T_1, T_2 \in \mathrm{Sym}^+(3)$ is given by

$$\gamma_W(t) = \big((1-t)I + tT_2^{\frac{1}{2}}(T_2^{\frac{1}{2}}T_1T_2^{\frac{1}{2}})^{-\frac{1}{2}}T_2^{\frac{1}{2}}\big)T_1\big((1-t)I + tT_2^{\frac{1}{2}}(T_2^{\frac{1}{2}}T_1T_2^{\frac{1}{2}})^{-\frac{1}{2}}T_2^{\frac{1}{2}}\big),$$

and the geodesic distance from T_1 to T_2 is given by

$$d_W(T_1, T_2) = W_2(\mathcal{N}(\mathbf{0}, T_1), \mathcal{N}(\mathbf{0}, T_2)) = \mathrm{tr}(T_1) + \mathrm{tr}(T_2) - 2\,\mathrm{tr}\sqrt{T_2^{\frac{1}{2}}T_1T_2^{\frac{1}{2}}}.$$

In the more general case of normal distributions with positive *semi*definite covariance matrices, the Wasserstein distance induces a stratified space geometry, which is analyzed in detail in [30].

The Wasserstein distance has a nice intuitive interpretation, it is well-understood mathematically, and it is immensely popular. However, looking at Figures 3 and 4, we observe that in practice, when interpolating two thin ellipsoidal tensors with identical shape but different orientation, the tensors go through significant swelling.

3.2 The affine-invariant metric

The classical affine-invariant metric [7, 14] was introduced to avoid the swelling effect on $\mathrm{Sym}^+(3)$. For two matrices V_1, V_2 in the tangent space at $P \in \mathrm{Sym}^+(3)$, $V_1, V_2 \in T_P\,\mathrm{Sym}^+(3) \in \mathbb{R}^{3\times3}$, it defines the tangent inner product via the tangent space at the identity

$$\langle V_1, V_2\rangle_P = \langle P^{-\frac{1}{2}}V_1P^{-\frac{1}{2}}, P^{-\frac{1}{2}}V_2P^{-\frac{1}{2}}\rangle = \mathrm{Tr}\big(P^{-\frac{1}{2}}V_1P^{-1}V_2P^{-\frac{1}{2}}\big).$$

The geodesic between tensors $T_1, T_2 \in \mathrm{Sym}^+(3)$ is given analytically by

$$\gamma_{AI}(t) = T_1^{\frac{1}{2}}\exp\big(tT_1^{-\frac{1}{2}}T_2T_1^{-\frac{1}{2}}\big)T_1^{\frac{1}{2}}, \qquad t \in [0, 1],$$

and the geodesic distance from T_1 to T_2 is

$$d_{AI}(T_1, T_2) = \|\log(T_1^{-\frac{1}{2}}T_2T_1^{-\frac{1}{2}})\|_F.$$

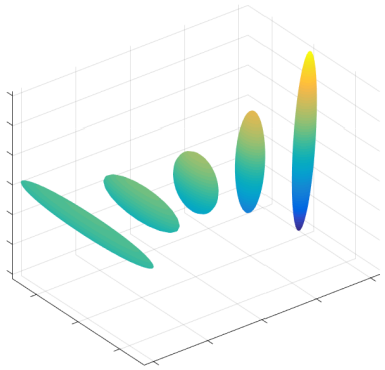


Figure 5: Samples from a geodesic interpolation between two identical line-like tensors at an angle of 85 degrees in the affine-invariant metric. The samples at the middle of the geodesic are less disc-like than in the Frobenius metric. Note, however, that they still have different (fatter) shape than the endpoint tensors.

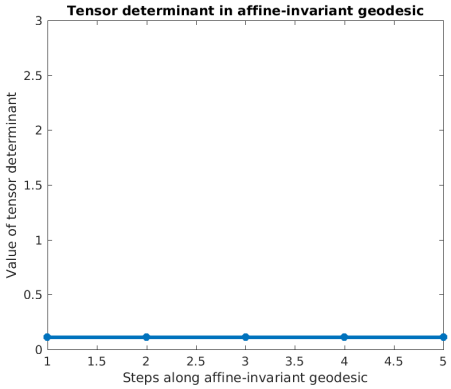


Figure 6: The tensor determinant is constant along the affine invariant geodesic illustrated in Figure 5. The affine-invariant metric thus avoids the swelling effect has a fattening effect throughout the geodesic.

The affine-invariant metric is a special case of the Fisher-Rao metric [3] on probability distributions, restricted to zero-mean multivariate Gaussian distributions with covariance in $\mathrm{Sym}^+(3)$, as in the previous section.

In Figure 5 we see an example of an affine invariant geodesic on the same example as above, and in Figure 6 we track the determinant of the interpolated tensors throughout the geodesic. We see that there is no swelling effect, in the sense that the determinant remains constant througout the geodesic. However, there is a noticeable *fattening* effect leading to rounder tensors in the middle of the geodesic than at the endpoints.

It is clear from the above equations that the affine-invariant metric is relatively computationally expensive, as it relies on a number of evaluations of matrix exponentials, logs, square roots and inverses, and for this reason, the more tractable *Log-Euclidean* metric was proposed in 2006 [5].

3.3 The Log-Euclidean metric

The Log-Euclidean metric on $\mathrm{Sym}^+(3)$ [5] utilizes the observation that the matrix exponential defines the one-to-one mapping $\exp\colon \mathrm{Sym}(n) \to \mathrm{Sym}^+(3)$ from the vector space of symmetric 3×3 matrices into the manifold $\mathrm{Sym}^+(3)$ of SPD 3×3 matrices, whose inverse is the matrix logarithm (which is well defined on $\mathrm{Sym}^+(3)$). This means that the differential structure of $\mathrm{Sym}(n) = \mathbb{R}^{(n^2+n)/2}$ can be pulled back to $\mathrm{Sym}^+(3)$ via the matrix logarithm. Equivalently, the logarithm provides a nonlinear transformation of $\mathrm{Sym}^+(3)$ into the Euclidean space $\mathbb{R}^{(n^2+n)/2}$, where analysis can take place. In this way, the Log-Euclidean metric defines a *Euclidean* structure on $\mathrm{Sym}^+(3)$. In this geometry, the distance between two tensors T_1 and T_2 is given by

$$d_{LE}(T_1, T_2) = \| \log(T_1) - \log(T_2) \|_F.$$

More importantly, the geodesic between T_1 and T_2 under the Log-Euclidean metric is given by the straight line between $\log(T_1)$ and $\log(T_2)$. However, this straight line resides in $\mathbb{R}^{(n^2+n)/2}$, not in $\mathrm{Sym}^+(3)$. It can be pulled back to $\mathrm{Sym}^+(3)$ using the matrix exponential, giving rise to the following analytical expression for the geodesic $\gamma_{LE}\colon [0,1] \to \mathrm{Sym}^+(3)$:

$$\gamma_{LE}(t) = \exp\big(\log(T_1) + (1-t)\log(T_2)\big), \qquad t \in [0,1].$$

The Log-Euclidean distances and geodesics thus avoid many of the involved matrix power computations needed in the affine-invariant framework, although they still require matrix exponentials and logarithms. Moreover, empirically, the Log-Euclidean geodesics are often found to be extremely similar to the affine invariant ones. This is illustrated in our geodesic interpolation example, shown in Figure 7. As we can see in Figure 8, the Log-Euclidean geodesic keeps the tensor determinant constant when interpolating identically shaped tensors, and it therefore does not exhibit the swelling effect – but we see in Figure 7 that it does, just like the affine-invariant metric, still lead to a fattening effect in the tensors at the middle of the geodesic. Note also the qualitative similarity between the affine-invariant geodesics in Figure 5 and the Log-Euclidean geodesics in Figure 7.

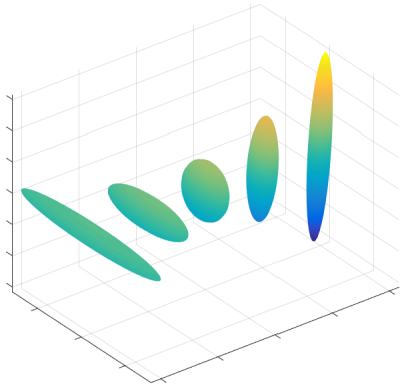


Figure 7: Samples from a geodesic interpolation between two identical line-like tensors at an angle of 85 degrees in the Log-Euclidean metric.

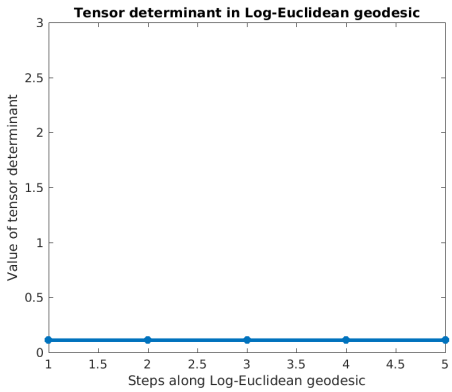
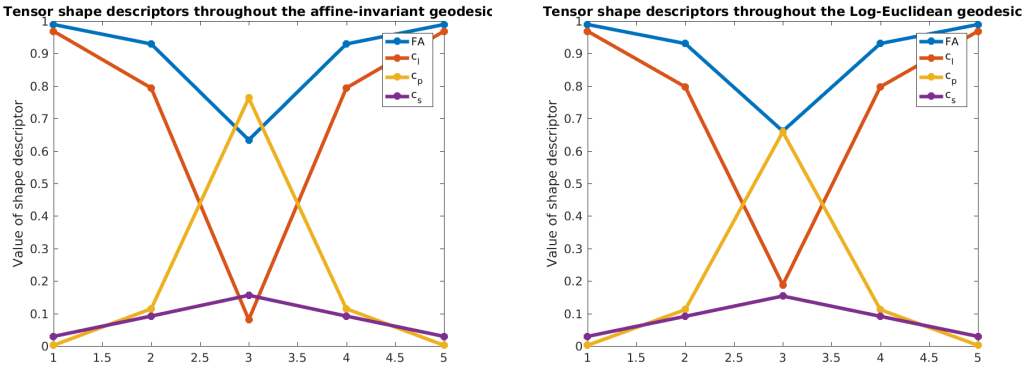


Figure 8: The tensor determinant is constant along the Log-Euclidean geodesic illustrated in Figure 7. It therefore also avoids the swelling effect, but exhibits a fattening effect as seen in Figure 7.

4 Avoiding the swelling, but not the fattening, effect

Figures 5 and 7 illustrate that while the affine-invariant and Log-Euclidean metrics do keep the tensor determinant fixed and therefore avoid the previously defined swelling effect, they do not preserve tensor shape when interpolating between identically shaped, but differently oriented tensors. Let us quantify this effect further by



(a) Affine-invariant **(b)** Log-Euclidean

Figure 9: We quantitatively confirm the fattening effect in the affine invariant and Log-Euclidean metrics: The shape is variable and the tensors become fatter towards the middle of the geodesic.

considering four different shape indices: The fractional anisotropy (FA) of a tensor T is given by

$$FA = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}},$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of T. In the same notation, the three Westin tensor shape indices [36] are given by

$$c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1}, \qquad c_p = \frac{\lambda_2 - \lambda_3}{\lambda_1}, \qquad c_s = \frac{\lambda_3}{\lambda_1},$$

and quantify the tensor's resemblence to a line, a plane or a sphere, respectively. For the same two endpoint tensors used previously, we plot the four different shape measures throughout the geodesic for the affine invariant and Log-Euclidean metrics in Figure 9.

The plots in Figure 9 clearly confirm the fattening effect: The tensors are less line-like and more plane- and sphere-like towards the middle of the geodesic, despite the fact that the endpoint tensors have identical ellipsoidal shape. To avoid this effect, several approaches have appeared that aim to decouple tensor shape and tensor rotation.

5 Decoupling shape and rotation

The eigenvalue decomposition of an SPD matrix

$$T = Q\Lambda Q^{-1} \tag{5.1}$$

into a rotation matrix $Q \in \mathrm{SO}(3)$ given by column-wise eigenvectors and a diagonal matrix Λ containing the eigenvalues, provides a natural way of splitting T into its shape- and orientation properties. This presents an attractive opportunity to avoid the fattening effect, and it is also useful for interpretation. An abundance of approaches to decouple shape and rotation have appeared [23, 16, 35, 28, 22, 18, 9], seemingly independent of each other, and we shall review some of the most important ones below, in order of increasing complexity.

5.1 The shape-and-orientation rotation metric

In the *shape-and-orientation* metric [35] the tensor $T = Q\Lambda Q^{-1}$ is considered to reside on the Riemannian product manifold $\mathrm{SO}(3) \times \mathbb{R}^3_+$, where $\mathrm{SO}(3)$ is given the angular geodesic metric and $\mathbb{R}^3_+$ is given the bi-invariant metric. A geodesic γ between two SPD matrices T_1 and T_2 in this metric is a product geodesic $\gamma = \gamma_r \times \gamma_s$, where γ_r is a geodesic between Q_1 and Q_2 in $\mathrm{SO}(3)$, and γ_s is a geodesic between Λ_1 and Λ_2, where $T_1 = Q_1\Lambda_1 Q_1^{-1}$ and $T_2 = Q_2\Lambda_2 Q_2^{-1}$ as in (5.1).

Such geodesics are given analytically by the formulas

$$\gamma_r(t) = Q_1 \exp(t \cdot \log_{\mathrm{SO}(3)}(Q_1^{-1}Q_2)),$$
$$\gamma_s(t) = \Lambda_1 \exp(t \cdot \log(\Lambda_1^{-1}\Lambda_2)),$$

where $\log_{\mathrm{SO}(3)}$ denotes the log map on $\mathrm{SO}(3)$, and the corresponding geodesic $\Gamma\colon [0,1] \to \mathrm{Sym}^+(3)$ from T_1 to T_2 is given by

$$\Gamma(t) = \gamma_r(t) \cdot \gamma_s(t) \cdot \gamma_r(t)^{-1}.$$

An example of a shape-and-orientation geodesic is shown in Figures 10 and 11, and the determinant and shape indices are shown in Figures 12 and 13. The corresponding geodesic distance between two tensors T_1 and T_2 is given by

$$d^2_{SAO}\left((Q_1,\Lambda_1),(Q_2,\Lambda_2)\right) = d^2_{\mathrm{SO}(3)}(Q_1,Q_2) + d^2_{\mathbb{R}^3_+}(\Lambda_1,\Lambda_2)$$

where

$$d_{\mathrm{SO}(3)}(Q_1,Q_2) = \frac{1}{\sqrt{2}}\|\log_{\mathrm{SO}(3)}(Q_2Q_1^T)\|_F, \qquad d_{\mathbb{R}^3_+}(\Lambda_1,\Lambda_2) = \|\log(\Lambda_2\Lambda_1^{-1})\|_F.$$

As we see from the simulations in Figures 10 and 11, where we compute the geodesic between identically shaped but differently oriented tensors, the tensor shape is kept constant throughout the deformation – as intended. This is confirmed by the constant determinant and shape indices shown in Figures 12 and 13. Moreover, the construction is fast and apparently easy to work with as all the formulas are analytical, and the metric is intuitive and preserves tensor shape properties well. However, the construction ignores a fundamental problem: There exist multiple decompositions $T_i = Q_i\Lambda_i Q_i^{-1}$ of the endpoint tensors T_i, $i = 1, 2$, and different choices of Q_i and Λ_i generate different interpolations $\Gamma\colon [0,1] \to \mathrm{Sym}^+(3)$. We see this effect very clearly in the example geodesic in Figures 10 and 11, where the path chosen turns 95 degrees and is – while locally a geodesic – *not* the shortest path between the two tensors (the shortest path turns 85 degrees).

The underlying problem is that while any point on $\mathrm{SO}(3) \times \mathbb{R}^3_+$ does, indeed, correspond to a tensor, the map $F\colon \mathrm{SO}(3) \times \mathbb{R}^3_+ \to Sym^+(3)$ defined by $F(Q,\Lambda) = Q\Lambda Q^{-1}$ is *not* injective, because the eigenvalue decomposition is not unique. The map F is the map that takes an eigenvalue decomposition to its corresponding tensor.

There are several ways in which the eigenvalue decomposition of a tensor, as in Equation (5.1), is not unique. First, given any eigenvector e_i of a tensor T, its

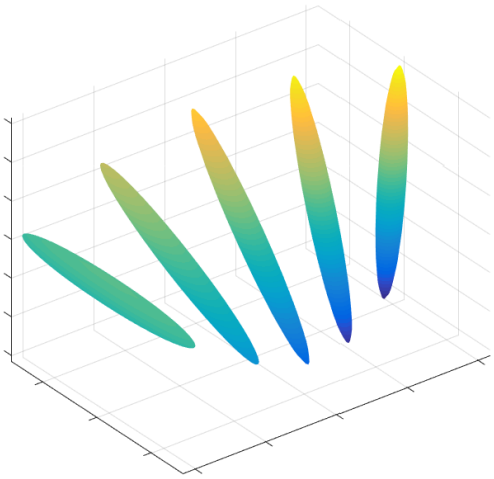


Figure 10: Samples from a geodesic interpolation between two identical line-like tensors at an angle of 85 degrees in the shape-and-orientation metric.

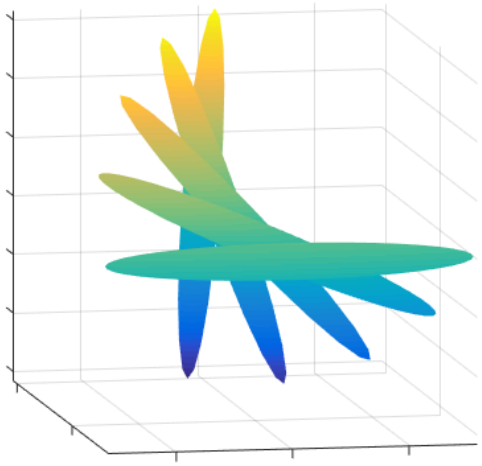


Figure 11: The same geodesic interpolation as in Figure 10, from a different viewpoint. It is clear that the interpolation is not the shortest possible 85 degree interpolation, but a longer 95 degree interpolation.

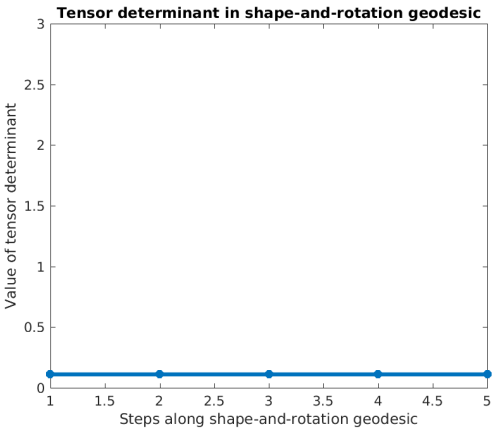


Figure 12: The tensor determinant is constant along the shape-and-orientation geodesic illustrated in Figure 10.

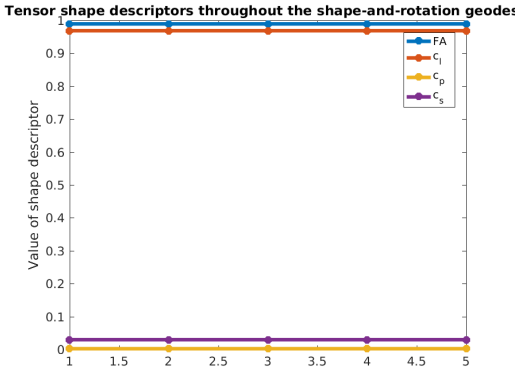


Figure 13: The FA and Westin shape descriptors are also constant along the shape-and-orientation geodesic illustrated in Figure 10.

antipode $-e_i$ is also an eigenvector. The orientation part Q of the tensor decomposition consists of eigenvectors of T, but not all combinations of eigenvectors lead to a matrix $Q \in \mathrm{SO}(3)$. In particular, multiplying an eigenvector by -1 gives another eigenvector, but flips the sign of Q, and only those eigenvector matrices that give $\det(Q) = 1$ actually reside in $\mathrm{SO}(3)$. Moreover, even those eigenvector sets that do define a $Q \in \mathrm{SO}(3)$, are not unique. This is what happens in Figures 10 and 11. Another source of non-uniqueness is the order of the eigenvalues and eigenvectors in the decomposition. This issue is usually avoided in practice by requiring $\lambda_1 \geq \lambda_2 \geq \lambda_3$ in $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. However, this is also problematic, in particular when two eigenvalues approach each other. The problem with multiple representations of the same tensor becomes particularly complicated when the endpoint tensors have two or more identical eigenvalues [18]. To avoid the resulting technical complications, our experiments and code will assume that endpoint tensors have three distinct eigenvalues.

5.2 Scaling-rotation curves

While the shape-and-orientation metric as defined in [35] is a very simple way of defining a metric using eigenvalue decompositions, it is not the first appearance of the manifold $\mathrm{SO}(3) \times \mathbb{R}^3_+$ in an attempt to generate a geometric framework for $\mathrm{Sym}^+(3)$. In a series of papers [28, 22, 18], A. Schwartzman and collaborators define *scaling-rotation* curves and a corresponding geometry, as follows.

A scaling-rotation curve [28] between $T_1, T_2 \in \mathrm{Sym}^+(3)$ is a geodesic in $\mathrm{SO}(3) \times \mathbb{R}^3_+$ between representatives (Q_1, Λ_1) and (Q_2, Λ_2) of T_1 and T_2, where the metrics on $\mathrm{SO}(3)$ and $\mathbb{R}^3_+$ are the same bi-invariant metrics as in Section 5.1. Scaling-rotation curves suffer from the same non-uniqueness problems as the shape-and-orientation metric from Section 5.1. A more recent series of papers [22, 18] attempt to handle this by factoring out multiple representations of the same tensor as follows. Given the mapping

$$F\colon \mathrm{SO}(3) \times \mathbb{R}^3_+ \to \mathrm{Sym}^+(3), \qquad (Q, \Lambda) \mapsto Q\Lambda Q^{-1}, \tag{5.2}$$

one can define an equivalence relation $\sim$ on $\mathrm{SO}(3) \times \mathbb{R}^3_+$ by setting $(Q, \Lambda) \sim (Q'\Lambda')$ whenever $F(Q, \Lambda) = F(Q'\Lambda')$. This gives rise to a *quotient space*

$$(\mathrm{SO}(3) \times \mathbb{R}^3_+)/\!\sim \; = (\mathrm{SO}(3) \times \mathbb{R}^3_+)/F, \tag{5.3}$$

whose elements are equivalence classes, denoted $\overline{(Q, \Lambda)} = \{(Q', \Lambda') \in \mathrm{SO}(3) \times \mathbb{R}^3_+ : F(Q, \Lambda) = F(Q', \Lambda')\} = F^{-1}(Q\Lambda Q^{-1})$. In the quotient space, each tensor is represented exactly once.

The quotient space $(\mathrm{SO}(3) \times \mathbb{R}^3_+)/F$ can be identified with $\mathrm{Sym}^+(3)$, since the map F descends to a $1-1$ mapping $\bar{F}\colon (\mathrm{SO}(3) \times \mathbb{R}^3_+)/F \to \mathrm{Sym}^+(3)$. This quotient space is not a smooth manifold, but it is a *stratified space*, meaning that it is a union of smooth manifolds which are adjacent to each other in a "well-behaved" way (see [26] for details on stratified spaces). The *strata*, or manifold components, are given by elements of $\mathrm{Sym}^+(3)$ with a fixed number of eigenvalue degeneracies, meaning the top stratum consists of tensors with three distinct eigenvalues, the next stratum consists of elements with two identical eigenvalues, etc.

Groisser et al. [18] note that tensors in $\mathrm{Sym}^+(3)$, represented as equivalence classes $F^{-1}(T) = \overline{(Q, \Lambda)}$ in $\mathrm{SO}(3) \times \mathbb{R}^3_+$, can be interpolated by *minimal* scaling-rotation curves between equivalence classes $\overline{(Q, \Lambda)}, \overline{(Q', \Lambda')}$. This gives rise to the *scaling-rotation distance*

$$\begin{aligned} cd_{\mathcal{SR}}(T_1, T_2) &= d_{\mathrm{SO}(3)\times\mathbb{R}^3_+}\big(\overline{(Q_1, \Lambda_1)}, \overline{(Q_2, \Lambda_2)}\big) \\ &= \min\{d_{\mathrm{SO}(3)\times\mathbb{R}^3_+}\left((Q_1, \Lambda_1), (Q_2, \Lambda_2)\right) \mid T_1 = Q_1\Lambda_1 Q_1^{-1}, T_2 = Q_2\Lambda_2 Q_2^{-1}\}. \end{aligned} \tag{5.4}$$

5.2.1 The scaling-rotation distance is not a metric

As remarked in [22, 18], the scaling-rotation distance is *not* a metric, and in particular it is *not* the quotient (distance) metric on $(\mathrm{SO}(3) \times \mathbb{R}^3_+)/F$. The quotient metric [8, p. 65] defined on a quotient

$$X/\sim = \{\bar{x} \mid x \in X, \bar{x} = \bar{y} \text{ if } x \sim y\}$$

is given by

$$\bar{d}(\bar{x}, \bar{y}) = \inf\Big\{\sum_{i=1}^{n-1} d(\bar{z}_i, \bar{z}_{i+1}) \mid z_1, \ldots, z_n \in X, z_1 \in \bar{x}, z_n \in \bar{y}\Big\}.$$

Here, the equivalence classes $\bar{x}, \bar{y}$ are viewed as subsets of X and $d(\bar{x}, \bar{y}) = \inf\{d(x, y) \mid x \in \bar{x}, y \in \bar{y})$ is the set distance between subsets of X. Meanwhile, the scaling-rotation distance is a special case of

$$\tilde{d}(\bar{x}, \bar{y}) = \inf\{d(x, y) \mid x \in \bar{x}, y \in \bar{y}\}.$$

The quotient metric is the minimal total cost of hopping from $\bar{x}$ to $\bar{y}$ with stopovers at a finite set of equivalence classes $\bar{z}_i$, while the scaling-rotation distance does not allow stopovers. The lack of stopovers causes the scaling-rotation distance to *not* satisfy the triangle inequality, which is why it does not satisfy the criteria for being a metric.

This is easy to see via a visual example. The scaling-rotation distance between the left- and rightmost tensors in Figure 14 would be given by a 90 degrees rotation with no shape component, while the distance for either of those two tensors to the round tensor in the center would consist in a very small shape change with no rotational component. The sum of the scaling-rotation distances from the leftmost tensor, via the middle one, to the rightmost, is smaller than the scaling-rotation distance between the two endpoint tensors.

Nevertheless, the scaling-rotation distance *does* define "minimal" interpolations that solve some of the problems of the shape-and-orientation metric from [35, 28]. The quotient metric on $(\mathrm{SO}(3) \times \mathbb{R}^3_+)/F$ remains unexplored to the best of our knowledge. This is likely due to the computational complexity and non-Riemannian structure of the space.

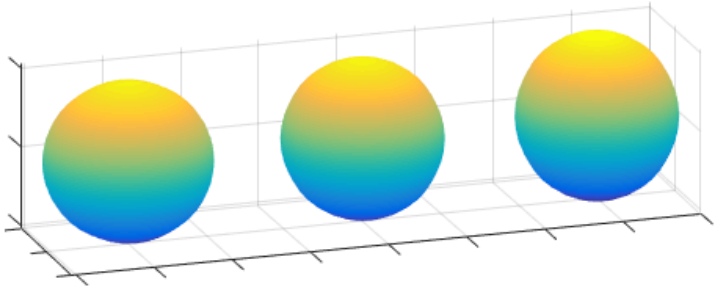


Figure 14: An example of two tensors for which the scaling-rotation distance does not satisfy the triangle inequality. The left- and rightmost tensors have identical slightly ellipsoidal shape, but are rotated at an angle of 90 degrees. The tensor in the middle is spherical.

5.3 Linear invariant tensor interpolation

Based on a set of tensor invariants proposed by Ennis et al. [11], Kindlmann *et al.* [23] proposed the *geodesic loxodromes* framework in order to avoid fattening when interpolating tensors. A *tensor invariant* is a scalar which depends only on the shape of the tensor – or mathematically speaking, on its eigenvalues. A classical example is the fractional anisotropy (FA). A geodesic loxodrome between two tensors T_1 and T_2 is the shortest path from T_1 to T_2 for which certain tensor shape parameters (either K_1-K_3 or R_1-R_3, as reviewed below) are linearly interpolated. In the original geodesic loxodromes framework, the interpolations were found by optimization, which is inexact and sometimes computationally expensive.

Gahm et al [16] utilize the following combination of R- and K-invariants from [23], which allows an analytical reconstruction of the eigenvalues along the interpolated path[2]:

$$\begin{aligned} K_1(T) &= \mathrm{tr}(T); \qquad R_2(T) = FA(T) = \sqrt{\frac{3}{2}}\frac{\|\tilde{T}\|_F}{\|T\|_F}; \\ R_3(T) &= 3\sqrt{6}\det\left(\frac{\tilde{T}}{\|\tilde{T}\|_F}\right), \end{aligned} \tag{5.5}$$

where $\|T\|_F$ is the Frobenius norm of T, and

$$\tilde{T} = T - \frac{\mathrm{tr}(T)}{3} I_3 \tag{5.6}$$

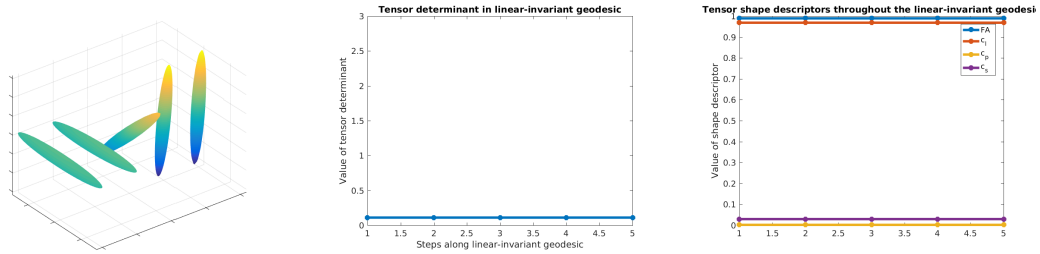
is the anisotropic part of T, the so-called deviatoric tensor. Note that R_2 is the tensor FA used previously, while K_1 measures tensor scale. Now, the path through SPD matrices in $\mathrm{Sym}^+(3)$ with eigenvalues $\lambda_1, \lambda_2, \lambda_3$ given by

$$\lambda_i(t) = \tfrac{1}{3}\big(K_1(t)\big) + \frac{2K_1R_2(t)}{3\sqrt{3-2(R_2(t))^2}} \cos\left(\frac{\arccos(R_3(t)) + P_i}{3}\right), \tag{5.7}$$

where $P_i = 0, -2\pi, 2\pi$ for $i = 1, 2, 3$, results in a linear interpolation

$$\begin{aligned} K_1(t) &= (1-t)K_1(T_1) + tK_1(T_2), \\ R_2(t) &= (1-t)R_2(T_1) + tR_2(T_2), \\ R_3(t) &= (1-t)R_3(T_1) + tR_3(T_2), \end{aligned} \tag{5.8}$$

[2]Note that there are some typos in the definitions of R_2 and R_3 in [16].



(a) The LIT interpolation. **(b)** Tensor determinant. **(c)** Shape descriptors.

Figure 15: The LIT interpolation between the same two thin, ellipsoidal tensors as before. We see that the determinant and shape of the tensor is unchanged throughout the geodesic, but we also see that the orientation is not interpolated at a constant rate. To see the reason for this effect, we refer back to the Frobenius interpolation in Figure 1, where the tensor orientation also does not change at a constant rate.

in the corresponding three tensor invariants. This tells us how to interpolate the shape of the tensors while linearly, and in particular, monotonically, interpolating the corresponding tensor invariants.

However, we do not yet know how to interpolate the orientation of the tensors. In [16] this is handled by using the orientation component of a Frobenius geodesic. Let $\gamma_F(t)$ be the tensor at time t in the Frobenius geodesic $\gamma_F \colon [0,1] \to \mathrm{Sym}^+(3)$ from T_1 to T_2, and let

$$Q(t)\tilde{\Lambda}(t)Q(t)^{-1} = \gamma_F(t) \tag{5.9}$$

be an eigenvalue decomposition of the time t tensor $\gamma_F(t)$. Let the diagonal matrix

$$\Lambda(t) = \begin{pmatrix} \lambda_1(t) & 0 & 0 \\ 0 & \lambda_2(t) & 0 \\ 0 & 0 & \lambda_3(t) \end{pmatrix}, \tag{5.10}$$

consist of the eigenvalues obtained in Eq. 5.7. The LIT interpolation $\gamma_{\mathrm{LIT}} \colon [0,1] \to \mathrm{Sym}^+(3)$ from T_1 to T_2 is defined as

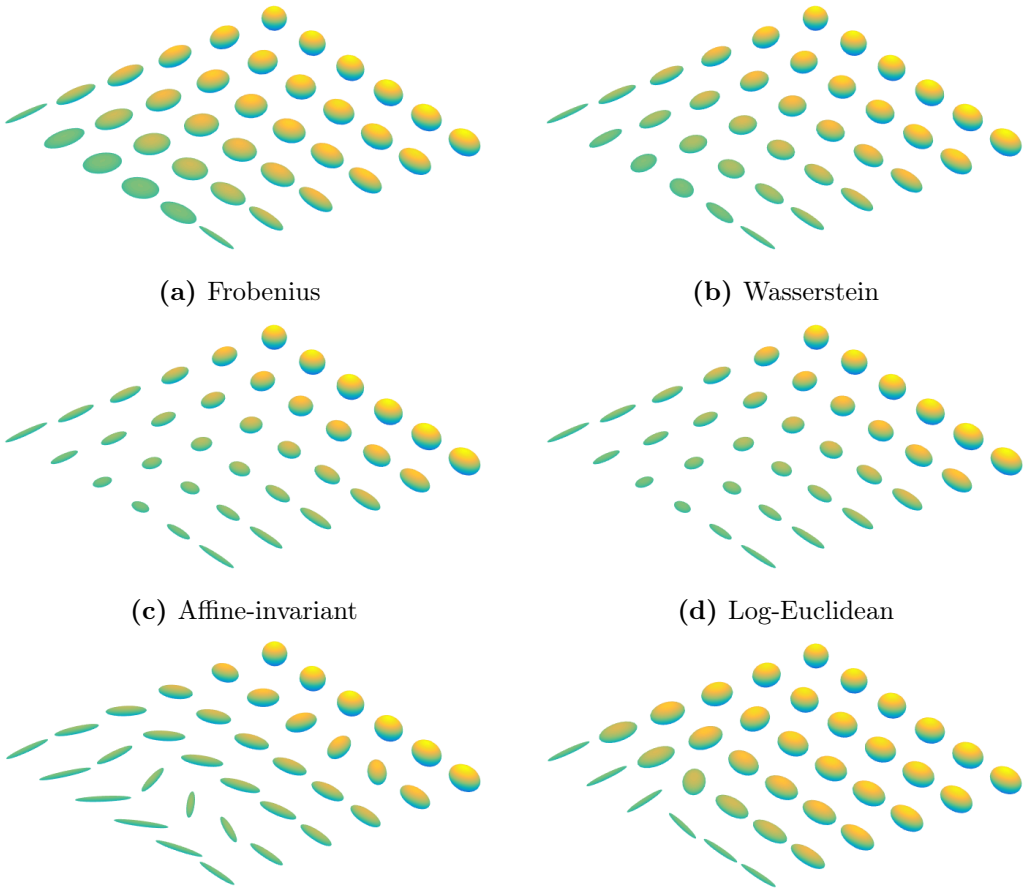
$$\gamma_{\mathrm{LIT}}(t) = Q(t)\Lambda(t)Q(t)^{-1}. \tag{5.11}$$

That is, the orientation component of the interpolation from T_1 to T_2 coincides with that of a Frobenius interpolation from T_1 to T_2, and the shape interpolation component is the one which linearly interpolates K_1, R_2 and R_3, given in Equation 5.7.

We see that, at least in our running example, the fattening effect is resolved using the LIT interpolation. Tensor invariant interpolation methods have the advantage that shape is interpolated monotonically, and even linearly, with respect to the prespecified shape invariants. For the combination of invariants used in γ_{LIT} [16], there is an analytical solution, making the interpolation scheme computationally feasible. These are attractive properties.

On the negative side, the particular choice of angular interpolation is not ideal when the difference in orientation is large – this is clear from Figure 15. This could be resolved by choosing the angular component differently, e.g. from one of the alternative tensor metrics. Moreover, to the best of our knowledge, the interpolations do not correspond to geodesics in a given geometric space. This makes it unclear to what extent the derived interpolations can be used in a geometric statistics framework to obtain Fréchet means, principal components, regression etc.

5.4 Further simulations



(a) Frobenius **(b)** Wasserstein

(c) Affine-invariant **(d)** Log-Euclidean

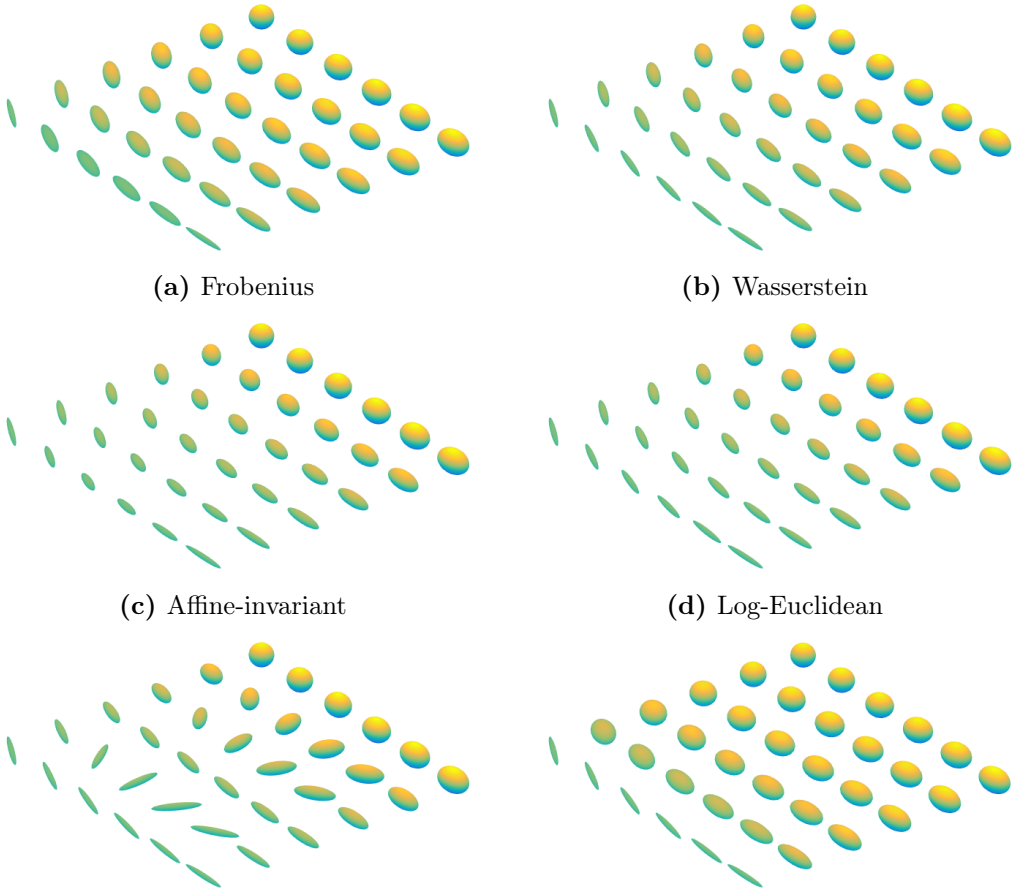
(e) Shape-and rotation **(f)** Linear invariant inteprolation

Figure 16: Interpolation between four tensors, all with three distinct eigenvalues but different shapes and orientations.

In the above, we have used as a running example an interpolation between two identically shaped ellipsoidal tensors at an 85 degree angle of each other. Figures 16 and 17 illustrate additional simulations,[3] illustrating how the different metrics and interpolation schemes handle interpolation between tensors of different shape and less extreme orientation differences.

In Figures 16 and 17, all four tensors have three distinct eigenvalues to give optimal conditions for the shape-and-orientation metric. Two of the tensors are thin and ellipsoidal; in Figure 16 one of them is rotated 85 degrees. The two remaining tensors have the same orientation as the first, but different shape: one is a bit fatter, and one is almost spherical. In Figure 17 the experiment is repeated with a 30 degree rotation instead of 85 degrees. Code for the experiment will be made available online at `https://sites.google.com/site/aasaferagen/home/software`.

[3]For the square geodesic interpolations the optimal interpolation would have been made as a weighted Fréchet mean; however, this strategy does not apply to the linear invariant tensor interpolation. Therefore, all interpolations were made by first interpolating pairs of corners to obtain two "side" interpolations and then interpolating the elements of the sides to obtain the remainder of the square.



(a) Frobenius **(b)** Wasserstein

(c) Affine-invariant **(d)** Log-Euclidean

(e) Shape-and rotation **(f)** Linear invariant inteprolation

Figure 17: Interpolation between four tensors, all with three distinct eigenvalues but different shapes and orientations. Two tensors are thin and ellipsoidal; one of them is rotated 30 degrees. The two remaining tensors have the same orientation as the first, but different shape: One is less thin and one is almost spherical.

Note the difference between the first and second columns of the interpolation results for the shape-and-orientation metric in Figure 16(e): The inconsistencies with respect to choice of representation of tensors results in consecutive pairwise interpolations being radically different geodesics in $\mathrm{Sym}^+(3)$, some of them not being shortest paths. This corresponds to our remarks in Section 5.1.

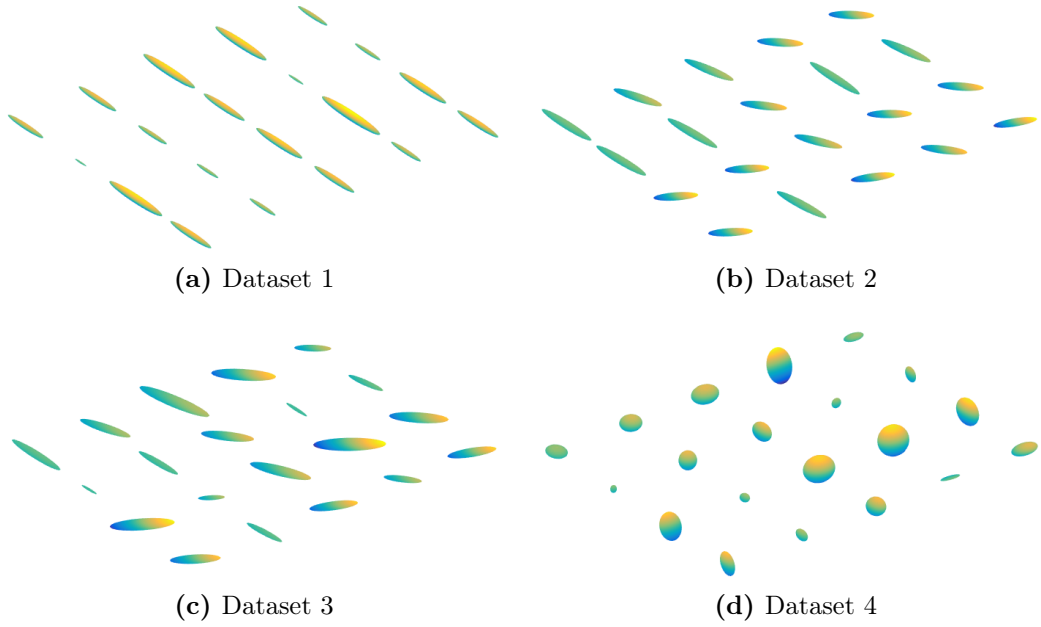
Note the abrupt change in orientation in the leftmost column of Figure 16(f). This is an effect of the choice of rotation component in the linear invariant interpolation, which is the rotation found in the corresponding Frobenius geodesic. The connection is visually evident by comparing to the Frobenius geodesic in Figure 16(a).

5.5 Tensor statistics and tensor decomposition for visualization purposes

This paper was largely motivated by a discussion of visualization of tensor populations at the Dagstuhl seminar "Multidisciplinary Approaches to Multivalued Data: Modeling, Visualization, Analysis" (16142). In a recent paper, Zhang et al [37] visu-

alize tensor population variation by separately visualizing variation in scale, shape and orientation.

This leads to the question of how scale, shape and orientation can be decomposed in different geometric frameworks. In this section we present simulations that aim to make clear that the choice of metric on $\mathrm{Sym}^+(3)$ affects data variation both quantitatively in the notion of variance, and in the extent to which it is possible to decompose the variation into scale, shape and orientation without misrepresenting the metric. We do so by visualizing the first component of Fletcher principal geodesic analysis [13] on four different simulated tensor populations, shown in Figure 18, for four of the above discussed metrics.



(a) Dataset 1 **(b)** Dataset 2

(c) Dataset 3 **(d)** Dataset 4

Figure 18: Four similated datasets.

Each dataset consists of 20 tensors. The tensors in dataset 1 have identical shape and rotation, but variable scale (defined as the Frobenius norm of the tensor matrix); the tensors in dataset 2 have identical shape and scale but variable rotation; the tensors in dataset 3 have identical shape but the scale and rotation of the corresponding tensors in datasets 1 and 2, respectively. Dataset 4 has variable shape, scale and orientation.

Figure 19 shows sampled tensors along the first geodesic principal component for the Frobenius, affine-invariant, Log-Euclidean and shape-and-rotation metrics. The middle sample for each dataset and metric is the Fréchet mean. The first conclusions to be made from Figure 19 is that the Riemannian metric frameworks are rather different, and that neither of the first three metrics capture the dataset variability very well except for in the case where orientation is kept constant. In particular, while every tensor in the datasets 1–3 has identical shape, the geodesic principal components of these datasets for the first three Riemannian metrics indicate shape variation. In dataset 4, however, there *is* shape variation in the dataset, but the tensors sampled along the geodesic principal component do not exhibit much shape variation. The tendency holds for all of the first three metrics. This indicates that

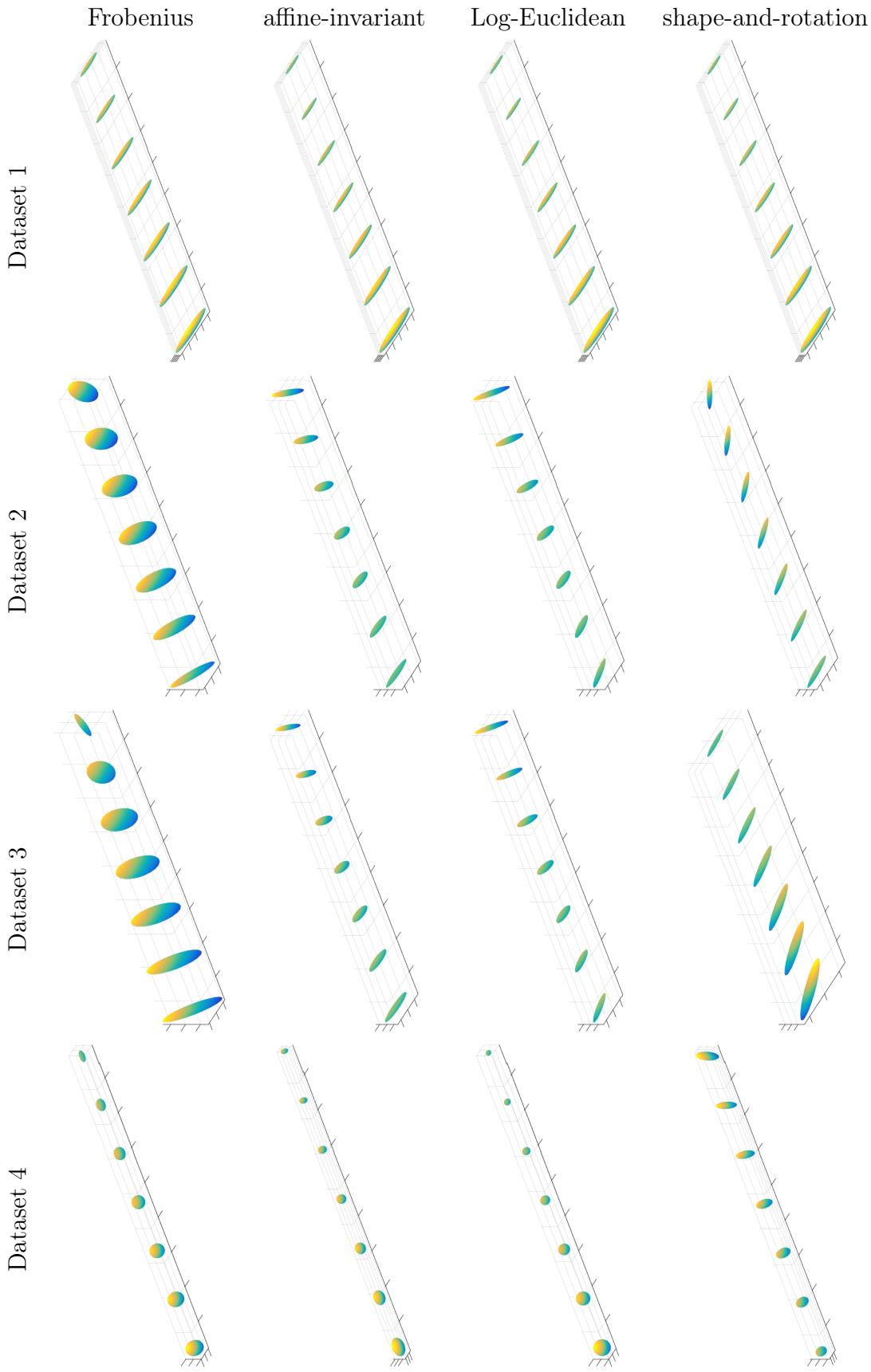


Figure 19: First geodesic principal components in four Riemannian metrics, for the four simulated datasets.

the Riemannian metrics do not capture tensor shape very well in the presence of high orientation variation. This is not surprising.

The shape-and-rotation metric is better at capturing the shape variation (or lack of it), as expected. However, we already know that due to its multiple representation of single tensors, it will overestimate variance and we know the results to be incorrect.

Note that while the shape and scale of the tensors in datasets 1 and 3 are identical, the variation of the shape and scale along the first geodesic principal component is quite different both between the metrics and the datasets. This indicates that one should think carefully about how to separate scale, orientation and shape for visualization of variance when using a geometric framework. For instance, factoring out orientation before applying the geodesic PCA as in dataset 1 changes the captured shape variance dramatically. This might be surprising.

6 Further related work

Tensor interpolation and geometric frameworks for analysis of tensors has been an active field of research for a number of years, and in the above we have only touched upon some of the most classical and most recent approaches to tensor interpolation. As an extension of the geometric framework, several approaches have appeared that utilize *divergences* [10, 29]. Divergences are not generally symmetric and therefore do not lead to a geodesic space in an obvious way. They can, however, be closely linked to Riemannian metrics, as in the case of the KL divergence, which infinitesimally coincides with the Fisher-Rao metric [3]. In the context of machine learning, kernel methods have also been proposed for SPD matrices [21]; however, these have been shown to just consist of Gaussian kernels on Euclidean features extracted from the SPD matrices [12]. Aside from tensors, approaches that try to separate rotation from other properties have also appeared [9, 1].

Several surveys and comparisons exist for geometries and interpolations for $\mathrm{Sym}^+(3)$. Moakher et al. [24] compare the affine-invariant metric to the closely related KL divergence in both geometric properties and in the context of statistics and visualization. Both Peeters et al. [25] and Zhou et al. [38] compare different distance/similarity measures for DTI, including several simple measures along with the affine-invariant and Log-Euclidean metrics representing the Riemannian approaches; the latter has a focus on regularization. What our survey has to offer in comparison is an extensive discussion of nonlinear geometries on $\mathrm{Sym}^+(3)$, a comparison with non-geometric approaches such as LIT [16], a thorough mathematical discussion of the current status of approaches that decompose shape and orientation – and a mathematical explanation why this is not trivial. Moreover, we offer publically available software[4] online to enable any reader to start working with geometries on $\mathrm{Sym}^+(3)$.

[4]Upon publication, the software will be available at `https://sites.google.com/site/aasaferagen/home/software`.

7 Discussion and conclusion

7.1 Geometry versus shape preservation

We have investigated a number of different geometric frameworks for tensor computation: The Frobenius, Wasserstein, affine-invariant and log-Euclidean metrics are all Riemannian metrics on $\mathrm{Sym}^+(3)$. However, they all exhibit the fattening effect, which can lead to unwanted smoothing effects when used for tensor interpolation or statistics. The shape-and-orientation metric is also a Riemannian metric. However, it comes with multiple representations of the same tensor, which leads to inconsistencies and unreliable statistics in practice, as illustrated by our experiments. This can be handled by factoring out the multiple representations in a quotient space. This quotient space is still a geodesic metric space, but it is no longer a Riemannian manifold – it has singularities and is a *stratified space* [26, 18]. The geodesics in this space could be used for interpolation, but a remaining problem for a practical investigation of its usefulness is to obtain an algorithm for computing the quotient metric geodesics. We expect this to be computationally demanding.

We have also reviewed frameworks that let go of geometry and simply ask for tensor interpolations that preserve shape well. These interpolations would very likely be good at preserving signal, but do not give the geometric framework of a geodesic metric space. In particular, this means that we do not have access to statistical approaches such as Fréchet means, hypothesis tests, geodesic PCA or regression, as we have with the more geometric approaches. This also makes interpolation of multiple tensors less well-defined, as we have already observed in Figures 16 and 17.

7.2 Why are second order tensors still interesting?

Second order tensors were key objects in DTI, but with the advent of HARDI imaging [32], they are often considered "too simple" to warrant further study. We argue the opposite. If you want to build geometries or tools that can handle the challenges of higher order tensors in diffusion-weighted imaging, these tools have better also perform well on second order tensors in DTI. Some of the most natural choices of metrics on the distributions returned by fODF model estimators such as constrained spherical deconvolution [31] or Q-ball [33], are the Wasserstein and Fisher-Rao metrics [17], as these are defined and theoretically well understood for general probability distributions. However, as we have seen in this survey, these metrics have unwanted smoothing effects when used for interpolation or statistics on second order tensors, and should not be expected to behave better on higher order tensors [9].

Another route to comparison of higher-order tensors comes through representations of higher order tensors based on second order tensors. Such representations are given by multi-compartment models [2] as well as through Finsler geometry, where any higher order ODF can be represented through a Finsler norm, and any Finsler norm can be represented as an orientation-dependent second order tensor [15, 6]. In both of these representations, a well-chosen metric for second order tensors may be extended to higher order tensors by integrating over orientation. This is an interesting direction for future work.

Finally, second order tensors are in 1-1 correspondence with covariance matrices, and any metric on second order tensors therefore also defines a metric on centered multivariate normal distributions. Statistics on probability distributions have many possible applications, from population statistics on uncertain tractography results represented as Gaussian Processes [27, 20] via evolutionary algorithms for optimization [19], to information geometry [3].

The quest for a descriptive geometric framework for $\mathrm{Sym}^+(3)$ therefore continues.

Acknowledgements

This research was supported by Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.

References

[1] *IEEE Conference on Computer Vision and Pattern Recognition, CVPR Workshops 2008, Anchorage, AK, USA, 23-28 June, 2008*. IEEE Computer Society, 2008.

[2] D. Alexander, G. Barker, and S. Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.

[3] S.-I. Amari and H. Nagaoka. *Methods of Information Geometry*. Transl. Math. Monogr., 2000.

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[5] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A Log-Euclidean framework for statistics on diffeomorphisms. In *MICCAI*, 2006.

[6] L. Astola and L. Florack. Finsler geometry on higher order tensor fields and applications to high angular resolution diffusion imaging. *International Journal of Computer Vision*, 92(3):325–336, 2011.

[7] P. G. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly. A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine*, 53(1):221–225, 2005.

[8] M. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*. Springer, 1999.

[9] H. ÇetingÃijl, B. Afsari, M. Wright, P. Thompson, and R. Vidal. Group action induced averaging for HARDI processing. In *ISBI*, 2012.

[10] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos. Jensen-bregman logdet divergence with application to efficient similarity search for covariance matrices. *IEEE transactions on pattern analysis and machine intelligence*, 35(9):2161–2174, 2013.

[11] D. B. Ennis and G. Kindlmann. Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images. *Magnetic Resonance in Medicine*, 55(1):136–146, jan 2006.

[12] A. Feragen, F. Lauze, and S. Hauberg. Geodesic exponential kernels: When curvature and linearity conflict. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2015.

[13] P. Fletcher, C. Lu, S. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *TMI*, 23:995–1005, 2004.

[14] P. T. Fletcher and S. Joshi. *Principal Geodesic Analysis on Symmetric Spaces: Statistics of Diffusion Tensors*, pages 87–98. 2004.

[15] L. Florack and A. Fuster. *Riemann-Finsler Geometry for Diffusion Weighted Magnetic Resonance Imaging*, pages 189–208. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[16] J. K. Gahm, N. Wisniewski, G. Kindlmann, G. L. Kung, W. S. Klug, A. Garfinkel, and D. B. Ennis. *Linear Invariant Tensor Interpolation Applied to Cardiac Diffusion Tensor MRI*, pages 494–501. 2012.

[17] A. Goh, C. Lenglet, P. Thompson, and R. Vidal. A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry. *NeuroImage*, 56(3):1181–1201, 2011.

[18] D. Groisser, S. Jung, and A. Schwartzman. Eigenvalue stratification and minimal smooth scaling-rotation curves in the space of symmetric positive-definite matrices. *arXiv preprint*, 2016.

[19] N. Hansen. The cma evolution strategy: a comparing review. In *Towards a new evolutionary computation*, pages 75–102. Springer Berlin Heidelberg, 2006.

[20] S. Hauberg, M. Schober, M. Liptrot, P. Hennig, and A. Feragen. A random riemannian metric for probabilistic shortest-path tractography. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Munich, Germany, Sept. 2015.

[21] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi. Kernel methods on the riemannian manifold of symmetric positive definite matrices. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 73–80. IEEE, 2013.

[22] S. Jung, A. Schwartzman, and D. Groisser. Scaling-rotation distance and interpolation of symmetric positive-definite matrices. *SIAM J. Matrix Analysis Applications*, 36(3):1180–1201, 2015.

[23] G. L. Kindlmann, R. S. J. EstÃľpar, M. Niethammer, S. Haker, and C.-F. Westin. Geodesic-loxodromes for diffusion tensor interpolation and difference measurement. In N. Ayache, S. Ourselin, and A. Maeder, editors, *MICCAI (1)*, volume 4791 of *Lecture Notes in Computer Science*, pages 1–9. Springer, 2007.

[24] M. Moakher and P. G. Batchelor. *Symmetric Positive-Definite Matrices: From Geometry to Applications and Visualization*, pages 285–298. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[25] T. Peeters, P. Rodrigues, A. Vilanova, and B. ter Haar Romeny. Analysis of distance/similarity measures for diffusion tensor imaging. In *Visualization and Processing of Tensor Fields*, pages 113–136. Springer Berlin Heidelberg, 2009.

[26] M. Pflaum. *Analytic and Geometric Study of Stratified Spaces*. Springer Lecture Notes in Mathematics, 2001.

[27] M. Schober, N. Kasenburg, A. Feragen, P. Hennig, and S. Hauberg. Probabilistic shortest path tractography in dti using gaussian process ode solvers. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Boston, USA, Sept. 2014.

[28] A. Schwartzman, R. F. Dougherty, and J. E. Taylor. False discovery rate analysis of brain diffusion direction maps. *Ann. Appl. Stat.*, 2(1):153–175, 03 2008.

[29] S. Sra. Positive definite matrices and the s-divergence. *Proceedings of the American Mathematical Society*, 144(7):2787–2797, 2016.

[30] A. Takatsu. Wasserstein geometry of gaussian measures. *Osaka J. Math.*, 48(4):1005–1026, 12 2011.

[31] J. Tournier, F. Calamante, D. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.

[32] D. Tuch, T. G. Reese, M. Wiegell, N. Makris, J. Belliveau, and V. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Res. Med.*, 48(4):577–582, 2002.

[33] D. S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, 2004.

[34] O. Tuzel, F. Porikli, and P. Meer. Region covariance: A fast descriptor for detection and classification. *European Conference on Computer Vision (ECCV)*, pages 589–600, 2006.

[35] Y. Wang, H. Salehian, G. Cheng, and B. C. Vemuri. Tracking on the product manifold of shape and orientation for tractography from diffusion MRI. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014*, pages 3051–3056, 2014.

[36] C.-F. Westin, S. Peled, H. Gudbjartsson, R. Kikinis, and F. A. Jolesz. Geometrical diffusion measures for MRI from tensor basis analysis. In *ISMRM '97*, page 1742, Vancouver Canada, April 1997.

[37] C. Zhang, T. Schultz, K. Lawonn, E. Eisemann, and A. Vilanova. Glyph-based comparative visualization for diffusion tensor fields. *IEEE Trans. Vis. Comput. Graph.*, 22(1):797–806, 2016.

[38] D. Zhou, I. L. Dryden, A. A. Koloydenko, K. M. Audenaert, and L. Bai. Regularisation, interpolation and visualisation of diffusion tensor images using non-euclidean statistics. *Journal of Applied Statistics*, 43(5):943–978, 2016.