INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Fast and Simple Computations on Tensors with Log-Euclidean Metrics.

Vincent Arsigny — Pierre Fillard — Xavier Pennec — Nicholas Ayache

N° 5584

May 2005

Thème BIO

Rapport de recherche

ISSN 0249-6399

Fast and Simple Computations on Tensors with Log-Euclidean Metrics.

Vincent Arsigny , Pierre Fillard , Xavier Pennec , Nicholas Ayache

Thème BIO — Systèmes biologiques
Projets Epidaure

Rapport de recherche n° 5584 — May 2005 — 42 pages

Abstract: Computations on tensors, i.e. symmetric positive definite real matrices in medical imaging, appear in many contexts. In medical imaging, these computations have become common with the use of DT-MRI. The classical Euclidean framework for tensor computing has many defects, which has recently led to the use of Riemannian metrics as an alternative. So far, only affine-invariant metrics had been proposed, which have excellent theoretical properites but lead to complex algorithms with a high computational cost. In this article, we present a new familly of metrics, called Log-Euclidean. These metrics have the same excellent theoretical properties as affine-invariant metrics and yield very similar results in practice. But they lead to much more simple computations, with a much lighter computational cost, very close to the cost of the classical Euclidean framework. Indeed, Riemannian computations become Euclidean computations in the logarithmic domain with Log-Euclidean metrics. We present in this article the complete theory for these metrics, and show experimental results for multilinear interpolation, dense extrapolation of tensors and anisotropic diffusion of tensor fields.

Key-words: Symmetric positive definite matrices, Medical Imaging, DT-MRI, Riemannian Metrics, PDEs, Interpolation, Extrapolation, Anisotropic Filtering.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Calculs simples et rapides sur les tenseurs avec les métriques Log-Euclidiennes.

Résumé : Les calculs sur les tenseurs, c'est-à dire les matrices réelles symétriques définies positives en l'imagerie médicale, apparaissent dans de nombreux contextes. En imagerie médicale, ces calculs sont devenus courants avec l'utilisation des images de tenseurs de diffusion obtenues par résonance magnétique nucléaire. Le cadre euclidien classique pour les calculs sur les tenseurs ayant de nombreux défauts, ceci a mené récemment à l'utilisation de métriques Riemanniennes comme alternative. Jusqu'à présent, seules des métriques affine-invariantes avaient été proposées, qui ont d'excellentes propriétés théoriques mais mènent à des algorithmes complexes et coûteux en temps de calcul. Dans cet article, nous présentons une nouvelle famille de métriques, appelées Log-Euclidiennes. Ces métriques ont les mêmes excellentes propriétés théoriques que les métriques affines-invariantes et produisent des résultats très similaires en pratique. Mais elles mènent à des calculs bien plus simples, avec un coût algorithmique bien plus faible, très proche du coût du cadre euclidien classique. En effet, les calculs Riemannens deviennent des calculs euclidiens dans le domaine logarithmique avec les métriques Log-Euclidiennes. Nous présentons dans cet article la théorie complète de ces métriques et nous donnons des résultats expérimentaux pour l'interpolation multilinéaire, l'extrapolation dense et la diffusion anisotrope de champs de tenseurs.

Mots-clés : Matrices symmétriques définies positives, imagerie médicale, DT-MRI, métriques Riemanniennes, EDPs, interpolation, extrapolation, filtrage anisotrope.

Contents

1 Motivation **4**

2 Overview of the Theory of Log-Euclidean Metrics **6**

3 Classical Properties of the Tensor Space **8**
3.1 Notations 9
3.2 Matrix Exponential 9
3.3 Algebraic Properties 11
3.4 Differential Properties 12
3.5 Compatibility Between Algebraic and Differential Properties 12

4 Log-Euclidean Metrics on the Tensor Space **13**
4.1 Multiplication in the Tensor Space 14
4.2 Tensor Space as a Lie group with the Logarithmic Multiplication 15
4.3 Log-Euclidean Metrics on the Tensor Lie group 16
4.4 A Vector Space Structure on the Tensor Space 19

5 Probabilities and Statistics with Log-Euclidean Metrics **20**
5.1 General Riemannian Statistical Framework 20
5.2 Random Tensors 21
5.3 Fréchet Means and Covariances with Log-Euclidean Metrics 21
5.4 General Log-Euclidean Statistical Framework 23

6 Comparison Between Log-Euclidean and Affine-Invariant Metrics **23**
6.1 Elementary Operations and Invariance 24
6.2 Log-Euclidean and Affine-Invariant Means 25
6.3 Geometric Interpolation of Determinants 25
6.4 Criterion for the Equality of the Two Means 27
6.5 Larger Anisotropy in Log-Euclidean Means 27

7 Application to Interpolation, Dense Extrapolation and Anisotropic Filtering **29**
7.1 Multilinear Interpolation 30
7.2 Solving PDEs in a Riemannian Tensor Space 31
7.3 Dense Extrapolation with Harmonic Diffusion 32
7.4 Anisotropic Filtering 33

8 Conclusions **34**

1 Motivation

Tensors, i.e. symmetric positive definite real matrices, appear in many contexts in medical imaging: Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) [1], modeling of anatomical variability, statistics on registration results... They are also a general tool in image analysis, especially for segmentation, grouping, motion analysis and texture segmentation [4]. Furthermore, they appear in other research fields such as solid mechanics or fluid dynamics, where their visualization and processing is needed [22]. In order to carry out computations with these objects, one needs to define the distance between two tensors and a complete operational framework. In particular, this is necessary when one wishes to generalize to the tensor case the usual tools used on scalar or vector-valued images such as anisotropic filters, geodesic active contours [7] and other Partial Differential Equations-(PDEs) based methods to restore, enhance, segment tensor-valued images. One can also generalize many statistical tools when a differential geometry approach is adopted, as shown in [15]. Otherwise, one has to rely solely on operations performed on *features* extracted from tensors like first eigenvectors [6], rotations matrices [24], which unfortunately cannot take into account all the information carried by tensors.

The tensor space is part of the vector space of square matrices. So one can directly use a Euclidean structure on square matrices to define a metric on the tensor space. For instance, one can use the following Euclidean distance:

$$d(S,T) = \left(Trace((S-T)^2)\right)^{\frac{1}{2}}. \tag{1}$$

This is straightforward to use and leads *a priori* to simple computations. But this framework is practically and theoretically unsatisfactory for three main reasons.

First, this framework has a boundary problem: it does not prevent the appearance of singular matrices during computations. There are many good reasons why they should not appear. Physically, a diffusion exactly equal to zero in DT-MRI is impossible. It can be very small, but definitely *not* zero. As long as the temperature is higher than 0 Kelvin, molecules *will* move in *all* directions. Mathematically, a zero eigenvalue in a covariance matrix implies absolute certainty in a direction, which is either physically impossible or implies that the dimensionality of the space where the random vector lives is too large. In this context, a null eigenvalue is the sign of error or improper modeling.

Typically this problem appears as soon as one begins to solve Ordinary Differential Equations (ODEs) or even worse PDEs. This is the case for the estimation of tensors from raw DT-MRI data, or the regularization of tensors fields, the dense extrapolation of sparse tensor data... When these equations are being solved, negative eigenvalues appear and a re-projection step is necessary. Even with an orthogonal projection on non-negative symmetric matrices, matrices with null eigenvalues cannot easily be suppressed. This problem has to be addressed, because solving PDEs on tensor fields is a crucial issue in applications. In [5], geometric constrains (i.e. positive definiteness) are integrated in the numerical schemes used to run PDEs and non-positive eigenvalues are *a priori* avoided. However, the mathematical analysis of issues such as existence of solutions of such PDEs is still a research subject, and these evolutions could blow up in finite time, i.e. go out of the tensor space in finite time.

Second, a tensor corresponds typically to a *covariance matrix*. The value of its determinant is a measure of the dispersion of a Gaussian random vector which has the same covariance. The classical linear interpolation of two tensors leads in most cases to an interpolation of the determinants which is *not* satisfactory: a global maximum is usually reached between the two tensors. This means that the associated interpolated Gaussian random vector has a strictly larger dispersion at some point between the two extremities. In the case of DT-MRI, this means that the interpolated diffusion process can be strictly larger between the two extremities! This is a serious defect, and it a monotonic interpolation of determinants would be much more preferable. We will refer to this effect as *tensor swelling*. It appears in most attempts to restore or filter tensor-valued images with Euclidean metrics. See for instance the smoothing results obtained in [7] on page 17, or the results given by the rank/signature preserving flow in [5]. To address this defect, it was proposed in [23, 24, 5] to modify *only* the orientation of tensors and *not* their eigenvalues. This is not completely satisfactory since one would also wish to restore eigenvalues which are subject to noise like orientations.

Third, the Euclidean metric is also unsatisfactory in terms of symmetry. The Euclidean mean for N tensors $S_1, .., S_N$ is given by:

$$\mathbb{E}(S_1, .., S_N) = \frac{1}{N}\left(\sum_{i=1}^{N} Si\right). \tag{2}$$

But this *arithmetic* mean does not yield the identity for a tensor and its inverse! If tensors model variability, one would like in many cases to take a *geometric* mean. An isotropic variance of σ^2 and another isotropic variance of $1/\sigma^2$ should in many applications cancel out and give an isotropic mean of variance 1. This symmetry principle can be expressed in the following way: *the inversion mapping $S \mapsto S^{-1}$ should be an isometry.* In particular, the mean of two tensors S and S^{-1} should be the identity.

To circumvent these difficulties, other metrics have been proposed for the tensor space, this time not Euclidean but Riemannian. In the approach proposed in [16, 8, 13, 14] , the tensor space has no boundaries: null eigenvalues are at infinite distance. The symmetry principle is respected. Thus, the shortcomings mentioned above are overcome and moreover affine-invariance is obtained. This means that computations are invariant with respect to an affine change of coordinates. Last but not least, the swelling effect typical of Euclidean calculus on tensors has disappeared in this Riemannian framework. These theoretical properties are excellent, but the computational burden is high. This is essentially due to the curvature of the Riemannian space: the computation of geodesics and energy gradients all have substantial corrections with respect to their usual Euclidean counterparts. In spite of these technicalities, a fully operational framework for PDEs is presented in [16], exemplified with dense extrapolation and anisotropic filtering of tensors.

We propose a new Riemannian framework to fully overcome these computational limitations while conserving the same excellent theoretical properties. Contrary to the affine-invariant case, we obtain a space with a null curvature. The new Riemannian metrics described in this work are called **Log-Euclidean metrics**. The complexity of associated computations is almost identical to those of the classical Euclidean framework!

The key idea of this approach is to define two new structures on tensors. First a *Lie group structure* , i.e. a group structure in which algebraic operations like inversion and multiplication are smooth mappings. Second, a *Vector Space structure* which complements the Lie group structure with a new scalar multiplication. To our knowledge, these structures are completely new in the literature and shed a new light on the tensor space. The Riemannian framework associated to the biinvariant metrics on this novel structure is particularly simple. It results in classical Euclidean computations in the logarithmic domain. This is why we have baptized such metrics *Log-Euclidean.* In particular, the Riemannian Fréchet mean is simply obtained by computing the arithmetic mean in the logarithm domain and mapping it back to the tensor domain with the matrix exponential.

Theoretically, this leads to excellent invariance and interpolation properties, very similar to those of affine-invariance metrics. In terms of computational efficiency, the difference is very significant. Experimentally, in the applications studied in this article, computations with Log-Euclidean metrics are at least 6 times faster than in the affine-invariant case. For the resampling of tensor fields, Log-Euclidean computations can even be 50 times faster or even more.

Let us now detail the contents of this article. In Section 2, we give an overview of the new results obtained here. These results are easier to grasp when considered globally, without the technicalities on which they are based. Then we present in Section 3 the fundamental theoretical properties of tensors we rely on afterwards. Next, we proceed to the theory of the Log-Euclidean metrics in Section 4, which shows how one can construct a Lie group and a Vector Space structures on tensors. On these structures, a special class of metrics called Log-Euclidean are particularly straightforward to use. In Section 5, we present the associated probabilistic and statistical framework, in which Riemannian notions on tensors coincide with Euclidean notions on their logarithms. In Section 6, we highlight the differences and similarities between the Log-Euclidean framework presented in Section 4 and the recently introduced affine-invariant framework. Then we turn to applications in Section 7 where we compare results in interpolation, extrapolation and anisotropic filtering obtained with the Log-Euclidean and affine-invariant frameworks. The results obtained are very similar, but are obtained much faster in the Log-Euclidean case and are also much simpler to implement.

2 Overview of the Theory of Log-Euclidean Metrics

The theory developed in the next sections is based on quite sophisticated mathematical materials. In this section, we summarize the main results obtained in Sections 4, 5, 6 and 7. This way, the reader can have a global view on the theory developed in this article.

Existence and Uniqueness of the Logarithm As shown in Section 4, to a tensor S is associated a unique *logarithm* L which is *symmetric.* It verifies $S = \exp(L)$ where exp is the matrix exponential. To each symmetric matrix is associated by the exponential a tensor. In an orthonormal basis in which S is diagonal, L is obtained by transforming the eigenvalues of S into their scalar logarithm.

A Vector Space Structure on Tensors Since their is a one-to-one mapping between the tensor space, denoted $Sym_\star^+(n)$ and the space of symmetric matrices $Sym(n)$, one can give $Sym_\star^+(n)$ a *vector space structure* by transporting the addition + and the scalar multiplication . on $Sym_\star^+(n)$ with the exponential.

This defines on $Sym_\star^+(n)$ the *logarithmic multiplication* $\odot$ and the *logarithmic scalar multiplication* $\circledast$, given by:

$$\begin{cases} S_1 \odot S_2 := \exp\left(\log(S_1) + \log(S_2)\right) \\ \lambda \circledast S := \exp\left(\lambda . \log(S)\right) = S^\lambda . \end{cases} \tag{3}$$

The logarithmic multiplication is *commutative* and coincides with the matrix multiplication whenever the two tensors S_1 and S_2 commute in the matrix sense.

With $\odot$ and $\circledast$, the tensor space has by construction a *vector space structure*. One should note that this is *not* the usual vector space structure derived from the addition and scalar multiplication on square matrices. With the latter structure, the tensor space is *not* a vector space, whereas with this new structure it *is* a vector space. The notion of vector space depends on the structure one considers, and not only on the space itself.

Log-Euclidean Metrics When one considers only the multiplication $\odot$ on the tensor space, one has a *Lie group* structure, i.e. a space which is both a hypersurface and a group in which algebraic operations are *smooth mappings*. This viewpoint is important, because the theory of Riemannian metrics can be applied to Lie groups to define metrics, i.e. distances, in a framework where usual analysis and statistical tools can be generalized (see [15]).

Among Riemannian metrics in Lie groups, the most convenient in practice, when they exist, are *biinvariant metrics*, i.e. distances that are invariant by multiplication and inversion. For the tensor Lie group, biinvariant metrics exist and are particularly simple. We have baptized such metrics *Log-Euclidean* metrics, since they correspond to Euclidean metrics in the logarithmic domain. From a Euclidean norm $\|.\|$ on $Sym(n)$, they can be written:

$$d(S_1, S_2) = \|\log(S_1) - \log(S_2)\|. \tag{4}$$

Boundary Problems and Symmetry Principle Contrary to the classical Euclidean framework on tensors, one can clearly see from Eq. (4) that matrices with null or negative eigenvalues are at an infinite distance from tensors and will not appear in practical computations.

Moreover, distances are not changed by inversion, i.e. the *symmetry principle* is verified. As a consequence, the Log-Euclidean mean between tensors will be a generalization of the *geometric mean* and not of the arithmetic mean as in the classical Euclidean case. This is crucial for instance for a correct interpolation of determinants when two tensors are interpolated.

Invariance by Similarity Among all Log-Euclidean metrics on the tensor space, some are invariant by similarity (rotation plus scaling). This means that if tensors are covariance matrices, computations on tensors using such a metric will be invariant with respect to a change of coordinates obtained by a similarity. A similarity-invariant Log-Euclidean metric (maybe the simplest one!) is given by:

$$d(S_1, S_2) = \left(\text{Trace}\left(\{\log(S_1) - \log(S_2)\}^2\right)\right)^{\frac{1}{2}}. \tag{5}$$

Euclidean Calculus in the Logarithmic Domain The tensor vector space with a Log-Euclidean metric is in fact isomorphic and *isometric* with the corresponding Euclidean space of symmetric matrices. As a consequence, the Riemannian framework for statistics and analysis is extremely simplified, much simpler than in [8, 13, 14]. In particular, the Log-Euclidean mean of N tensors is given by:

$$\mathbb{E}_{LE}(S_1, ..., S_N) = \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(S_i)\right). \tag{6}$$

This is remarkable: in this framework, the interpolation, extrapolation, anisotropic diffusion, etc. of tensors can simply be performed in a Euclidian way in the logarithm domain, and final results are mapped back to the tensor domain with the exponential. Hence, statistical tools or PDEs can be generalized to tensors very readily. Mathematical issues such as existence and uniqueness of such PDEs on tensors are simply particular cases of the classical PDE theory on vector fields [19]. This is a striking result since so far the theory of the existence and uniqueness of general PDEs on tensors is still an open research field [5].

Comparison with Affine-Invariant Metrics In [16], a complete framework for the use of affine-invariant metrics on tensors is presented. Log-Euclidean and affine-invariant metrics yield very similar results in applications like resampling, dense extrapolation of sparse data or anisotropic filtering (see Section 7). The only difference is slightly more anisotropic tensors in the Log-Euclidean results. The real differences are in the computational cost and the simplicity of implementation. Indeed, Log-Euclidean computations are faster by a factor of at least 6 in the applications considered in this article. Moreover, all Log-Euclidean computations are *simple* since Euclidean in the logarithmic domain, and do not use curvature corrections as in the affine-invariant case.

3 Classical Properties of the Tensor Space

We begin with a description of the fundamental properties and tools used in this work. First, we recall the elementary properties of the matrix exponential. Then we examine the general properties of tensors. These properties are of two types: algebraic and differential. On the one hand, tensors have algebraic properties because they are a special kind of invertible matrices, and on the other hand they can be considered globally as a hypersurface (or *manifold*) and therefore have differential geometry properties. These properties are not independent: on the contrary, they are compatible in a profound way. This compatibility is the core of the approach developed here.

3.1 Notations

Definition 1. *We will use the following definitions and notations:*

- *We call tensor space of dimension* n *and write* $Sym_\star^+(n)$ *the space of real* $n*n$ *matrices that are symmetric positive and definite.*
- *We write* $Sym(n)$ *the vector space of real* $n*n$ *symmetric matrices.*
- $GL(n)$ *is the group of real invertible* $n*n$ *matrices.*
- $M(n)$ *is the space of real* $n*n$ *square matrices.*
- $Diag(\lambda_1, ..., \lambda_n)$ *will be the diagonal matrix constructed with the real values* $(\lambda_i)_{i\in 1...n}$ *in its diagonal.*
- *For any square matrix* M, $Sp(M)$ *is the spectrum of* M, *i.e. the set of its eigenvalues.*
- *Let* $\phi : E \to F$ *be a differentiable mapping between two smooth manifolds. Its differential at a point* $m \in E$ *acting on a infinitesimal displacement* dM *in the tangent space to* E *at* m *is written:* $D_m\phi.dM$.

The denomination "Tensor Space" comes from DT-MRI ([1]), where diffusion covariance matrices are called tensors.

3.2 Matrix Exponential

The exponential plays a central role in *Lie groups* (see [3, 21, 9]). We will consider here only the matrix version of the exponential, which is a tool that we extensively use in the next sections. We recall its definition, and give its elementary properties. Last but not least, we give the Baker-Campbell-Hausdorff formula. It is a very fine tool that provides deep insights in Lie Groups. We will find it highly useful to compare Log-Euclidean means to affine-invariant means in Sec. 6.

Definition 2. *The exponential* $\exp(M)$ *of a matrix* M *is given by* $\exp(M) = \sum_{n=0}^{\infty} \frac{M^k}{k!}$. *Let* $G \in GL(n)$. *If there exists* $M \in M(n)$ *such that* $G = \exp(M)$, *then* M *is said to be a logarithm of* N *and we write* $G = \log(M)$.

In general, the logarithm of a real invertible matrix may not exist, and if it exists it may not be unique. The lack of existence is a general phenomenon in connected Lie groups. One generally needs *two* exponentials to reach every element [25]. The lack of uniqueness is essentially due to the influence of rotations: rotating of an angle α is the same as rotating of an angle $\alpha + 2k\pi$ where k is an integer. Since the logarithm of a rotation matrix directly depends on its rotation angles (one angle suffices in 3D, but several angles are necessary when $n > 3$), it is not unique. However, when a matrix is sufficiently close to the identity, then a logarithm always exists, and moreover among all existing logarithms, there always exists one with smallest norm, called *principal logarithm.* It is given by the following series:

$$\log(Id + M) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} M^k. \tag{7}$$

This series is absolutely convergent for $\|M\| < 1$ where $\|.\|$ is an algebra norm, i.e. verifying $\|M.N\| \leq \|M\|.\|N\|$ for all square matrices $M,\ N$.

Theorem 1. $\exp : M(n) \to GL(n)$ *is a* $\mathcal{C}^\infty$ *mapping. Its differential map at a point* $M \in M(n)$ *acting on an infinitesimal displacement* $dM \in M(n)$ *is given by:*

$$D_M \exp .dM = \sum_{k=1}^{\infty} \frac{1}{k!} \left(\sum_{l=0}^{k-1} M^{k-l-1}.dM.M^l \right). \tag{8}$$

Proof. The smoothness of exp is simply a consequence of the uniform absolute convergence of its series expansion in any compact set of $M(n)$. The differential is obtained classically by a term by term derivation of the series defining the exponential.

We see here that the non-commutativity of the matrix multiplication complicates seriously the differentiation of the exponential, which is much simpler in the scalar case. However, taking the trace in Eq. (8) yields:

Corollary 1. *Since* $\forall M,\ N \in M(n) :\ Trace(M.N) = Trace(N.M)$*, we have:*

$$Trace(D_M \exp .dM) = Trace(\exp(M).dM). \tag{9}$$

We will also use in the following this property on determinants:

Proposition 1. *Let* $M \in M(n)$*. Then* $det(\exp(M)) = \exp(Trace(M))$*.*

Proof. This is easily seen in terms of eigenvalues of M. The Jordan decomposition of M [12] assure that Trace(M) is the sum of its eigenvalues. But the exponential of a triangular matrix transforms the diagonal values of this matrix into their scalar exponential. The determinant of $\exp(M)$ is simply the product of its eigenvalues, which is precisely the exponential of the trace of M.

Baker-Campbell-Hausdorff Formula. *See [11]. We will only state the theorem in the matrix case. Let* $M, N \in M(n)$ *and* $t \in \mathbb{R}$*. When* t *is small enough, we have the following development, in which the logarithm used is the principal logarithm:*

$$\begin{aligned}\log(\exp(t.M).\exp(t.N)) \; = \; & t.(M+N) + t^2/2([M,N]) \\ & +t^3/12([M,[M,N]]+[N,[N,M]]) \\ & +t^4/24([[M,[M,N]],N]) + O(t^5).\end{aligned} \quad (10)$$

We recall that $[M,N] = MN - NM$ *is the Lie Bracket of* M *and* N.

The Baker-Campbell-Hausdorff formula shows how much $\exp(\log(M).\log(N))$ deviates from $M+N$ due to the non-commutativity of the matrix product. What is remarkable is that the deviation can be expressed only in terms of Lie Brackets between M and N.

3.3 Algebraic Properties

Tensors have very remarkable algebraic properties. First, there always exists a unique *symmetric*[1] matrix logarithm for a tensor. Second, if the tensor space is not a subgroup of $GL(n)$, it is stable with respect to inversion. Moreover, its spectral decomposition is particularly simple[2].

Theorem 2. *For any* $S \in Sym(n)$*, there exists an orthonormal coordinate system in which* S *is diagonal. This is in particular the case for tensors.* $Sym_\star^+(n)$ *is not a subgroup of* $GL(n)$*, but it is stable by inversion. Moreover, the matrix exponential* $\exp : Sym(n) \rightarrow Sym_\star^+(n)$ *is one-to-one.*

Proof For the first part, see elementary linear algebra manuals, or [12]. For the second part, see for example [16] chapter 3.

Thanks to the existence of an orthonormal basis in which a tensor (resp. a symmetric matrix) is diagonal, the logarithms (resp. the exponential) has a particularly simple expression. In such a basis, taking the $\log$ (resp. the $\exp$) of a is simply done by applying their scalar version to eigenvalues:

$$\begin{cases} \log(R.\mathrm{Diag}(\lambda_1,...,\lambda_N).R^T) = R.\mathrm{Diag}(\log(\lambda_1),...,\log(\lambda_N)).R^T \\ \exp(R.\mathrm{Diag}(\lambda_1,...,\lambda_N).R^T) = R.\mathrm{Diag}(\exp(\lambda_1),...,\exp(\lambda_N)).R^T. \end{cases}$$

These formulae make numerical computations of logarithms and exponentials here particularly efficient as compared to those of general matrices. Moreover, this viewpoint gives also a straightforward method for computing inverses: it suffices to inverse the eigenvalues in an orthonormal basis to obtain the inverse.

[1] On the contrary, an infinity of non-symmetric matrix logarithms may exist for a tensor. For example, the identity has an infinity of *skew* matrix logarithms, since it can be viewed as a rotation whose angles are arbitrary multiples of 2π.

[2] This is due to the fact that tensors are *normal operators*, like rotations and antisymmetric matrices [12].

3.4 Differential Properties

This time from the viewpoint of topology and differential geometry, the tensor space has also many particularities.

Proposition 2. *$Sym_\star^+(n)$ is an open convex half-cone of $Sym(n)$ and is therefore a submanifold of $Sym(n)$.*

Proof. First, let us recall that a symmetric matrix is *positive definite* if and only if for each vector $x \neq 0$, $x^T.S.x > 0$. With this in mind, let us consider one assertion at a time:

- $Sym_\star^+(n)$ is an open. Let $\|.\|$ be a norm on $\mathbb{R}^n$. Let $\|.\|$ be the associated norm on matrices, seen as linear forms on $\mathbb{R}^n$: $\|M\| = \sup_{\|x\| \leq 1} \|M.x\|$. For any tensor S, we have: $|x^T.S.x| \geq \lambda_{\min}\|x\|^2$ where $\lambda_{\min} := \min_{\lambda \in Sp(T)} \lambda > 0$. As a consequence, as long as a symmetric matrix dS is such that $\|dS\| < \lambda_{\min}$, we have for $x \neq 0$: $x^T.(S + dS).x > (\lambda_{\min} - \|dS\|)\|x\|^2 > 0$. This means that $S + dS$ is still a tensor.
- $Sym_\star^+(n)$ is convex: let $0 < \lambda < 1$ and $S_1,\ S_2 \in Sym_\star^+(n)$. Then $(1 - \lambda)S_1 + \lambda S_2$ is still positive definite.
- $Sym_\star^+(n)$ is a half-cone: $\forall \lambda > 0$ and $S \in Sym_\star^+(n)$, λS is positive definite.
- $Sym_\star^+(n)$ is a submanifold of $Sym(n)$: every open subset of a manifold is a submanifold.

This allows us to think of tensors as belonging to a smooth manifold. Thus $Sym_\star^+(n)$ is some kind of hypersurface, whose dimension is $n(n + 1)/2$ like that of $Sym(n)$.

3.5 Compatibility Between Algebraic and Differential Properties

We have seen that exp is a smooth bijection. But the logarithm, i.e. its inverse, is also smooth. As a consequence, all the algebraic operations on tensors presented before are also smooth, in particular the inversion. Thus the two structures are fully compatible.

Theorem 3. $\log : Sym_\star^+(n) \to Sym(n)$ *is* $\mathcal{C}^\infty$*. Thus* exp *and its inverse* log *are both smooth, i.e. they are diffeomorphisms. This is due to the fact that the differential of* exp *is nowhere singular.*

Proof In fact, we only need to prove the last assertion. If it is true, the Implicit Function Theorem [20] applies and assures that log is also smooth. Since the differential of exp at 0 is simply given by the identity, it is invertible by continuity in a neighborhood of 0. We now show that this propagates to the entire space $Sym(n)$. Indeed, let us then suppose that for a point M, the differential $D_{M/2}\exp$ is invertible. We claim that then $D_M \exp$ is also invertible, which suffices to prove the point. To show this, let us take $dM \in Sym(n)$ such that $D_M \exp .dM = 0$. If $D_M \exp$ is invertible, we should have $dM = 0$. To see this, remark that $\exp(M) = \exp(M/2).\exp(M/2)$. By differentiation and applying to dM, we get:

$$D_M \exp .dM = 1/2((D_{M/2} \exp .dM).\exp(M/2) + \exp(M/2).(D_{M/2} \exp .dM)) = 0.$$

This implies by multiplication by $\exp(-M/2)$:

$$\exp(-M/2)(D_{M/2} \exp .dM).\exp(M/2) + (D_{M/2} \exp .dM) = 0.$$

Since $A^{-1}.\exp(B).A = \exp(A^{-1}.B.A)$ we have also by differentiation $A^{-1}.D_B \exp(dB).A = D_B \exp(A^{-1}.dB.A)$. Using this simplification and the hypothesis that $D_{M/2} \exp$ is invertible, we obtain:

$$\exp(-M/2).dM.\exp(M/2) + dM = 0.$$

Let us transpose this equation in an orthonormal basis in which M is diagonal with a rotation matrix R. Let (λ_i) be the eigenvalues of M. We get:

$$dN := R.dM.R^T = -\mathrm{Diag}(\exp(-\lambda_1/2), ..., \exp(-\lambda_N/2)).dN.\mathrm{Diag}(\exp(\lambda_1/2), ..., \exp(\lambda_N/2)).$$

Coordinate by coordinate this writes:

$$\forall i, j : dN_{i,j}(1 + \exp(-\lambda_i/2 + \lambda_j/2)) = 0.$$

Hence $\forall i, j :\ dN_{i,j} = 0$ and thus $dM = 0$. And we are done.

Corollary 2. *In the tensor space, for all $\alpha \in \mathbb{R}$, the power mapping: $S \mapsto S^\alpha$ is smooth. In particular, this is true for the inversion mapping (i.e. when $\alpha = -1$).*

Proof. We have $S^\alpha = \exp(\alpha \log(S))$. The composition of smooth mappings is smooth.

4 Log-Euclidean Metrics on the Tensor Space

We focus in this section on the description of a two new structures on tensors.

The first is a Lie group structure, i.e. an algebraic group structure that is *compatible* with the differential structure of the Tensor Space. This means that inversion and multiplication are smooth mappings in this space. Moreover, this new algebraic structure naturally complements the usual algebraic properties of tensors: the new inverse is the classical matrix inverse, the new product of two tensors coincides with the matrix product when the two tensors commute, and last but not least the Lie group exponential coincides with the matrix exponential! The key idea of this approach is the novel notion of *logarithmic multiplication* of two tensors. We will not recall the definition of Lie groups and their elementary properties. For a complete account on Lie group theory, see [3]. This theory is quite arduous, because it is part of both advanced algebra and differential geometry.

The second structure is a *vector space structure*. Indeed, one can define a *logarithmic scalar multiplication* that complements the Lie group structure to form a vector space structure on the tensor space.

We present all biinvariant metrics on the Tensor Lie group, and show that among them there are metrics invariant by similarity. These metrics correspond to classical Euclidean metrics in the logarithmic domain, and we have therefore baptized them *Log-Euclidean metrics*.

The very striking result obtained in this section is that this Lie Group and Vector Space approach leads to extremely simple computations. In practice, Riemannian computations are Euclidean computations in the logarithmic domain. With a Log-Euclidean metric, the tensor space is isomorphic, diffeomorphic and isometric to the associated Euclidean vector space of symmetric matrices. This means that these two spaces with the associated structures are essentially identical.

The Riemannian framework associated to Log-Euclidean metrics fully meets the requirement stated in Section 1. First, symmetric matrices with null or negative eigenvalues are at infinite distance. Second, the *symmetry principle* is respected: the inversion of tensors does not change the metric.

4.1 Multiplication in the Tensor Space

It is definitely not obvious how one could define a multiplication on tensor spaces compatible with previous algebraic and differential properties. Intuitively, the question we ask here is: how can we combine smoothly two tensors to make a third one, in such a way that Id is still the identity and the usual inverse is remains its inverse? Moreover, if we obtain a new Lie group structure, we would also like the matrix exponential to be the exponential associated to the Lie group structure, which *a priori* can be different.

First Idea: Matrix Multiplication The first idea that comes to mind is to use matrix multiplication. But then the non-commutativity of matrix multiplication between tensors stops the attempt: if $S_1, S_2 \in Sym_\star^+(n)$, $S_1.S_2$ is a tensor (or equivalently, is symmetric) if and only if S_1 and S_2 commute.

Second Idea: Forcing Symmetry in the Matrix Product To overcome the possible asymmetry of the matrix product of two tensors, one can simply take the *symmetric part* of the product and define the new product $\diamond$:

$$S_1 \diamond S_2 := \frac{1}{2}(S_1.S_2 + S_2.S_1).$$

This product is smooth, conserves the identity and the inverse, but it is *not associative*! Indeed, we have in general:

$$(S_1 \diamond S_2) \diamond S_3 \neq S_1 \diamond (S_2 \diamond S_3).$$

This makes everything fail, because associativity is an essential requirement of group structure. Without it, many fundamental properties disappear. For Lie groups, the notion of adjoint representation does not exist anymore without associativity.

Third Idea: Inspiration from the Affine-Invariant Metrics In [16], the distance between two tensors S_1, S_2 is given by:

$$d(S_1, S_2) = \| \log(S_1^{-1/2}.S_2.S_1^{-1/2}) \|. \tag{11}$$

where $\|.\|$ is a Euclidean norm defined on $Sym(n)$. Let us define the following multiplication $\circledcirc$:

$$S_1 \circledcirc S_2 := S_1^{1/2}.S_2.S_1^{1/2}.$$

With this multiplication, the affine-invariant metric constructed in [16] can be interpreted then as a left-invariant metric. Moreover, this multiplication is smooth, compatible with matrix inversion and matrix exponential. Everything works fine, except that it is also *not associative*! Again this comes from the non-commutativity of the matrix multiplication on tensors.

Thus, we see that defining a group structure directly from the matrix multiplication is difficult.

4.2 Tensor Space as a Lie group with the Logarithmic Multiplication

A Multiplication Based on Logarithms Theorem 3 points a very important fact: $Sym_\star^+(n)$ is diffeomorphic to its tangent space at the identity, $Sym(n)$. But $Sym(n)$ has an additive group structure, and to obtain a group structure on the tensor space, one can simply transport the additive structure of $Sym(n)$ to $Sym_\star^+(n)$ with the exponential. More precisely, we set:

Definition 3. *Let $S_1, S_2 \in Sym_\star^+(n)$. We define their* **logarithmic multiplication** $S_1 \odot S_2$ *by:*

$$S_1 \odot S_2 := \exp(\log(S_1) + \log(S_2)). \tag{12}$$

Proposition 3. *$(Sym_\star^+(n), \odot)$ is a group. The neutral element is the usual identity matrix, and the inverse of a tensor is the inverse in the matrix sense. Moreover, whenever two tensors commute in the matrix sense, then the logarithmic product is equal to their matrix product. Last but not least, the multiplication is* **commutative**.

Proof. The multiplication is defined by addition on logarithms. It is therefore associative and commutative. Since $\log(Id) = 0$, the neutral element is Id and since $\log(S^{-1}) = -\log(S)$, the new inverse is the matrix inverse. Finally, we have $\exp(\log(S_1) + \log(S_2)) = \exp(\log(S_1)).\exp(\log(S_2)) = S_1.S_2$ when $[S_1, S_2] = 0$.

Theorem 4. *The logarithmic multiplication $\odot$ on $Sym_\star^+(n)$ is compatible with its structure of smooth manifold: $(S_1, S_2) \mapsto S_1 \odot S_2^{-1}$ is $\mathcal{C}^\infty$. Therefore, $Sym_\star^+(n)$ is given a commutative Lie group structure by $\odot$.*

Proof. $(S_1, S_2) \mapsto S_1 \odot S_2^{-1} = \exp(\log(S_1) - \log(S_2))$. But since exp and log and the addition are smooth, their composition is also smooth. By definition (see [9], page 29), $Sym_\star^+(n)$ is a Lie group.

Definition 4. $(Sym_\star^+(n), \odot)$ *is called the Tensor Lie group.*

Proposition 4. $\exp : (Sym(n), +) \to (Sym_\star^+(n), \odot)$ *is a Lie group isomorphism. In particular, one-parameter subgroups of $Sym_\star^+(n)$ are obtained by taking the matrix exponential of those of $Sym(n)$, which are simply of the form $(t.V)_{t\in\mathbb{R}}$ where $V \in Sym(n)$. As a consequence, the Lie group exponential in $Sym_\star^+(n)$ is given by the classical matrix exponential on the Lie Algebra $Sym(n)$.*

Proof. We have explicitly transported the group structure of $Sym(n)$ into $Sym_\star^+(n)$ so exp is a morphism. It is also a bijection, and thus an isomorphism. The smoothness of exp then assures its compatibility with the differential structure.

Let us recall the definition of one-parameter subgroups. $(S(t))_{t\in\mathbb{R}}$ is such a subgroup if and only if we have $\forall t, s : S(t+s) = S(t) \odot S(s) = S(s) \odot S(t)$. But then $\log(S(t+s) = \log(S(t) \odot S(s)) = \log(S(t)) + \log(S(s))$ by definition of $\odot$. Therefore $\log S(t)$ is also a one-parameter subgroup of $(Sym(n), +)$, which is necessarily of the form $t.V$ where $V \in Sym(n)$. V is the *infinitesimal generator* of $S(t)$. Finally, the exponential is obtained from one-parameter subgroups, which are all of the form $(\exp(t.V))_{t\in\mathbb{R}}$ (see [21], Chap. V).

Thus we have given the tensor space a structure of Lie group that leaves unchanged the classical matrix notions of inverse and exponential. The new multiplication used, i.e. the logarithmic multiplication, generalizes the matrix multiplication when two tensors do not commute in the matrix sense.

The associated Lie Algebra is the space of symmetric matrices, which is diffeomorphic and isomorphic to the group itself. These technical terms mean that the two spaces have essentially the same algebraic and differential structures.

The reader should note that this Lie group structure is to our knowledge completely new in the literature. For a space as commonly used as tensors, this is on the face of it very surprising. This probably comes from the fact that the Tensor Lie group is *not* a multiplicative matrix group, contrary to most Lie groups.

4.3 Log-Euclidean Metrics on the Tensor Lie group

Now that we have given $Sym_\star^+(n)$ a structure of Lie group, we turn to the task of exploring metrics compatible with this new structure. The framework of Riemannian metrics is very powerful: it provides a simple framework for statistics in manifolds. See [15] for a detailed account on this topic.

Among Riemannian metrics in Lie groups, *biinvariant* metrics are the most convenient. We have the following definition:

Definition 5. *A metric $<,>$ defined on a Lie group G is said to be biinvariant if $\forall m \in G$, the left- and right-multiplication by m do not change distances between points, i.e. are isometries.*

Theorem 5. *From [21, Chapter V], biinvariant metrics have the following properties:*

1. *A biinvariant metric is also invariant w.r.t inversion*
2. *It is biinvariant if and only if $\forall m \in G, Ad(m)$ is an isometry of the Lie algebra $\mathfrak{g}$, where $Ad(m)$ is the adjoint representation of m.*
3. *One-parameter subgroups of G are geodesics for the biinvariant metric. Conversely, geodesics are simply given by left- or right-translations of one-parameter subgroups.*

Corollary 3. *Any metric $<,>$ on $T_{Id}Sym_\star^+(n) = Sym(n)$ extended to $Sym_\star^+(n)$ by left- or right-multiplication is a biinvariant metric.*

Proof. The commutativity of the multiplication implies that $Ad(Sym_\star^+(n)) = \{Id\}$, which is trivially an isometry group.

This result is striking. In general Lie groups, the existence of biinvariance metrics is not guaranteed. More precisely, it is guaranteed if and only if the adjoint representation $Ad(G)$ is relatively compact, i.e. (the dimension is assumed finite) if the group of matrices given by $Ad(G)$ is *bounded* ([21], Theorem V.5.3.). This is trivially the case when the group is commutative, like here, since $Ad(G) = \{e\}$, which is obviously bounded. Other remarkable cases where $Ad(G)$ is bounded are compact groups, like rotations. But for non-compact non-commutative groups, there is in general *no* biinvariant metric, as in the case of rigid transformations.

Definition 6. *Any biinvariant metric on the tensor Lie group is also called a* **Log-Euclidean** *metric, because it corresponds to a Euclidean metric in the logarithmic domain as is shown in Corollary 4.*

Corollary 4. *Let $<,>$ be a biinvariant metric on $Sym_\star^+(n)$. Then its geodesics are simply given by the translated versions of one-parameter subgroups, namely:*

$$(\exp(V_1 + t.V_2))_{t\in\mathbb{R}} \text{ where } V_1,\ V_2 \in Sym(n). \tag{13}$$

The exponential and logarithmic maps associated to the metric can be expressed in terms of matrix exponential and logarithms in the following way:

$$\begin{cases} \log_{S_1}(S_2) = D_{\log(S_1)} \exp .(\log(S_2) - \log(S_1)) \\ \exp_{S_1}(L) = \exp(\log(S_1) + D_{S_1} \log .L). \end{cases} \tag{14}$$

The scalar product between two tangent vectors V_1, V_2 at a point S is given by:

$$< V_1, V_2 >_S = < D_S \log . V_1, D_S \log . V_2 >_{Id} . \tag{15}$$

From this equation, we get the distance between two tensors:

$$d(S_1, S_2) = \| \log_{S_1}(S_2) \|_S = \| \log(S_2) - \log(S_1) \|_{Id}. \tag{16}$$

where $\|.\|$ is the norm associated to the metric.

Proof. Theorem 5 states that geodesics are obtained by translating one-parameter subgroups and Prop. 4 gives the form of these subgroups in terms of matrix exponential. By definition, the metric exponential $\exp_{S_1} : T_{S_1}Sym_\star^+(n) \to Sym_\star^+(n)$ is the mapping that associates to a tangent vector L the value at time 1 of the geodesic starting at time 0 from S_1 with an initial speed vector L. Differentiating the geodesic equation Eq. (13) at time 0 yields an initial vector speed equal to $D_{V_1} \exp . V_2$. As a consequence, $\exp_{S_1}(L) = \exp(\log(S_1) + (D_{\log(S_1)} \exp)^{-1}.L)$. The differentiation of the equality $\log \circ \exp = Id$ yields: $(D_{\log(S_1)} \exp)^{-1} = D_{S_1} \log$. Hence the formula for $\exp_{S_1}(L)$. Solving in L the equation $\exp_{S_1}(L) = S_2$ provides the formula for $\log_{S_1}(S_2)$.

The metric at a point S is obtained by propagating by translation the scalar product on the tangent space at the identity. Let $L_S : Sym_\star^+(n) \to Sym_\star^+(n)$ be the logarithmic multiplication by S. We have: $< V_1, V_2 >_S = < D_S L_{S^{-1}}.V_1, D_S L_{S^{-1}}.V_2 >$. But simple computations show that $D_S L_{S^{-1}} = D_S \log$. Hence Eq. (15). Finally, we combine Eq. (14) and Eq. (15) to obtain the (simple this time!) formula for the distance.

Corollary 5. *Endowed with a biinvariant metric, the tensor space is a flat Riemannian space: its sectional curvature (see [9], page 107) is null everywhere.*

This is clear, since it is isometric to the $Sym(n)$ endowed with the Euclidean distance associated to the metric.

In [16], the metric defined on the tensor space is affine-invariant. The action $\mathrm{act}(A)$ of an invertible matrix A on the tensor space is defined by:

$$\forall S, \ \mathrm{act}(A)(S) = A.S.A^T.$$

Affine-invariance means that $\forall A$ invertible, $\mathrm{act}(A) : Sym_\star^+(n) \to Sym_\star^+(n)$ is an isometry. This group action describes how a tensor, assimilated to a covariance matrix, is affected by a general affine change of coordinates.

Here, the Log-Euclidean Riemannian framework will not yield full affine-invariance. However, it is not very far from it, because we can obtain invariance by similarity (rotation plus a scaling).

Proposition 5. *We can endow $Sym_\star^+(n)$ with a similarity-invariant metric, for instance by choosing $< V_1, V_2 > := Trace(V_1.V_2)$ for $V_1, V_2 \in Sym(n)$.*

Proof. Let $R \in SO(n)$ be a rotation and $s > 0$ be a scaling factor. Let S be a tensor. V is transformed by the action of $s.R$ into $\text{act}(sR)(S) = s^2.R.S.R^T$. From Eq. (16), the distance between two tensors S_1 and S_2 transformed by sR is:

$$\begin{aligned} d(\text{act}(sR)(S_1), \text{act}(sR)(S_2)) &= \text{Trace}(\{\log(\text{act}(sR)(S_1)) - \log(\text{act}(sR)(S_2))\}^2) \\ &= \text{Trace}(\{\log(s^2)Id + R.\log(S_1).R^T - \log(s^2)Id - R.\log(S_2).R^T\}^2) \\ &= \text{Trace}(\{R.(\log(S_1) - \log(S_2)).R^T\}^2) \\ &= \text{Trace}(\{\log(S_1) - \log(S_2)\}^2) \\ &= d(S_1, S_2). \end{aligned}$$

Hence the result.

Invariance in the Tensor Lie group Thus, we see that the Tensor Lie group with a Log-Euclidean metric has many invariance properties: Lie group biinvariance and similarity-invariance. Moreover, Theorem 5 shows that the *symmetry principle* is respected: the inversion mappin: $S \mapsto S^{-1}$ is an isometry. In Section 5, we will see another invariance property, called *exponential invariance*: for instance the Log-Euclidean mean of square roots of tensors will be square root of the Log-Euclidean mean of the same tensors.

4.4 A Vector Space Structure on the Tensor Space

We have already seen that the Tensor Lie group is isomorphic and diffeomorphic to the additive group of symmetric matrices. We have also seen that with a Log-Euclidean metric, the Tensor Lie group is also isometric to the space of symmetric matrices endowed with the associated Euclidean metric. This is much already, but there is still more!

A Logarithmic Scalar Multiplication The Lie group isomorphism exp from the Lie Algebra of symmetric matrices to the tensor space can be extended into an isomorphism of vector spaces. Indeed, we set:

Definition 7. *The logarithmic scalar multiplication $\circledast$ of a tensor by a scalar $\lambda \in \mathbb{R}$ is:*

$$\lambda \circledast S = \exp(\lambda. \log(S)) = S^\lambda. \tag{17}$$

For example, multiplying a tensor by a scalar factor of $1/2$ is equivalent to taking its square root.

When we assimilate the logarithmic multiplication to an addition and the logarithmic scalar multiplication to a usual scalar multiplication, we have all the properties of a vector space. By construction, the mapping $\exp : (Sym(N), +, .) \to (Sym_\star^+(n), \odot, \circledast)$ is a vector space isomorphism. This of course does not mean that the tensor space is a subvector space of the vector space of square matrices. But this shows that we can view this space as a vector space when we identify a tensor to its logarithm. The question of whether or not the tensor space is a vector space depends on the vector space structure we are considering, and *not* on the space itself. With the classical Euclidean structure of square matrices, it is not a vector space. But with the logarithmic multiplication and the logarithmic scalar multiplication, it is.

Definition 8. $(Sym^+_\star(n), \odot, \circledast)$ *is called the Tensor Vector Space.*

From this viewpoint, biinvariant metrics on the Tensor Lie group are simply derived from the classical Euclidean metrics on the vector space $(Sym(n), +, .)$. **Thus we have in fact defined a new Euclidean structure on the tensor space** by transporting that of its Lie Algebra $Sym(n)$ on tensors. But this Euclidean structure has not the defects mentioned in the Section 1 of this article: matrices with null eigenvalues are at infinite distance and the symmetry principle is respected. Moreover, many invariance properties are valid, like similarity-invariance.

5 Probabilities and Statistics with Log-Euclidean Metrics

In this section, we present the Riemannian statistical framework for the Tensor Space with Log-Euclidean metric. It is particularly simple, since it is the same as the usual Euclidean framework when one identifies a tensor to its logarithm.

Practically, one simply uses the usual tools of Euclidean statistics on the logarithms and maps the results back to the Tensor Vector Space with the exponential. We recall that theoretically fully justified because the Tensor Lie group endowed with a biinvariant metric (i.e. here a Log-Euclidean metric) is isomorphic, diffeomorphic and isometric to the additive group of symmetric matrices with the associated[3] Euclidean norm.

Applications of the statistics on tensors include for example DT-MRI segmention: see [13].

5.1 General Riemannian Statistical Framework

The tensor space with a Log-Euclidean metric is a Riemannian space, exactly as the tensor space with an affine-invariant metric. In [15], the statistical framework in Riemannian spaces is fully presented from a geometrical point of view.

In these spaces, one usually generalizes the classical expectation of a random-variable with the notion of Fréchet expectation. It is defined as the set of points which minimize the metric dispersion of the random variable. Let $(G, d(.,.))$ be a Riemannian space with its distance and $S : \Omega \rightarrow G$ a G-valued random variable. Let dP be the probability measure associated to S defined on the space of all possible outcomes Ω (see [2] for a complete description of the classical probabilistic framework, particularly for technical requirements such as measurability, which will not be mentionned here).

[3]By *associated*, we mean that the metric on the space of symmetric is the same as that used on the space of symmetric matrices viewed as the Lie Algebra of the Tensor Lie group.

With these notations, the Fréchet mean $\mathbb{E}(S)$ of S is defined by:

$$\mathbb{E}(S) = \arg\min_T \int_\Omega d^2(T, S(\omega))dP(\omega). \tag{18}$$

A priori, existence and uniqueness are only guaranteed when the values taken by S are contained in a region of G that is small enough. See [15] page 13 for the statement of Karcher's theorem.

When the Fréchet expectation is uniquely defined, one can also compute moments of superior order like the covariance. This is done this time using vectors, namely the logarithms centered on the mean. More precisely, the covariance matrix (see [15] page 17) is defined by:

$$Cov_{\mathbb{E}(S)}(S) = \mathbb{E}(\log_{\mathbb{E}(S)}(S)) = \int_\Omega \log_{\mathbb{E}(S)}(S(\omega)).\log_{\mathbb{E}(S)}(S(\omega))^T dP(\omega). \tag{19}$$

Within this framework, many usual statistical tools can be used, like Mahalanobis distance, generalizations of the normal law, etc. See [15] for more details.

5.2 Random Tensors

Thanks to the isometric isomorphism between the tensor space with a Log-Euclidean metric and the Euclidean vector space of symmetric matrices, the theory of tensor-valued random variables is greatly simplified. Every notion of probabilities and statistics on vectors is readily generalized in the tensor case.

Indeed, one can define the classical vector spaces of random tensors:

Definition 9. *We can define for $\alpha \geq 1$ the Banach vector space $(L^\alpha(\Omega, Sym_\star^+(n)), \odot, \circledast, \|.\|_\alpha)$ of L^α-integrable tensor-valued random variables by identification with the vector space of symmetric matrices-valued random variables with the same integrability requirement. $S \in L^\alpha(\Omega, Sym_\star^+(n))$ if and only if:*

$$\int_\Omega \|\log(S(\omega))\|^\alpha dP(\omega) < \infty. \tag{20}$$

The associated norm $\|S\|_\alpha$ is simply:

$$\|S\|_\alpha = \left(\int_\Omega \|\log(S(\omega))\|^\alpha dP(\omega)\right)^{\frac{1}{\alpha}}. \tag{21}$$

One can also compute characteristic functions, etc.

5.3 Fréchet Means and Covariances with Log-Euclidean Metrics

Computing means and expectations is particularly simple with Log-Euclidean metrics, much more than in the general Riemannian case. Indeed, we have:

For any L^2 tensor-valued random variable S, its Log-Euclidean Fréchet mean $\mathbb{E}_{\mathrm{LE}}(S)$, also called *Log-Euclidean mean*, is defined and uniquely so. It is given as in the Euclidean case by:

Theorem 6. *Let $<,>$ be a Log-Euclidean metric on the tensor space. Let S be a L^2 tensor-valued random variable. Then its Fréchet mean is well defined and writes:*

$$\mathbb{E}_{LE}(S) = \exp\left(\int_\Omega \log(S(\omega))dP(\omega)\right). \tag{22}$$

In particular, The Log-Euclidean mean of N tensors is given by:

$$\mathbb{E}_{LE}(S_1, ..., S_N) = \exp\left(\frac{1}{N}\sum_i \log(S_i)\right). \tag{23}$$

In the L^1 case, one can generalize the Fréchet expectation by defining the expectation directly from Eq. (22).

Proof. When one expresses everything in the logarithm domain, one is faced with the classical computation of expectations and means in a Euclidean vector space. Hence the result by mapping back the results with exp in the tensor domain.

This theorem shows that it is not necessary to suppose that the random tensor hits almost surely a small enough region of the tensor space. The usual integrability condition on vectors apply.

As announced in Section 4, there is a last invariance property associated to Log-Euclidean metrics, the *exponential-invariance*:

Proposition 6. *The Log-Euclidean mean in the tensor space is exponential-invariant. By this, we mean that if a scaling is applied in the logarithmic domain to a random tensor, then the resulting mean is scaled identically with respect to the Log-Euclidean mean of the same tensor-valued random variable. For example, the mean of the square root of a random tensor is the square root of its mean.*

Proof. This simply results from the factorization of a scalar factor for in the expectation of random vector.

Like for the mean, there are considerable simplifications in the computations of other moments. Their form is exactly the usual one. For covariances, we get:

Proposition 7. *Let us endow the Tensor Vector Space with a Log-Euclidean metric. Let $S : \Omega \to Sym_\star^+(n)$ be a tensor-valued random variable. Then the associated covariance matrix $Cov(S)$ is:*

$$Cov(S) = \mathbb{E}(\log_{\mathbb{E}_{LE}(S)}(S)) = \int_\Omega (\log(S(\omega)) - \log(\mathbb{E}_{LE}(S))) \otimes (\log(S(\omega)) - \log(\mathbb{E}_{LE}(S)))dP(\omega). \tag{24}$$

Proof. This is done by replacing $\log_{\mathbb{E}_{\mathrm{LE}}(S)}(S(\omega))$ by its value and taking into account the scalar product at the Fréchet mean from Eqs. (14) and (15). The differential of the logarithm is canceled out by the differential of the exponential like in the computation of distances.

5.4 General Log-Euclidean Statistical Framework

All probabilistic and statistical notions of the Euclidean framework are directly transposed by the logarithm into the Tensor Vector Space with a Log-Euclidean metric. Hence Kolmogorov's Strong Law of Large Number applies for Log-Euclidean means in the Tensor Vector Space. So does the Central Limit theorem and all the others usual tools. One can for example use Principal Component Analysis to analyze data, use Gaussian distributions, etc. This yields a much simpler framework than in [8, 13] where part of the affine-invariant framework statistics for tensors is presented. One can proceed in this domain exactly like for vector-valued random variables. As examples, we formulate the Strong Law of Large Numbers for tensors and the generalize the normal law to tensors. Other generalizations are straightforward and left to the reader.

Kolomogorov's Strong Law of Large Numbers for Tensors. *Let (S_n) be sequence of independent identically distributed L^1 tensor-valued random variables. Then in the L^1 sense and almost surely for $\omega \in \Omega$:*

$$\frac{1}{N}\sum_{k=1}^{N} \log(S_i)(\omega) \underset{N\to\infty}{\longrightarrow} \mathbb{E}(\log(S)). \tag{25}$$

which by continuity of exp *yields with the Log-Euclidean means:*

$$\mathbb{E}_{LE}\left((S_i)_{i=1..N}(\omega)\right) \underset{N\to\infty}{\longrightarrow} \mathbb{E}_{LE}(S) \textit{ almost surely.} \tag{26}$$

Proof. As mentioned in the theorem, only the continuity of the exponential is necessary to extend the vector case to the tensor case. For a vector proof, see [2]. We especially recommend to the interested reader the martingale and ergodic theory proofs.

Definition 10. *A random tensor T is said to follow a normal Law of mean S_0 and covariance Σ if and only if $\log(S)$ follows a classical normal law of mean $\log(S_0)$ and covariance Σ.*

6 Comparison Between Log-Euclidean and Affine-Invariant Metrics

A complete framework for affine-invariance computations in the tensor space is presented in [16]. In Section 5, we have emphasized how much the general Riemannian framework is

Properties	Affine-Invariant Metrics	Log-Euclidean Metrics
$\exp_{S_1}(L) =$	$S_1^{1/2}.\exp(S_1^{-1/2}.L.S_1^{-1/2}).S_1^{1/2}$	$\exp(\log(S_1) + D_{S_1}\log.L)$
$\log_{S_1}(S_2) =$	$S_1^{1/2}.\log(S_1^{-1/2}.S_2.S_1^{-1/2}).S_1^{1/2}$	$D_{\log(S_1)}\exp.(\log(S_2) - \log(S_1))$
$< L_1, L_2 >_S =$	$< S^{-1/2}.L_1.S^{-1/2}, S^{-1/2}.L_2.S^{-1/2} >_{Id}$	$< D_S\log.L_1, D_S\log.L_2 >_{Id}$
$d(S_1, S_2) =$	$\\| \log(S_1^{-1/2}.S_2.S_1^{-1/2}) \\|$	$\\| \log(S_2) - \log(S_1) \\|$
Geodesic between S_1 and S_2	$S_1^{1/2}.\exp(tW).S_1^{1/2}$ with $W = \log\left(S_1^{-1/2}.L.S_1^{-1/2}\right)$	$\exp((1-t)\log(S_1) + t\log(S_2))$
Invariances	Affine-invariance	Lie Group Biinvariance, Similarity-Invariance, Exponential-Invariance

Table 1: Comparison between affine-invariant and Log-Euclidean metrics for the Tensor Space.

simplified in the Log-Euclidean case. In this section, we first recall the differences between affine-invariant metrics and Log-Euclidean metrics in terms of elementary operators, distance and geodesics. Then we turn to a finer study of the properties of Fréchet means in the Log-Euclidean and affine-invariant cases. The comparison between the two types of means is important, because many algorithms use averaging procedures on tensors. As a consequence, differences in averaging have a direct influence on results in applications.

We show that Log-Euclidean and affine-invariant means are very similar. In particular, they have the same determinant and they coincide in a number of cases. But there are many cases were they are not equal. When different, the Log-Euclidean mean tends to be more anisotropic then the affine-invariant mean. We prove this locally when tensors are sufficiently isotropic: in this case, the trace of the Log-Euclidean mean is always greater than the trace of its counterpart when they differ. This is theoretically only a local result, but this phenomenon has been numerically globally verified in applications. The proof in the global case will be the object of future work.

6.1 Elementary Operations and Invariance

In terms of elementary operations like distance computations, geodesics and means, the Log-Euclidean case provides much simpler formulae than in the affine-invariant case. We also recall the respective invariance properties in Table 1.

However, we see that the exponential and logarithmic mappings are complicated in the Log-Euclidean case by the use of the differentials of the matrix exponential and logarithm. This is the price to pay to obtain very simple distances and geodesics. In most applications, this is of no consequence, because distances, geodesics and means can be used directly without using explicitly the metric. See Section 7.

6.2 Log-Euclidean and Affine-Invariant Means

Definition 11. *Let S be a random tensor. When they are well-defined, we will write $\mathbb{E}_{Aff}(S)$ its affine-invariant mean and we recall that we write $\mathbb{E}_{LE}(S)$ the Log-Euclidean mean.*

Theorem 6 has to be compared with the results obtained generally on Fréchet means and in particular on affine-invariant means. In the latter case, there is in general *no* closed form for the mean of more than two tensors. But here, in the Log-Euclidean metrics framework, there is one, thanks to the commutativity of the additive product between tensors. One should notice that the mean *does not depend on the metric* as long as the metric is Log-Euclidean. We have the same phenomenon for affine-invariant means. This generalizes the classical result on Euclidean means which are independent from the Euclidean distance used.

From [15], we see that if there is no closed form for the affine-invariant mean of tensors, there is rather a *barycentric equation* that defines it implicitly:

$$\int_\Omega \log(\mathbb{E}_{Aff}(S)^{-1/2}.S(\omega).\mathbb{E}_{Aff}(S)^{-1/2})dP(\omega) = 0. \tag{27}$$

The solution to this equation exists and is unique and when the tensor-valued random variable S is bounded. Practically, this leads to iterative numerical methods to solve the equation, which are far more computationnaly expensive than the direct computation of the Log-Euclidean mean using Eq. (22).

6.3 Geometric Interpolation of Determinants

The definition of the Log-Euclidean mean given by Eq. (23) is extremely similar to that of the classical scalar geometrical mean. Indeed, we classically have:

Definition 12. *The geometrical mean of positive numbers $d_1, ..., d_N$,is given by*

$$\mathbb{E}(d_1, ..., d_N) = \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(d_i)\right).$$

The Log-Euclidean and affine-invariant Fréchet means can both be considered as generalizations of the geometric mean. Indeed, their determinant is the geometric mean of the geometric mean of the random determinants!

Theorem 7. *Let S be a L^1 random tensor. Then the determinant of its Log-Euclidean mean is the geometric mean of their random determinant. Moreover, if S is L^∞, its affine-invariant mean is defined and the same result holds for its determinant.*

Proof. From Prop. 1 we know that $\det(\exp(M)) = \exp(\mathrm{Trace}(M))$ for any square matrix M. Then for the geometric mean, we get:

$$\begin{aligned}
\det(\mathbb{E}_{LE}(S)) &= \exp(\mathrm{Trace}(\log(\mathbb{E}_{LE}(S)))) \\
&= \exp\left(\mathrm{Trace}(\int_\Omega \log(S(\omega))dP(\omega))\right) \\
&= \exp\left(\int_\Omega \mathrm{Trace}(\log(S(\omega)))dP(\omega)\right) \\
&= \exp\left(\int_\Omega \log(\det(S(\omega)))dP(\omega)\right) \\
&= \exp\left(\mathbb{E}(\log(\det(S)))\right).
\end{aligned}$$

The commutation between the integral and the trace done in the computations in justified because the trace is a linear operator on matrices. Hence the result.

For affine-invariant means, there is no closed form for the mean. *But* there is the barycentric equation given by Eq. (27). By applying the same formula as before after having taken the exponential and using $\det(S.T) = \det(S).\det(T)$ we obtain the result.

Theorem 7 shows that the Log-Euclidean and affine-invariant means of tensors are very similar. In terms of interpolation, this result is satisfactory, since it implies that the interpolated determinant, i.e. the volume of the associated interpolated ellipsoids, will vary between the values of the determinants of the source tensors. Indeed, we have:

Proposition 8. *Let S be a L^∞ random tensor. Then the determinant of the Fréchet Log-Euclidean and affine-invariant means are within the interval* $[\inf_{\omega\in\Omega}(S(\omega)), \sup_{\omega\in\Omega}(S(\omega))]$

Proof. This is simply a consequence of the monotonicity of the scalar exponential and of the scalar integral.

Corollary 6. *Let S_1 and S_2 be two tensors. The geodesic interpolations provided by the affine-invariant and Log-Euclidean metrics lead to a geometric interpolation of determinants. As a consequence, this interpolation of determinants is monotonic.*

Proof. Indeed, in both cases, the interpolated determinant $\mathrm{Det}(t)$ is the geometric mean of the two determinants, i.e. at $t \in [0, 1]$: $\mathrm{Det}(t) = \exp((1-t)\log(\det(S_1)) + t\log(\det(S_2)))$. This interpolation is monotonic, since the differentiation yields:

$$\frac{d}{dt}\mathrm{Det}(t) = \mathrm{Det}(t)\log(\det(S_2.S_1^{-1})).$$

The sign of $\frac{d}{dt}\mathrm{Det}(t)$ is constant and given by $\log(\det(S_2.S_1^{-1}))$.

6.4 Criterion for the Equality of the Two Means

In general, Log-Euclidean and affine-invariant means are similar, but they are *not* identical. Nonetheless, there are a number of cases where they are identical, for example when the random logarithms all commute with one another almost surely. In fact, we have more:

Proposition 9. *Let S be a L^∞ random tensor. If the Euclidean mean of the associated random logarithm commutes almost surely with* $\log(S)$*, then the Log-Euclidean and the affine-invariant means are identical.*

Proof. Let $\bar{L} := \int_\Omega \log(S(\omega))dP(\omega)$. The hypothesis is that $[\bar{L}, \log(S(\omega))] = 0$ almost surely. This implies that $\log(\exp(-1/2\bar{L}).S(\omega).\exp(-1/2\bar{L})) = \log(S(\omega)) - \bar{L}$ almost surely. We see then that $\exp \bar{L}$, i.e. the Log-Euclidean mean, is a solution of Eq. (27), i.e. is also the affine-invariant mean.

So far, we have not been able to prove the converse part of this proposition. It may be true in general and the proof of this assertion will be the subject of future work. However, the next subsection provides a partial proof, valid only when tensors are not too anisotropic, i.e. close to a scaled version of the identity.

6.5 Larger Anisotropy in Log-Euclidean Means

In Section 7, we see that affine-invariant means tend to be less anisotropic then Log-Euclidean means. Indeed, this is a general result. The following theorem explains this phenomenon when tensors are isotropic enough.

Theorem 8. *Let S be a random tensor close enough to the identity, so that we can apply the Baker-Campbell-Hausdorff formula almost surely (see Section 3). When the logarithm of the Log-Euclidean mean does not commute almost surely with* $\log(S)$*, then we have the following inequality:*

$$\mathit{Trace}(\mathbb{E}_{Aff}(S)) < \mathit{Trace}(\mathbb{E}_{LE}(S)). \tag{28}$$

Proof. The idea is to see how the two means differ close to the identity. To do this, we introduce a small scaling factor t and see how the two means vary when t is close to zero. Let S_t be the version of S scaled by t in the logarithmic domain. Around the identity, we can use the Baker-Campbell-Hausdorff (BCF) formula to simplify the barycentric equation (Eq. (27)). Let us use the following notations: $\log(S(\omega)) := L(\omega)$, $\bar{L}_{t;Aff} := \log(\mathbb{E}_{Aff}(S_t))$ and $\bar{L}_{LE} := \log(\mathbb{E}_{LE}(S))$.

First, we use twice the BCF formula to we obtain the following approximation:

$$\begin{aligned}\log(\mathbb{E}_{Aff}(S_t)^{-1/2}.S_t(\omega).\mathbb{E}_{Aff}(S_t)^{-1/2}) = \quad & tL(\omega) - \bar{L}_{t;Aff} - t^3\tfrac{1}{12}[L(\omega),[L(\omega),\bar{L}_{t;Aff}]] \\ & + t^3\tfrac{1}{24}[\bar{L}_{t;Aff},[\bar{L}_{t;Aff},L(\omega)]] + O(t^5).\end{aligned} \tag{29}$$

Then we integrate over Ω to obtain the following approximation Lemma:

Lemma 1. *When t is small enough, we have:*

$$\bar{L}_{t;Aff} = t\bar{L}_{LE} + \frac{t^3}{12}\int_\Omega [L(\omega), [\bar{L}_{LE}, L(\omega)]]dP(\omega) + O(t^5). \tag{30}$$

Proof. To obtain the approximation, note that the second factor $t^3\frac{1}{24}[\bar{L}_{t;Aff}, [\bar{L}_{t;Aff}, L(\omega)]]$ in Eq. (29) becomes a $O(t^5)$. Indeed, when the integration over Ω is done, $L(\omega)$ becomes $\bar{L}_{LE}$. But we can replace $\bar{L}_{LE}$ by its value in term of affine-invariance mean using Eq. (29). Then, using the fact that $[\bar{L}_{t;Aff}, \bar{L}_{t;Aff}] = 0$ we see that we obtain a $O(t^5)$.

Note also that thanks to the *symmetry principle*, $\bar{L}_{t;Aff}$ becomes $-\bar{L}_{t;Aff}$ when t is changer into $-t$, i.e. $t \mapsto \bar{L}_{t;Aff}$ is odd. As a consequence, only odd terms appear in the development in powers of t.

Next, we take the exponential of Eq. (30) and differentiate the exponential to obtain:

$$\mathbb{E}_{Aff}(S_t) = \mathbb{E}_{LE}(S_t) + D_{t\bar{L}_{LE}}\exp.\left(\frac{t^3}{12}\int_\Omega [L(\omega), [\bar{L}_{LE}, L(\omega)]]dP(\omega)\right) + O(t^5).$$

Then we use several properties to approximate the trace of affine-invariant means. First, we use Corollary 1 to simplify the use of the differential of the exponential. Then we approximate the exponential by the first two terms of its series expansion. We get:

$$\begin{aligned}
\mathrm{Trace}(\mathbb{E}_{Aff}(S_t)) &= \mathrm{Trace}(\mathbb{E}_{LE}(S_t)) + \mathrm{Trace}\left(\exp(t\bar{L}_{LE}).\frac{t^3}{12}\int_\Omega [L(\omega), [\bar{L}_{LE}, L(\omega)]]dP(\omega)\right) + O(t^5) \\
&= \mathrm{Trace}(\mathbb{E}_{LE}(S_t)) + \mathrm{Trace}\left((Id + t\bar{L}_{LE}).\frac{t^3}{12}\int_\Omega [L(\omega), [\bar{L}_{LE}, L(\omega)]]dP(\omega)\right) + O(t^5) \\
&= \mathrm{Trace}(\mathbb{E}_{LE}(S_t)) + \frac{t^4}{12}\int_\Omega \mathrm{Trace}\left(\bar{L}_{LE}.[L(\omega), [\bar{L}_{LE}, L(\omega)]]\right)dP(\omega) + O(t^5) \\
&= \mathrm{Trace}(\mathbb{E}_{LE}(S_t)) - \frac{t^4}{12}\int_\Omega \mathrm{Trace}\left(L(\omega)^2.\bar{L}_{LE}^2 - (L(\omega).\bar{L}_{LE})^2\right)dP(\omega) + O(t^5).
\end{aligned}$$

To conclude, we use the following Lemma:

Lemma 2. *Let A, $B \in Sym(n)$. Then:*

$$\mathit{Trace}(A^2.B^2 - (A.B)^2) \geq 0.$$

The inequality is strict if and only if A and B do not commute.

Proof. Let (A_i) (resp. (B_i)) be the column vectors of A (resp. B). Let $<,>$ be the usual scalar product. Then we have:

$$\begin{cases} \mathrm{Trace}(A^2.B^2) = \sum_{i,j} < A_i, A_j >< B_i, B_j > \\ \mathrm{Trace}((A.B)^2) = \sum_{i,j} < A_i, B_j >< B_i, A_j > . \end{cases}$$

Let us now chose a rotation matrix R that makes A diagonal: $R.A.R^T = Diag(\lambda_1, ..., \lambda_n) =:$ D. Let us define $C := R.B.R^T$ and use the notations (C_i) and (D_i) for the column vectors of C and D. We get:

$$\begin{cases} \text{Trace}(A^2.B^2) = \sum_{i,j} < D_i, D_j >< C_i, C_j >= \sum_i \lambda_i^2 < C_i, C_i > \\ \text{Trace}((A.B)^2) = \sum_{i,j} < D_i, C_j >< C_i, D_j >= \sum_{i,j} \lambda_i.\lambda_j < C_i, C_j > . \end{cases}$$

By the Cauchy-Schwarz inequality, we have $|\sum_{i,j} \lambda_i.\lambda_j < C_i, C_j > | \leq \sum_i \lambda_i^2 < C_i, C_i >$, which proves the first point. But the Cauchy-Schwarz inequality is an equality if and only if there is a constant μ such that $D.C = \mu C.D$. But only $\mu = 1$ allows the inequality of the lemma to be an equality. This is equivalent to $C.D = D.C$, which is equivalent in turn to $A.B = B.A$. Hence the result.

End of Proof of Theorem 8. When we apply Lemma 2 to the obtained estimation for the trace, we see that for a $t \neq 0$ small enough, the trace of the affine-invariant mean is indeed strictly inferior to the trace of the Log-Euclidean mean whenever the mean logarithm does not commute almost surely with the logarithm of the random tensor.

Corollary 7. *By invariance of the two means with respect to scaling, the strict inequality given in Theorem 8 is valid in a neighborhood of any tensor of the form λId with $\lambda > 0$.*

Corollary 8. *When the dimension is equal to 2, the Log-Euclidean mean of tensors which are isotropic enough is strictly more anisotropic then their affine-invariant mean when those means do not coincide.*

Proof. In this case, there are only two eigenvalues for each mean. Their products are equal and we have a strict inequality between their sum. Consequently, the largest eigenvalue of the Log-Euclidean mean is strictly larger than the affine-invariant one, and we have the opposite result for the smallest eigenvalue.

7 Application to Interpolation, Dense Extrapolation and Anisotropic Filtering

In this section, we compare experimental results between Log-Euclidean metrics and affine-invariant metrics. We have selected three applications already presented in [16]: multilinear interpolation, dense extrapolation of sparse tensor data and anisotropic filtering.

The results obtained are qualitatively and quantitatively very similar. The only difference is a little more anisotropy in results obtained with the Log-Euclidean metric. However, there is a considerable difference in terms of computational cost: the resampling of tensor fields can be as much as 1000 times faster in the Log-Euclidean case. Experimentally, a factor of at least 6 is observed in applications such as dense extrapolation or anisotropic filtering.

Moreover, PDEs are much simpler in the Log-Euclidean case. Thanks to the isometric isomorphism between the tensor space and the Euclidean vector space of symmetric matrices, the evolution equations are classical vector PDEs in the logarithmic domain. Their resolution consists simply of mapping all tensors to the logarithmic domain, solving the PDEs with classical vector methods, and finally mapping back the result to the tensor space with the exponential. This very significant simplification allows for the direct generalization of all vector PDE theory and methods to tensor fields. This yields a much more flexible and powerful framework than in the affine-invariant case where each classical PDE has to be carefully re-adapted to the Riemannian framework taking into account substantial corrections due to the curvature.

7.1 Multilinear Interpolation

Usual Diffusion Tensor Images are anisotropic. A typical DT image has $128*128*30$ voxels, and a voxel has the following spatial dimensions: $2mm*2mm*4mm$. When one wishes to resample them isotropically, i.e. in the example obtain $128*128*60$ voxels whose spatial extension is $2mm*2mm*2mm$, one is faced with the choice of the interpolation method. This resampling step can be crucial, for instance in algorithms tracking fibers in the white matter of the brain (see [18]).

In Fig. 1 are presented the results obtained for the interpolation of two tensors with three methods: linear interpolation of coefficients and interpolation along geodesics in the affine-invariant and Log-Euclidean cases. One sees immediately that the classical interpolation of tensors is inadequate: the determinant of the interpolated tensors does not vary in a monotonic way between the two tensors. The volume of the interpolated ellipsoid reaches a global maximum between the two extremities, which is not satisfactory physically in terms of diffusion. This swelling effect is typical of interpolation methods based on classical Euclidean metrics. On the contrary, geodesic interpolation provides a monotonic geometric interpolation of determinants, which results in interpolated ellipsoid with a satisfying volume (the proof of this result is given by Corollary 6). Note that the interpolated ellipsoids given by the Log-Euclidean metric are somewhat more anisotropic than their affine-invariant counterpart. This effect is general and was discussed in Section 6.

For 2D images (resp. 3D images) of vector data, bilinear (resp. trilinear) interpolation is used very frequently. It provides a fast a rather accurate interpolation. Fig. 2 presents the results of geodesic bilinear interpolation in the affine-invariant and Log-Euclidean cases. For a position in space between the four tensor germs, the associated interpolated tensor is the Fréchet mean associated to the classical barycentric coefficients of the spatial position. In the affine-invariant case, the computation of the mean is iterative, whereas in the Log-Euclidean case, the closed form given by Eq. (23) is used directly. Again, results are very similar, with more anisotropy in the Log-Euclidean case. But in terms of computational performance, the difference is extremely in favor of Log-Euclidean interpolation. The resampling based on Log-Euclidean metrics can be as much as 1000 times faster than in the affine-invariant case. Practically, a factor at least between 10 and 50 is observed.

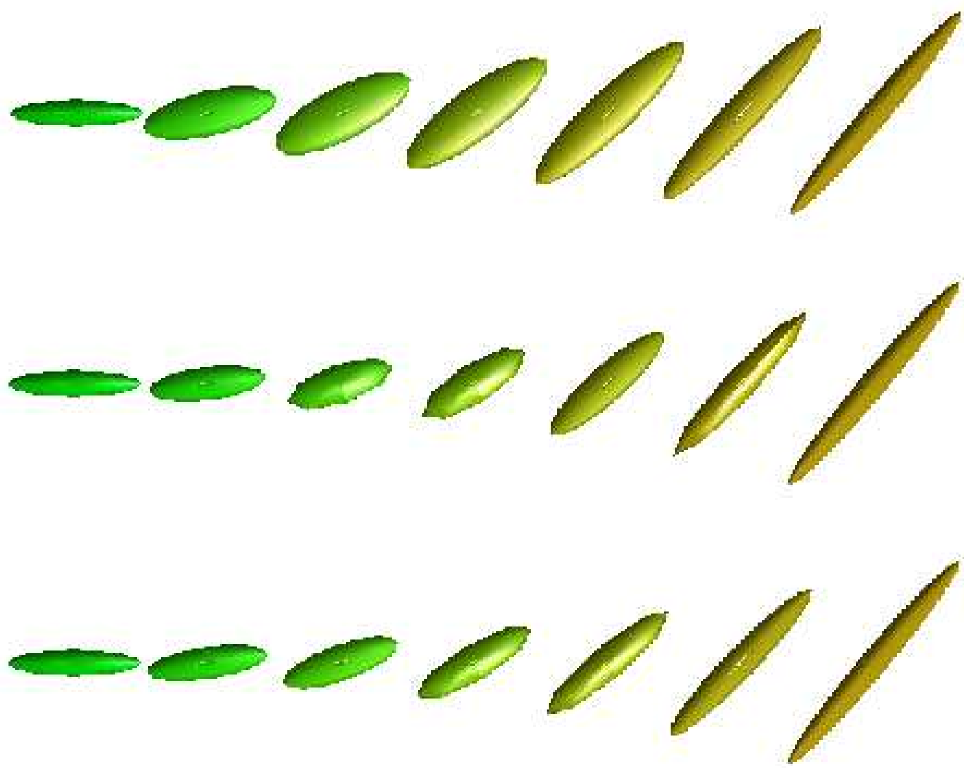

Figure 1: **Linear interpolation of two tensors**. **Above**: linear interpolation on coefficients. **Middle**: affine-invariant interpolation. **Below**: Log-Euclidean interpolation. Note the characteristic swelling effect observed in the Euclidean case, which is not present in both Riemannian frameworks. Note also that Log-Euclidean means are more anisotropic their affine-invariant counterparts.

7.2 Solving PDEs in a Riemannian Tensor Space

In [16] is shown how one can take into account the affine-invariant metric to solve PDEs in the Tensor Field. When one wishes to minimize a criterion of the type $E(S)$ where $S : \Omega \subset \mathbb{R}^3 \to Sym_\star^+(n)$ is a tensor Field, one can first compute the gradient of the criterion *relative to the metric* with the usual first variation of the criterion:

$$E(S + \epsilon dS)(x) = E(S)(x) + \epsilon < \nabla E(S)(x), dS(x) >_{S(x)} + o(\epsilon).$$

With this gradient, one can minimize the criterion with an *intrinsic* evolution equation of the type:

$$S_{t+1}(x) = \exp_{S_t(x)} \left(-\epsilon \nabla E(S_t)(x) \right).$$

In this *geodesic marching approach*, the idea is to adapt the classical first order gradient descent with a fixed time step ϵ to the Riemannian framework. This is done by shooting along the geodesic starting at $S(x)$ in the direction given by the opposite of the Riemannian gradient $\nabla E(S_t)(x)$ with a time step ϵ. This intrinsic evolution guarantees that at time t the field S_t is still a tensor field, which is not necessarily the case for a classical Euclidean gradient descent in which non-positive eigenvalues can appear.

7.3 Dense Extrapolation with Harmonic Diffusion

In [16], an optimization framework was proposed to densely extrapolate spare tensor information. This is useful in particular for the extrapolation of sparse first-order information on anatomical variability.

The criterion to minimize is the following:

$$E(S) = Sim(S) + \lambda Reg(S). \tag{31}$$

$Sim(S)$ is the *attachment term* to the sparse data, which is of the form:

$$Sim(S) = \int_\Omega \sum_{i=1}^{N} d^2(S(x), S_i) G_\sigma(x - x_i) dx. \tag{32}$$

Here, Ω is the image domain, the (x_i) are the N spatial positions where tensor measures (S_i) are known. This energy terms models the fact that values $S(x)$ near x_i should not differ too much from the measure S_i. The centered Gaussian G_σ of variance σ^2 quantifies how the influence of the measure S_i extends in space from x_i.

The proposed *regularization term* is the classical harmonic energy:

$$Reg(S) = \int_\Omega \|\nabla S(x)\|^2_{S(x)} dx. \tag{33}$$

More precisely, the squared norm of the spatial gradient $\|\nabla S(x)\|^2_{S(x)}$ is given by:

$$\|\nabla S(x)\|^2_{S(x)} = \sum_{j=1}^{3} \|\frac{\partial}{\partial x_j} S(x)\|^2_{S(x)}. \tag{34}$$

In [16], Eq. (31) is solved using the geodesic marching described in Section 7.2 with an adapted numerical scheme taking into account the Riemannian curvature.

Log-Euclidean Metric Case In the Log-Euclidean case, the resolution of Eq. (31) is greatly simplified. In effect, Eq. (32) becomes:

$$Sim(S) = \int_\Omega \sum_{i=1}^{N} \| \log(S(x)) - \log(S_i)\|^2 G_\sigma(x - x_i) dx. \tag{35}$$

Moreover, using the metric equation (Eq (15)) and the fact that:

$$\frac{\partial}{\partial x_j}\log(S(x)) = D_{S(x)}\log . \frac{\partial}{\partial x_j}S(x).$$

Eq. (33) writes now simply:

$$Reg(S) = \int_\Omega \|\nabla \log(S(x))\|^2 dx. \tag{36}$$

In Eqs. (35) and (36), it is remarkable that the norms do not depend on the position $S(x)$ any more. *The Log-Euclidean Riemannian framework leads to a classical vector optimization problem in the logarithmic domain.* As a consequence, one can use classical resolution methods in the logarithmic domain and finally map the result back in the tensor domain with the exponential.

Experimental results and comparisons are presented in Fig. 3 and Fig. 4. A dense extrapolation of 4 tensors in performed both in the affine-invariant and Log-Euclidean framework. As expected, the results are very similar, with slightly more anisotropy in the Log-Euclidean case. To measure anisotropy, we use the *fractional anisotropy* (FA), which is defined by:

$$FA(S) = \left(\frac{3}{2}\frac{\sum_{i=1}^3 \left(\lambda_i - \bar{\lambda}\right)^2}{\sum_{i=1}\lambda_i^2}\right)^{\frac{1}{2}}.$$

In terms of computational cost, the Log-Euclidean framework leads to computations 7 times faster than their affine-invariant counterparts. See Table 2 for computation times on a Pentium M 2 GHz, 1 Go RAM. The parameters chosen in the experiment are: $\sigma = 1/16 * W$ where W is the width of the grid, $\lambda = 0.05$, $\Delta t = 0.5$ (time step). The metric used is the similarity-invariant metric presented in Section 4.

7.4 Anisotropic Filtering

In [16], a generalization of the anisotropic filtering presented in [17, 10] to tensor fields is detailed. The idea is to regularize the tensor field (i.e. make the regularity energy given by Eq. (33) decrease) in an adaptive way so as to preserve strong variations as much as

Grid Size	Aff.-Inv. Metric	Biinv. Metric	# of Iterations	Speed Ratio (Aff/Bi)
16x16	1m27.544s	12.158s	600	7.2
32x32	5m18.420s	41.995s	600	7.6
64x64	20m4.875s	2m47.384s	600	7.2
32x32x32	31m7.086s	4m0.032s	200	7.8

Table 2: **Comparison of computational performance** between affine-invariant and Log-Euclidean metrics for dense extrapolation.

possible. This low-level processing can be crucial for the success of high-level algorithms based such as those tracking fibers tracts in the white matter of the brain.

This was done in [16] with simple modifications of the gradient of the regularization energy that result in an attenuation of the gradient wherever strong variations occur. This leads to a modified Laplace operator. In the classical Euclidean case, the modified Laplace operator writes:

$$\Delta_{aniso} S(x) = \sum_{j=1}^{3} c(\|\frac{\partial}{\partial x_j} S(x)\|) \frac{\partial^2}{\partial x_j^2} S(x) \tag{37}$$

Practically, the family of functions $c(x) = \exp(-x^2/\kappa^2)$ was used in [16].

Log-Euclidean Case Everything is simplified. Regularizing with the Log-Euclidean metric a tensor field is equivalent to a Euclidean regularization of the associated logarithmic field which is a vector field. Eq. (37) can be used directly since the Laplace operator is simply the usual Laplacian in the logarithmic domain.

As an example, we have filtered anisotropically a DTI slice of $128 * 128$ voxels in both Riemannian frameworks. The parameters used are: $\kappa = 1.0$, $\Delta t = 0.1$ (time step) and the number of iterations is 20. Anew, results are very similar, with slightly more anisotropy in the Log-Euclidean case. Computations are faster in the Log-Euclidean case by a factor at least 6. The results are shown in Fig. 5 and Fig. 6. As expected, there is *no* tensor swelling, as opposed to results in the literature obtained with classical Euclidean metrics as in [7, 5].

General Anisotropic Filtering in the Log-Euclidean Case Since regularizing with the Log-Euclidean metric a tensor field is equivalent to a Euclidean regularization of the associated logarithmic field, all classical filtering methods on vector fields apply directly to the tensor case. This will be the subject of future work.

From a mathematical point of view, the fact that such evolution PDEs on tensors are in fact PDEs on vector fields *guarantees the existence and uniqueness of solutions* for the tools used in the literature of PDEs on vector-valued images (see [19] for complete references on this theory). This is very satisfying when compared to the lack of current theory for general PDEs on tensor-valued images [5].

8 Conclusions

In this work, we have presented a new efficient Riemannian framework for tensor calculus. Based on Log-Euclidean metrics on the tensor space, or equivalently on Euclidean metrics on a new vector space structure on tensors, this framework transforms Riemannian computations on tensors into Euclidean computations on their logarithms. These logarithms are symmetric matrices, and thus can be considered as vectors. This leads to simple and efficient extensions of the classical tools of vector statistics and analysis to tensors.

Like affine-invariant metrics, Log-Euclidean metrics correct the defects of the classical Euclidean framework: non-positive eigenvalues are at an infinite distance, and matrix inversion does not change the distance between tensors. Like in the affine-invariant case, averaging procedures with Log-Euclidean metrics are a new generalization of *geometric* averaging. With this tool, one can use very simply and efficiently classical evolution PDEs usually applied to vector-valued images, without any swelling effect as in the classical Euclidean case. This is remarkable when compared to the usual complexity of restoration methods proposed in the literature. Moreover, these methods restore tensors in many cases only indirectly through the use of extracted features carrying less information than tensors. This is obviously not the case with Log-Euclidean metrics.

From a practical point of view, results are very similar to those of the affine-invariant framework and are obtained much faster, with an experimental computation time ratio of at least 6 and sometimes much more in favor of the Log-Euclidean framework for the applications presented in this article.

In future work, we will investigate in further details the generalization of anisotropic filtering to tensors in order to restore noisy DTI images. In particular, we will try to measure the impact of anisotropic filtering on the tracking of fibers of white matter in the human nervous system. We also intend to explore in more details how these new tools on tensors can lead to more accurate statistics and extrapolations of the anatomical variability of the human brain.

In this article, we have shown that within the family of Riemannian metrics, there are in fact several alternatives, each with its specificities. In the very new field of research on tensor calculus, it remains to be seen whether other frameworks remain to be discovered. As this work shows, new points of view on the tensor space can lead to significantly faster and simpler computations in a number of applications. Indeed it would be very desirable to continue to increase the number of tools available for tensor calculus to have an adapted tool for each situation.

Acknowledgements

The authors thank Denis Ducreux, MD, Kremlin-Bicêtre Hospital (France), for the DT-MRI data he kindly provided for this study.

The first author expresses his highest gratitude toward his young son Vincent-Thibault, who woke him up at 3 AM on a sunday, right on time to have the vision of the logarithmic multiplication and of the theory of Log-Euclidean metrics.

References

[1] Denis Le Bihan. Diffusion mnr imaging. *Magn Reson Q*, 7:1–30, 1991.

[2] Patrick Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, third edition edition, 1995.

[3] Nicolas Bourbaki. *Elements of Mathematics: Lie Groups and Lie Algebra, Chapters 1-3*. Springer-Verlag, 2nd printing edition, 1989.

[4] T. Broxand, M. Roussonand, R. Deriche, and J. Weickert. Unsupervised segmentation incorporating colour, texture, and motion. In *Lecture Notes in Computer Science*, volume 2756, pages 353–360. Springer, Berlin, 2003.

[5] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Regularizing flows for constrained matrix-valued images. *Journal of Mathematical Imaging and Vision*, 20(1-2):147–162, 2004.

[6] O. Coulon, D.C. Alexander, and S.R. Arridge. Principal diffusion direction field regularization for diffusion tensor magnetic resonance images. In *ISMRM'2001*, Glasgow, UK, 2001. International Society for Magnetic Resonance in Medicine.

[7] C. Feddern, J. Weickert, B. Burgeth, and M. Welk. Curvature-driven PDE methods for matrix-valued images. Technical Report 104, Department of Mathematics, Saarland University, Saarbrücken, Germany, April 2004.

[8] P.T. Fletcher and S.C. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Proc. of CVAMIA and MMBIA Workshops, Prague, Czech Republic, May 15, 2004*, LNCS 3117, pages 87–98. Springer, 2004.

[9] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian Geometry*. Springer-Verlag, 2nd edition edition, 1993.

[10] G. Gerig, R. Kikinis, O. Kübler, and F.A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11(2):221–232, June 1992.

[11] Roger Godement. *Introduction à la Théorie des Groupes de Lie*. Publications Mathématiques de l'Université Paris VII, 1982.

[12] Serge Lang. *Algebra*. Graduate Texts in Mathematics. Springer, 3rd rev. ed. 2002. corr. 4th printing edition, 2004.

[13] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on multivariate normal distributions: A geometric approach and its application to diffusion tensor MRI. Research Report 5242, INRIA, June 2004.

[14] M. Moakher. A differential geometry approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 2004. To appear.

[15] Xavier Pennec. Probabilities and Statistics on Riemannian Manifolds: A Geometric approach. Research Report 5093, INRIA, January 2004. submitted to Int. Journal of Mathematical Imaging and Vision.

[16] Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A Riemannian framework for tensor computing. Research Report 5255, INRIA, July 2004. To appear to the Int. Journal of Computer Vision.

[17] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 12(7):629–639, 1990.

[18] C. Poupon, C. A. Clark, V. Frouin, J. Regis, I. Bloch, D. Le Bihan, and J.-F. Mangin. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *Neuroimage*, 12(2):184–95, August 2000.

[19] G. Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, Cambridge, UK, 2001.

[20] Laurent Schwartz. *Analyse Tome 2 : Calcul Differentiel*. Hermann, octobre 1997.

[21] Shlomo Sternberg. *Lectures on Differential Geometry*. Prentice Hall Mathematics Series. Prentice Hall Inc., 1964.

[22] X. Tricoche, G. Scheuermann, and H. Hagen. Tensor topology tracking: A visualization method for time-dependent 2d symmetric tensor fields. *Computer Graphics Forum (Proceedings of EUROGRAPHICS 2001)*, 20(3):461–470, 2001.

[23] D. Tschumperlé and R. Deriche. Diffusion pde's on vector valued images. *IEEE Signal Processing Magazine - Special issue on Mathematical Methods in Imaging*, 19(5):15–25, 2002.

[24] D. Tschumperlé and R. Deriche. Orthonormal vector sets regularization with PDE's and applications. *International Journal on Computer Vision*, 50(3):237–252, 2002.

[25] Michael Wüstner. A connected lie group equals the square of the exponential image. *Journal of Lie Theory*, 13:307–309, 2003.

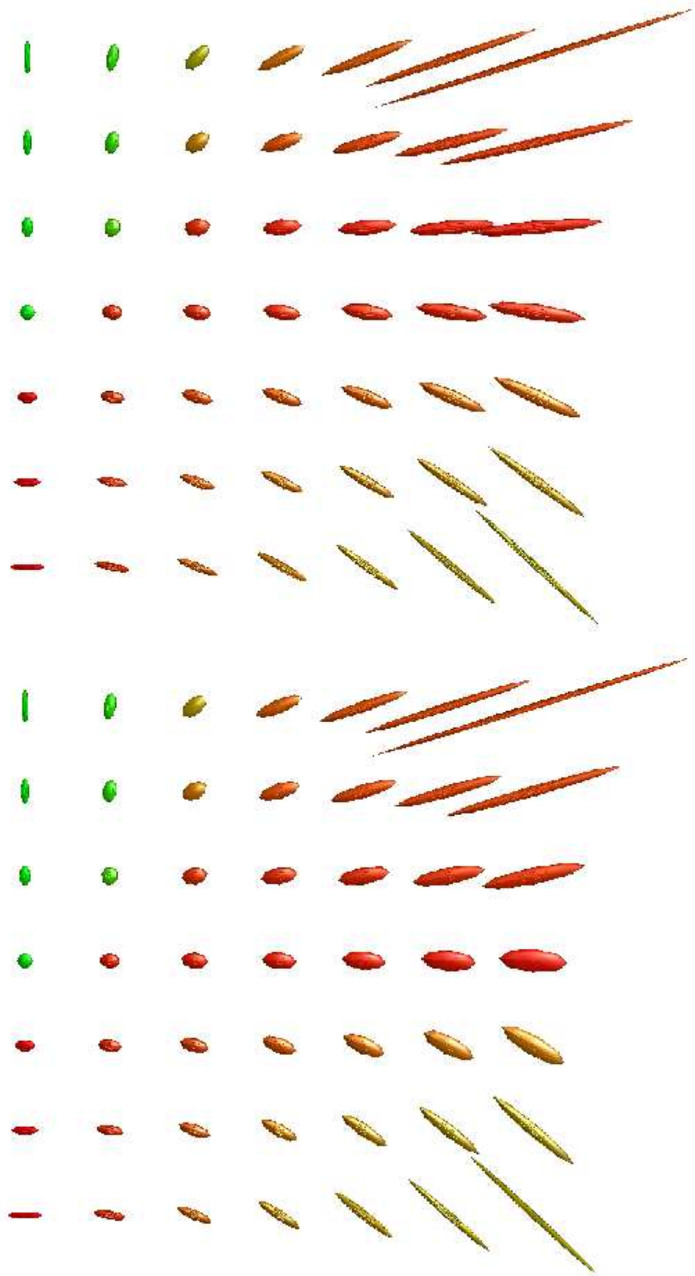

Figure 2: **Bilinear interpolation of four tensors**. **Above**: Log-Euclidean mean. **Below**: affine-invariant mean. Notice that the interpolation is the same to the left along the first column, because the two interpolated tensors commute. On the contrary, results are substantially different on the right: as expected, the anisotropy is larger in the Log-Euclidean case.

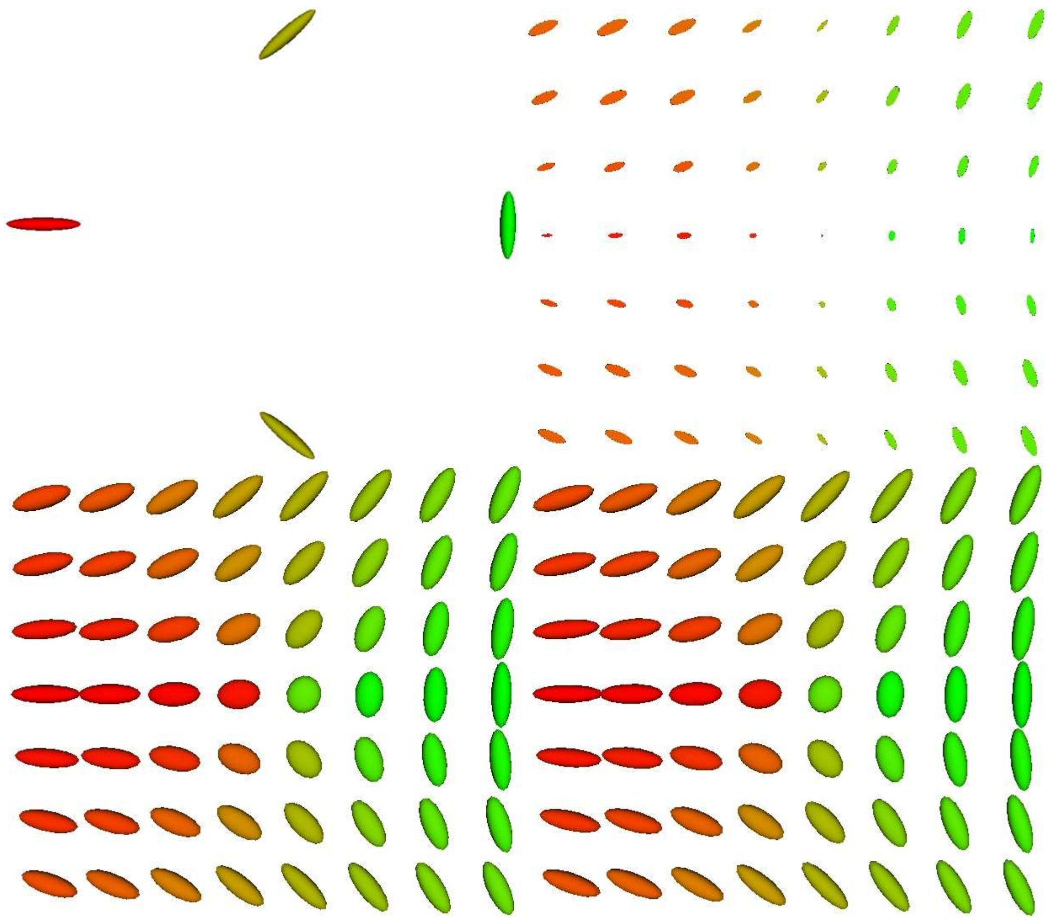

Figure 3: **Dense extrapolation of tensors with harmonic diffusion**. **Upper Left**: germs. **Upper Right**: magnified visualization of the absolute value of the Euclidean difference between the two extrapolations. **Lower Left**: affine-invariant extrapolation. **Lower Right**: Log-Euclidean extrapolation. Note that the Log-Euclidean extrapolation yields tensors a little more anisotropic. The differences are very small, and are mostly concentrated along the direction of anisotropy.

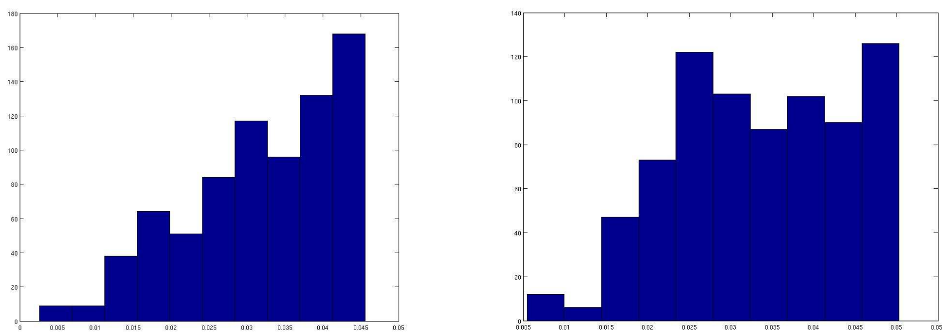

Figure 4: **Dense extrapolation of tensors with harmonic diffusion: comparison of anisotropy**. **Left**: histogram of the relative difference (i.e. the difference divided by the first term) between the fractional anisotropy of the Log-Euclidean extrapolation and its affine-invariant counterpart. **Right**: histogram of the relative difference between the largest eigenvalue of the Log-Euclidean extrapolation and its affine-invariant counterpart. Note that the differences are always positive: the anisotropy is larger in the Log-Euclidean case.

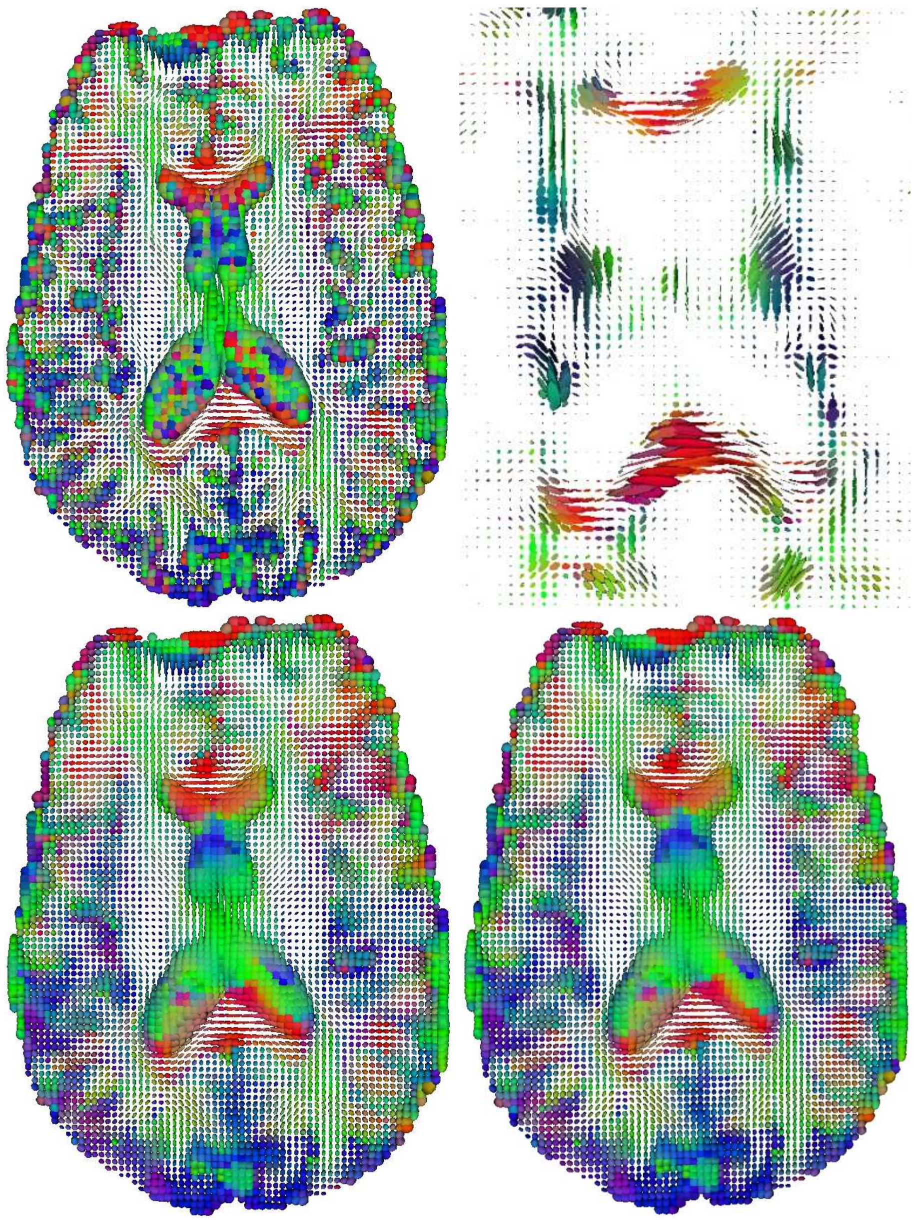

Figure 5: **Anisotropic filtering of a DTI slice. Upper Left**: raw DTI slice. **Upper Right**: zoom on a highly magnified visualization of the absolute value of the Euclidean differences between the two results (magnification factor of approximately 1000). **Lower Left**: affine-invariant result. **Lower Right**: Log-Euclidean result. Note that the differences are very small, and are concentrated along the directions of anisotropy. Note that there is *no* tensor swelling whatsoever, contrary to the Euclidean case.

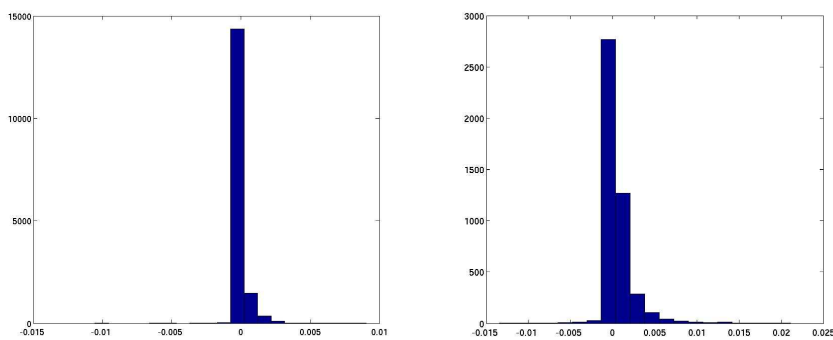

Figure 6: **Anisotropic filtering of a DTI slice .** **Above**: statistics on the relative difference(i.e. the difference divided by the first term) of fractional anisotropy between the affine-invariant and the Log-Euclidean cases. **Below**: statistics on the relative difference of largest eigenvalue between the affine-invariant and the Log-Euclidean cases. As expected, the anisotropy is generally larger in the Log-Euclidean case.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399