Riemannian processing of tensors for diffusion MRI and computational anatomy of the brain.

Pierre Fillard

UNIVERSITY OF NICE - SOPHIA ANTIPOLIS

GRADUATE SCHOOL STIC

INFORMATION AND COMMUNICATION TECHNOLOGIES AND SCIENCES

T H E S I S

to fulfill the requirements for the degree of

Ph.D. in Computer Science

from the University of Nice - Sophia Antipolis

Specialized in : CONTROL, SIGNAL AND IMAGE PROCESSING

presented and defended by

Pierre FILLARD

Riemannian Processing of Tensors for Diffusion MRI and Computational Anatomy of the Brain

Thesis supervised by Nicholas AYACHE and Xavier PENNEC,

defended the 8^{th} of February, 2008.

Jury:

Dr. Alfred ANWANDER,	Max Planck Institute	-	Invited member
Dr. Nicholas AYACHE,	Asclepios Project - INRIA	-	Director
Dr. Christian BARILLOT,	VisAGeS Project - IRISA	-	Examiner
Dr. Peter BASSER,	NIH Washington	-	Invited member
Dr. Rachid DERICHE,	Odyssée Project - INRIA	-	Examiner
Dr. Guido GERIG,	University of Utah	-	Reviewer
Dr. Thomas KNÖSCHE,	Max Planck Institute	-	Invited member
Dr. Jean-François MANGIN,	Neurospin - CEA	-	Reviewer
Dr. Xavier PENNEC,	Asclepios Project - INRIA	-	Co-director
Dr. Paul M. THOMPSON,	LONI - UCLA	-	Invited member
Dr. Carl-Fredrik WESTIN,	Harvard Medical School	-	Examiner

UNIVERSITÉ DE NICE - SOPHIA ANTIPOLIS

ECOLE DOCTORALE STIC
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

T H E S E

pour obtenir le titre de

Docteur en Sciences

de l'Université de Nice - Sophia Antipolis

Mention : AUTOMATIQUE, TRAITEMENT DU SIGNAL ET DES IMAGES

présentée et soutenue par

Pierre FILLARD

Traitement Riemannien des Tenseurs pour l'IRM de Diffusion et l'Anatomie Algorithmique du Cerveau

Thèse dirigée par Nicholas AYACHE et Xavier PENNEC,

soutenue le 8 février 2008.

Jury :

Dr. Alfred ANWANDER,	Max Planck Institute	-	Invité
Dr. Nicholas AYACHE,	Projet Asclepios - INRIA	-	Directeur
Dr. Christian BARILLOT,	Projet VisAGeS - IRISA	-	Examinateur
Dr. Peter BASSER,	NIH Washington	-	Invité
Dr. Rachid DERICHE,	Projet Odyssée - INRIA	-	Examinateur
Dr. Guido GERIG,	Univeristy of Utah	-	Relecteur
Dr. Thomas KNÖSCHE,	Max Planck Institute	-	Invité
Dr. Jean-François MANGIN,	Neurospin - CEA	-	Relecteyr
Dr. Xavier PENNEC,	Projet Asclepios - INRIA	-	Co-directeur
Dr. Paul M. THOMPSON,	LONI - UCLA	-	Invité
Dr. Carl-Fredrik WESTIN,	Harvard Medical School	-	Examinateur

Abstract

Symmetric, positive-definite matrices, or tensors, are nowadays a common geometrical tool for image processing and analysis. The recent emergence of diffusion tensor MRI (DT-MRI or DTI) and computational anatomy (CA) brought importance of tensors out to the medical community. In DTI, tensors are covariance matrices of the Brownian motion of water molecules within tissues, and allows to reconstruct white matter fibers through a process called tractography. In the context of CA (modeling of organs), tensors are covariance matrices depicting the anisotropy of shape variability. General processing of tensors is therefore highly demanded. However, working with those is difficult: the positive definiteness constraint must be satisfied at any cost, which cannot be ensured in general with standard matrix operations.

In this work, we propose two alternatives to the standard Euclidean calculus on tensors. Instead of seeing the tensor space as a vector space, we consider it as a manifold, i.e., a smooth curved space. Thanks to the Riemannian geometry, we are able to "unfold" this space, and to generalize any operation on tensors with astonishly simple implementations. In particular, we show how to recycle, at a very cheap cost, any image processing algorithms to tensors.

In a second step, we review the applications of such frameworks in the context of clinical DT-MRI and brain CA. In clinical DTI, the challenge is to exploit and perform tractography on data that may be very noisy, which is almost impossible with standard methods. To do so, we build a complete pipeline for this type of data, from the design of optimal processing by modeling the true nature of the MRI noise to their implementation in a convenient and ergonomic software. The utility of such approach is demonstrated with a clinical survey on the indications of DTI in spinal cord injuries.

In brain CA, we target the statistical modeling of the brain. We will show how, by considering simple brain anatomical landmarks - the sulcal lines - we are able to measure with tensors how much and in which prefered directions the brain varies among individuals. Then, we look at another type of statistics: the anatomical correlations between brain regions, telling how much two regions are statistically linked. A new type of tensors called total covariance matrices is involved in this study. We present results of so far unknown relationships between symmetric sucal positions, and between a-priori unrelated sulci, which raises new fundamental questions about the origin of such statistical dependencies.

Résumé

Les matrices symétriques définies positives (tenseurs) ont un rôle majeur en traitement et analyse des images. L'apparition de l'imagerie du tenseur de diffusion (ITD) et de l'anatomie algorithmique (AA) a fortement contribué à leur notoriété. Il est cependant difficile de les manipuler, car la propriété définie positive doit être vérifiée à tout prix, ce que le calcul Euclidien ne permet pas.

Dans cette thèse, nous proposons deux cadres Riemanniens pour remplacer le calcul Euclidien. Premièrement, nous proposons un cadre affine-invariant qui rend inaccessible toute matrice non positive, mais qui se révèle assez coûteux en temps de calcul. Nous proposons alors un second cadre, les métriques Log-Euclidiennes (LE), qui possède (presque) les mêmes propriétés que le précédent avec un coup algorithmique moindre. De manière surprenante, les deux cadres donnent des résultats similaires lors d'opérations d'interpolation ou de lissage.

Nous utilisons alors le cadre LE pour deux applications. Tout d'abord, nous proposons un panel de traitements des ITD cliniques. En particulier, nous repensons l'estimation des tenseurs avec le bon modèle de bruit IRM (Ricien). Ensuite, nous intégrons une partie des algorithmes développés dans le logiciel MedINRIA, dont le but est de rester simple tout en proposant des algorithmes puissants. Nous démontrons par ailleurs l'utilité d'une telle approche avec une étude clinique sur les indications de l'ITD dans les lésions de la moelle épinière. La seconde application vise l'AA du cerveau. Dans un premier temps, nous construisons un modèle de variabilité en mesurant et en extrapolant des matrices de covariance le long de lignes sulcales. Les cartes de variabilité obtenues sont cohérentes avec les résultats établis en neuroscience.

Dans un deuxième temps, nous nous intéressons aux corrélations anatomiques entre différentes zones du cerveau. Grâce à la matrice de covariance totale, nous calculons les corrélations entre un point de référence et le reste du cerveau. Nous montrons que la majorité du cerveau est corrélée symétriquement. D'autres régions comme l'aire de Broca ne présentent pas une telle symétrie (leurs fonctions dépendent de l'hémisphère également). Cela laisse penser que l'anatomie et la fonction sont liées, et que ces corrélations résultent de la présence d'un faisceau nerveux, ou ont une origine génétique ou environmentale.

Acknowledgments

First of all, I would like to thank Nicholas Ayache, my adviser during these 3 (almost 4) years of research in his group, for giving me the chance and the liberty to express my creativity in research without restrictions nor hesitations.

I would like to warmly thank Xavier Pennec, co-adviser of my thesis, who showed me the way when I was lost in the twilight zone of Riemannian geometry, and who transmitted me his appetite for mathematics. I could not have gone so far in my research without his precious and invaluable advices.

None of this could have been possible without Guido Gerig, who revealed to my eyes the fascinating world of research in the first place, and I cannot be thankful enough for this.

I would like to express my sincere gratitude to the members of the committee:

- Jean-François Mangin and Guido Gerig for kindly accepting to be reviewers of this thesis. I would like to thank them for the time and attention they devoted to my work, as well as for their encouraging compliments and judicious comments.
- Rachid Deriche, who kindly accepted to chair this jury, Carl-Fredrik Westin and Christian Barillot, who, by accepting to participate to the jury, demonstrated their interest in my work. I could not thank all of them enough for their compliments and advices.
- Peter Basser, Alfred Anwander and Thomas Knösche, who accepted three days before the defense to be part of the jury, despite the amount of work this required in such a short laps of time. I would like to thank them warmly for their exceptional kindness, their active participation, and their pertinent comments on my work.

I should not forget to express my immense gratitude to Paul M. Thompson, whose help was demanded more than once. I am thankful for the time he took to accurately analyze and interpret my results, and for the time taken to help me writing publications.

I would like also to thank two important people to me, Jean-Marie Becker and Catherine Mennessier, my professors in mathematics and computer science during engineer school, in particular for transmitting me their passion for image processing and analysis.

Naturally, none of this would have been possible without my parents, who gave me an education and a favorable environment, without which I could not have

been so far. I wanted to thank to my brother and sister, Clément and Hélène, who have been supporting me for now 28 years. I could not thank enough Ghislaine, my girlfriend and forever soulmate, for her unconditional support, patience, and love.

I would like to express my sympathy and gratitude to present and past labmates, with who I debated on various topics including research but not only: Stanley Durrleman, Olivier Clatz, Olivier Commowick, Nicolas Toussaint, Vincent Arsigny, Maxime Descoteaux, Maxime Sermesant, Ender Konukoglu, Jean-Marc Peyrat, Jean-Christophe Souplet, Céline Fouard, Radu Stephanescu, Isabelle Strobant, Heike Hufnagel, and Tommaso Mansi. I apologize to all people I forgot to mention the names.

Contents

Chapter 1

Introduction

Contents

The brain has been fascinating people for millennia. The role of this mysterious organ, however, was uncertain and very often misunderstood. In Egypt, in the late Middle Kingdom period (second millennium BC), in preparation for mummification, the brain was generally removed from the body through the nose, as it was the heart that was assumed to be the seat of intelligence. According to Herodotus (a Greek historian), during the first step of mummification: "The most perfect practice is to extract as much of the brain as possible with an iron hook, and what the hook cannot reach is mixed with drugs". Of course, over the next five thousand years, this view was reversed and the brain is now known to be the seat of intelligence. However, colloquial variations of the former remain as in "memorizing something by heart".

During the first millennium BC, the ancient Greeks developed different views on the function of the brain. It is said that it was the Pythagorean Alcmaeon of Croton who first considered the brain to be the place where the mind was located. Interestingly, during the 4th century BC, Aristotle thought that the brain was a cooling mechanism for the blood. He reasoned that humans are more rational than animals because they have a larger brain to cool their hot-bloodedness.

During the 1800s, scientists debated whether areas of the brain corresponded to specific functions, or if the brain functioned as a whole (the "aggregate field theory"). The work of Paul Broca, Karl Wernicke, and Korbinian Brodmann eventually helped to show that areas of the brain had specific functions. Today, modern neuroscience is rapidly growing due to the availability of computers capable of processing information of unprecedented complexity. Rapidly developing imaging

techniques such as magnetic resonance imaging (MRI), function MRI (fMRI) and diffusion tensor MRI (DT-MRI) allows scientists to study the brain in living humans and animals in ways their predecessors could not.

Up to now, only a little is known about the brain compared to other organs like the heart. The mistakes of our predecessors may be risible, but are we sure that our interpretations are correct? Most of our knowledge about brain anatomy are collections of observations made by neuroanatomists. The brain as we know it is identical as in the book "Gray's anatomy" first issued in...1858 [Gray 1958]! Today, thanks to the emergence of medical imaging for the past 20 years, a new hope arose to understand brain mechanisms. Brain can be observed in-vivo at very fine resolution with modern scanners. It can even be observed while "working" using fMRI or PET imaging. It is time to show our predecessors that brain's secrets can be revealed. Or maybe our successors will laugh of our mistakes. Future will tell.

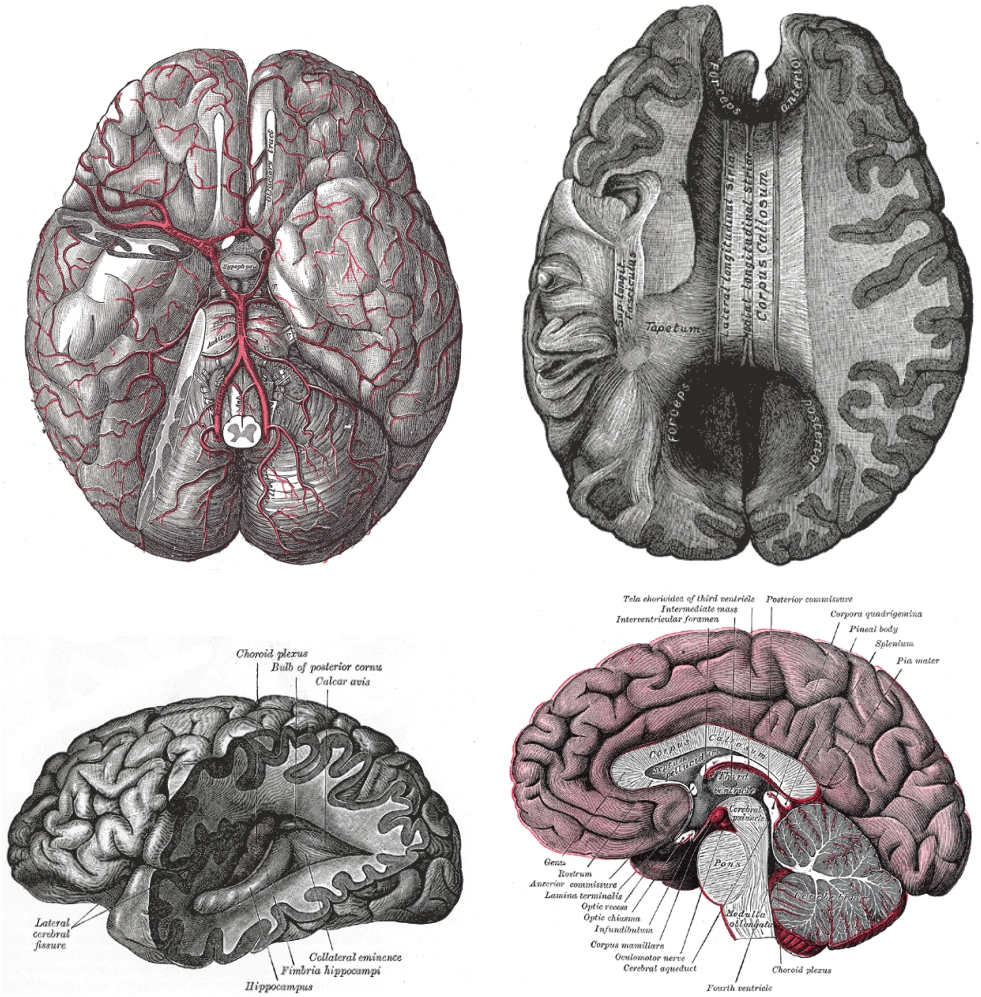


Figure 1.1: **Brain pictures taken from Gray's Anatomy (1858).**

1.1 Modeling of the Human Brain: A Neuroscience Challenge

Understanding brain's mechanisms, from a computer science point of view, means modeling those mechanisms. However, the complexity is such that a direct algorithmic modeling of the brain is almost impossible. Indeed, the brain is not a static arrangement of circuits, but a network of vastly interconnected neurons that are constantly changing their connectivity and sensitivity. More recent work in both neuroscience and artificial intelligence models the brain using the mathematical tools of chaos theory and dynamical systems. Current research has also focused on recreating the neural structure of the brain with the aim of producing human-like cognition and artificial intelligence but with little success.

Construction of computational models of the brain is the grand challenge of the neurosciences field. These models are the key for a better understanding of how the brain works, and essentially why some brains don't work as they should. Only little is known about many brain diseases like Parkinson, Alzheimer, schizophrenia, autism, multiple sclerosis, or even cancer. Having a better insight on the normal function of the brain would help detecting as early as possible these pathologies, and treating them. Our societies are getting older, and neurological pathologies are getting more and more frequent. The impact of such models, if they lead to actual treatment, will be huge.

Understanding brain also means understanding brain change over time, from birth to adult, and from adult to elderly. A better understanding of normal brain maturation can help early detect and treat severe pathologies, like autism or schizophrenia, that often appear during childhood. Another interesting point is neuroplasticity, which is the ability of the brain to change its organization (including localization of functions) as a result of experience. It is still mysterious and may contain the keys to function recovery. We could imagine giving drugs to a patient suffering from amnesia, or temporary hemiplegia (meaning that physical nerves are still present), to speed up the recovery of memory and motor functions.

There are different levels of modeling that need to be achieved, and we will discuss three of them. First, we need to model the functions, i.e., precisely delineate which areas of the brain are responsible for which functional tasks (like the speech area). Notice that this is a rather simplified description of the problem, since some higher functions may be distributed among different brain regions (and not only one). Moreover, there is an inter-individual variability of the localization of these regions, which makes the problem even more complex. Nevertheless, this information is accessible via fMRI, which is basically able to measure the neural response to a stimulus. Second, we need to model the anatomy, i.e., identify (delineate) and model the shape of brain structures. Modeling of anatomy (as well as functions)

should be done on a population of normal controls to characterize what "normality" is. Structural (or anatomical) MRI is the main tool for such level of modeling. Finally, there is a third level of modeling: the relationship between function and anatomy. We need to understand how function is linked to structure and shape. This modeling is making a bridge between the first two, and could help assessing how genes, experience-driven neuroplasticity, etc. affect brain organization. The emergence of diffusion MRI, a modality which eventually leads to a reconstruction of the white matter fibers, is a very promising tool to study such relationships.

1.2 Problems Investigated in this Thesis

In the previous section, we introduced the central theme of this work, which is the modeling of the human brain. Of course, this is a very complex problem that will not be fully solved here, and will require many more years of collaborative research between different domains (medicine, neuroscience, physics, mathematics and computer science) to be fully solved. In this thesis, we are investigating two aspects of it: diffusion MRI processing and computational anatomy of the brain, both having a major role to play in this long journey towards brain modeling.

First, diffusion MRI, as said earlier, is a promising tool to study the relationship between functions and anatomy as it gives access to the white matter fiber architecture. Those fibers physically connect functional regions. However, this imaging technique is relatively recent and not mature enough to be used for such precise modeling. Indeed, many problems inherent to this modality need to be tackled down first, especially in a clinical context. We will discuss these problems in Sec. 1.2.1, and a first contribution of this thesis is to address those problems.

Second, computational anatomy is a promising new field which specifically targets the study of an organ in a population, and whose goal is to create a model of variability of this organ. We introduce computational anatomy in Sec. 1.2.2, and a second contribution of this work is to build such model for the brain.

These two fields, that may sound unrelated at first glance, have something in common: the tensor. Tensors are a geometric tool to study statistical process, and are more complex than simple scalars to work with. We will show throughout the next two sections what are tensors, and which role they play in diffusion MRI and computational anatomy. In Sec. 1.2.3, we list the limitations of tensor computing. Finding a way to overcome these limitations will be the topic of another contribution of this thesis.

1.2.1 Diffusion Tensor MRI: In Vivo White Matter Fibers Characterization

The recent emergence of diffusion MRI, and in particular diffusion tensor MRI [Basser 1994a], has been received with enthusiasm by the medical image community. Indeed, diffusion MRI is the unique in-vivo technique to give an insight on the

anisotropy of tissues. Basically, diffusion MRI consists in measuring, in any direction of space, the diffusion of water molecules. This motion is better known as the *Brownian motion* formalized by Einstein in 1905 [Einstein 1956]. At a macroscopic scale, this phenomenon is perceived as a diffusion. Einstein related the diffusion coefficient to the distance of diffusion of particles:

$$d = \frac{1}{2\tau} \left\langle R^\top R \right\rangle,$$

where τ is the diffusion time, R is the displacement vector ($R = r - r_0$, r_0 being the initial position of a particle and r the final position), and $<>$ denotes the average over all particles. The scalar d is a measure of the diffusion of the medium and depends on the type of particles as well as the properties of the medium. However, it does not indicate a potentially different diffusion in a direction of space.

Certain media exhibit different diffusion properties. In particular, brain white matter fibers are protected by myelin sheaths that will "force" the water molecules to move in a direction tangential to them rather than orthogonally. In other words, water molecules move preferentially along the direction of the fibers than transvere to them. In this case, a simple scalar is not suitable to characterize the diffusion of such anisotropic tissue. Then, one can replace it by the covariance matrix D of the particles displacement:

$$D = \frac{1}{6\tau} \left\langle R R^\top \right\rangle.$$

D is a 3×3 symmetric, positive-definite matrix. The term *diffusion tensor* comes from [Stejskal 1965] and is related to the fact that it is obtained by the tensor product of the displacement vector R with itself. In rigorous mathematical notations, a scalar is a 0^{th} order tensor, a vector is a 1^{st} order tensor, and a matrix is a 2^{nd} order tensor (without any restriction on its symmetry and positivity). Thus, the denomination *tensor* is ambiguous since any matrix is a 2^{nd} order tensor! However, the term was adopted by the community and has been used for many years. Then, we will use the word tensor as synonym of symmetric and positive definite matrix from now on.

A symmetric, positive definite matrix is a symmetric matrix whose eigenvalues are all strictly greater than zero. Let us take an eigen decomposition of a tensor D: $D = U\,V\,U^\top$. U is the matrix of eigenvectors (it is an orthogonal matrix, i.e., $U^{-1} = U^\top$) and V is the matrix of eigenvalues (it is a diagonal matrix):

$$U = \begin{pmatrix} | & | & | \\ \overrightarrow{u}_1 & \overrightarrow{u}_2 & \overrightarrow{u}_3 \\ | & | & | \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

We say that eigenvalue λ_i is associated to eigenvector $\overrightarrow{u}_i$ for all i. Moreover, when working with diffusion tensors, we generally order the eigenvalues in descending order, i.e., λ_i is greater than λ_{i+1} $\forall i$. Consequently, λ_1 is the greatest eigenvalue,

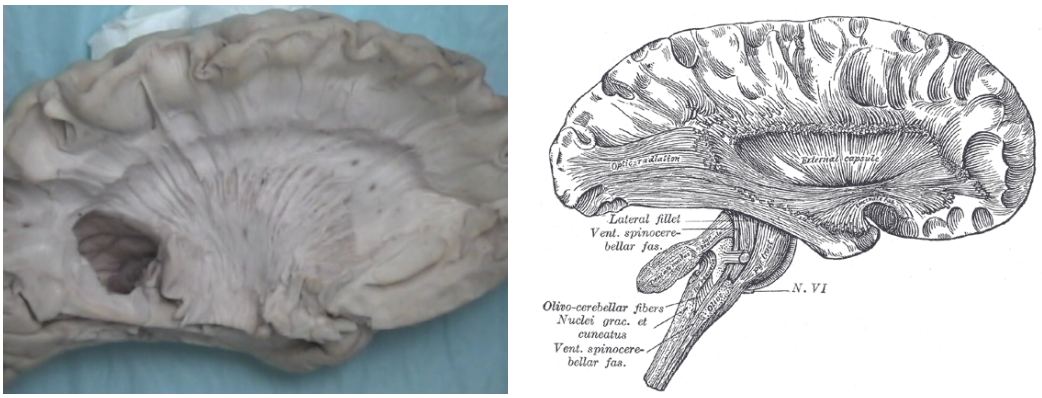


Figure 1.2: **Illustrations of brain white matter fibers. Left:** Dissection of a human brain showing some major fiber tracts (Courtesy of Lennart Heimer, MD, University of Virginia). **Right:** Image taken from Gray's Anatomy [Gray 1958].

and we call the associated eigenvector $\overrightarrow{u}_1$ the major eigenvector (also called the principal direction of diffusion (PDD) in case of diffusion tensors). Of course, all of this is valid for any $n \times n$ tensors, not only 3×3.
Representation of tensors is commonly done using ellipsoids. Indeed, tensors are nothing other than quadric forms. A quadric form whose eigenvalues are all positive is an ellipsoid. Actually, the rigorous definition states that an ellipsoid is a quadric form whose eigenvalues are all of the same sign, either positive or negative. But as we are modeling the diffusion phenomenon, the eigenvalues are homogeneous to a squared distance per second, so negative eigenvalues are physically not possible. The shape and orientation of the ellipsoid is determined by the eigen components of a tensor: the orientation of its axes are given by the eigenvectors, and the magnitude of each axis is given by the corresponding eigenvalue.

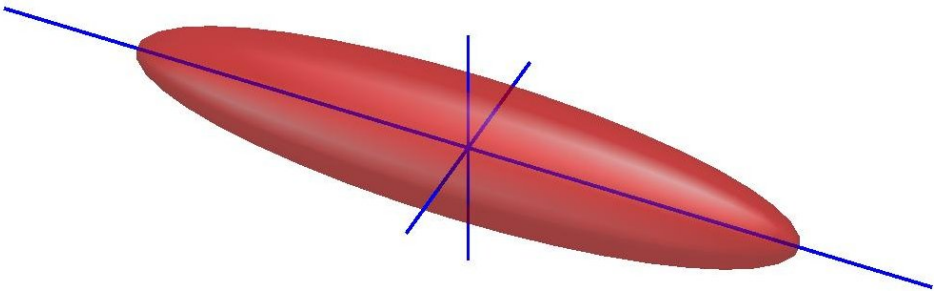


Figure 1.3: **A tensor represented by an ellipsoid.** The axes directions are given by the tensor's eigenvectors. The axes length is given by the eigenvalues.

We shall emphasize the fact that the diffusion tensor is a *model* of the covariance matrix of the particles displacement. This is not the unique way of modeling the diffusion. First, this model assumes a Gaussian distribution of the particle displace-

ments. The validity of such assumption could be questioned for instance in regions with crossing fibers: in this case, two or more Gaussian distributions are mixed up and nothing ensures the result to be Gaussian. A more elaborate model would be necessary in this situation and the literature offers a lot of alternatives to the simple tensor model: bi-tensor model [Alexander 2001a, Tuch 2002], CHARMED [Assaf 2005], Q-Ball imaging [Tuch 2004, Descoteaux 2007b], orientation density function [Tournier 2004a], or even higher order tensors [Barmpoutis 2007b]. However, these techniques requires a large number of MRI acquisitions (from several tens to several hundreds). Moreover, when one wants to have a robust estimate of the model parameters, one often needs to repeat the scans or acquire more images than the minimum number necessary. This is critical when working, for instance, in a clinical environment where the scanning time is rather short (average time of 15 minutes per patient). We will address this later in this manuscript. With a single tensor model, only six images are needed (in fact seven since a reference image is required) because a tensor has 6 degrees of freedom, which gives 6 parameters to estimate. Increasing the robustness of the model estimation (i.e., multiplying the acquisitions) becomes more reasonable in this case. Finally, we believe that the single tensor model still has a lot to offer.

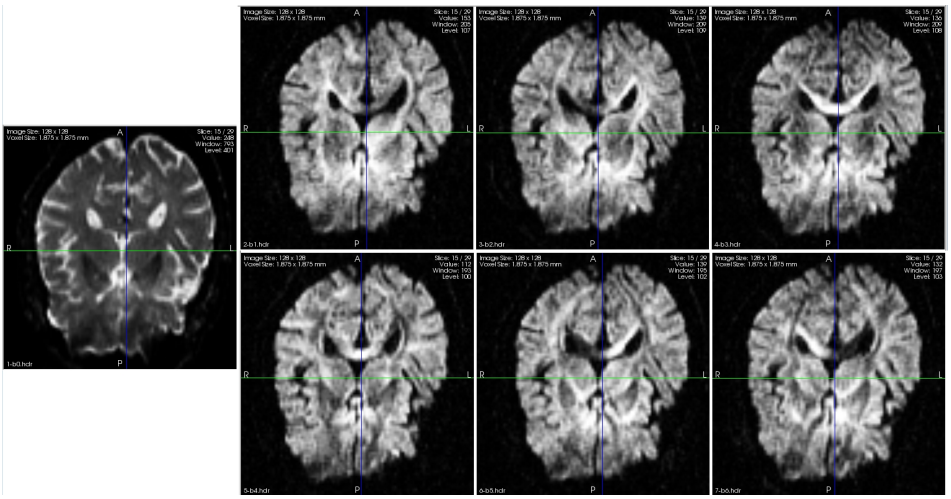


Figure 1.4: **Slices of a DT-MRI acquisition typical of a clinical environment.** (Courtesy of Denis Ducreux, MD, Bicêtre Hospital, Paris.) Image dimensions: $128 \times 128 \times 30$, spatial resolution: $1.8 \times 1.8 \times 4mm$. The dataset is made of one baseline image (the most-left image) plus 6 diffusion-weighted images (using Basser's sequence [Basser 1994a] was used).

Practically, the diffusion tensor D is related to the *diffusion-weighted* images by the Stejskal&Tanner diffusion equation:

$$S_i = S_0 \exp\left(-b\overrightarrow{g}_i^{\top} D \overrightarrow{g}_i\right), \tag{1.1}$$

where S_i is the diffusion-weighted image, S_0 is the *baseline* image (i.e., an image without displacement encoding gradients, typically a T_2-weighted image), $\overrightarrow{g}_i$ the diffusion gradient, and b the b-value in $s.mm^{-2}$ which depends on the scanning parameters (it is generally constant). For a more complete description of the diffusion MRI acquisition, we refer the reader to [Poupon 1999]. Basically, the coefficient $\overrightarrow{g}_i^{\top} D \overrightarrow{g}_i$ is a scalar giving a measure of the diffusion in the spatial direction $\overrightarrow{g}_i$. It is identical to the isotropic diffusion scalar d that we discussed previously except that it is measured in a spatial direction. $\overrightarrow{g}_i$ is called a *diffusion gradient* because it corresponds to an actual magnetic field gradient used in the MRI sequence. S_i is the signal returned by the scanner. The goal of diffusion tensor MRI is to reconstruct the field of diffusion tensors D from the measures S_i. Once again, since D has six degrees of freedom, at least six DWIs are required. Estimation is generally done by linearizing the diffusion equation 1.1 for all diffusion gradients, and solving this system of equations in a least-squares sense with algebraic methods. Once this estimation is done, one is able to represent the white matter architecture (in case of brain DTI) via a process called *tractography*.

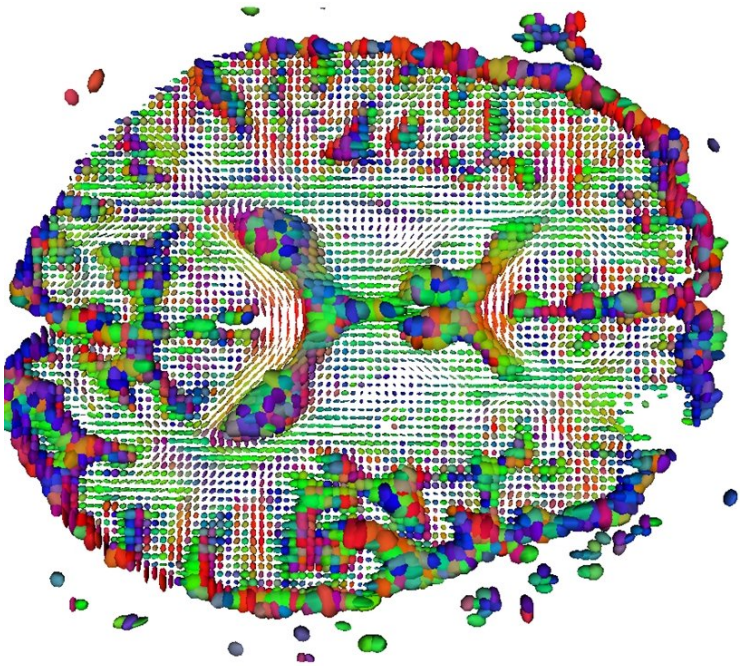


Figure 1.5: **Slice of the tensor field estimated from DWI of Fig. 1.4 with classical techniques.** We notice that the field looks rather noisy, and even some parts of it are missing (bottom right corner of this slice) due to distort of the acquisition.

Fiber reconstruction (or fiber tracking, or tractography), consists in building a geometric representation of white matter fibers. The hypothesis is made that diffusion tensors are aligned with the direction of oriented tissues like neural fibers

in case of brain DTI (although this hypothesis is obviously wrong in case of crossing fibers). Literature is abundant about fiber tracking and algorithms can be divided into two categories: streamline tracking and probabilistic tracking. Schematically, in streamline approaches, the field of principal direction of diffusion (i.e., the vector field composed with the principal tensors eigenvectors) is integrated from seed points and paths are then considered as fibers. The probabilistic approaches simulates many times the diffusion phenomenon from a seed point and consider the most likely path as a fiber.

In this thesis, we choose to place ourself in the context of clinical DT-MRI. The main reason is that in France, only a few MRI scanners are dedicated to research and most of the DTI acquisitions we get come from a clinical environment. Moreover, medical experts are really looking forward this relatively new technique as this is the unique in vivo modality allowing to study the integrity of white matter fibers. There are some constraints inherent to clinical acquisitions that must be taken into account though. First, the scanning time is rather short: 15 minutes maximum per patient in general. During this period of time, imaging modalities related to the pathology of the patient have to be acquired (T_1, T_2, proton density, etc.). And if there is some extra time, DWI sequence can be launched. This leaves most of the time 5 to 6 minutes for the diffusion sequences. This has a direct consequence on the acquisition: only very few gradients can be obtained with moderate to low signal-to-noise ratios. For instance, on a Siemens Sonata 1.5T scanner, acquisitions are limited to 25 gradient directions with 1 repeated scan. Diffusion tensor is then the best model for this type of acquisition, and switching to more elaborate models would be hazardous. Finally, like any MR acquisition, diffusion tensor MRI is subject to noise. This noise can seriously affect the quality of the diffusion tensors (Fig. 1.5), and consequently the result of tractography (Fig. 1.6): the estimation step can produce matrices that are not tensors (basically, nothing ensures the result to be positive definite, and noise, image distortions, or other acquisition-related artefacts - movement, etc. - affect the quality of the tensors), and noise puts uncertainty on tensors directions and magnitude.

In conclusion, clinical DTI faces a certain number of problems: acquisitions have only few encoding gradients (about 25) with relatively low SNRs. But the clinicians still need to be able to perform fiber reconstruction on their patient's images. How can we optimally exploit DTI typical of clinical acquisitions and give to clinicians the possibility to do fiber reconstruction?

1.2.2 Computational Anatomy: A Promising Emerging Field

The goal of computational anatomy is to develop algorithms to model and analyze the biological shape of tissues and organs. The goal is not only to estimate representative organ anatomies across species, populations, diseases, aging, ages, etc. but also to model the organ development across time (growth) and to establish

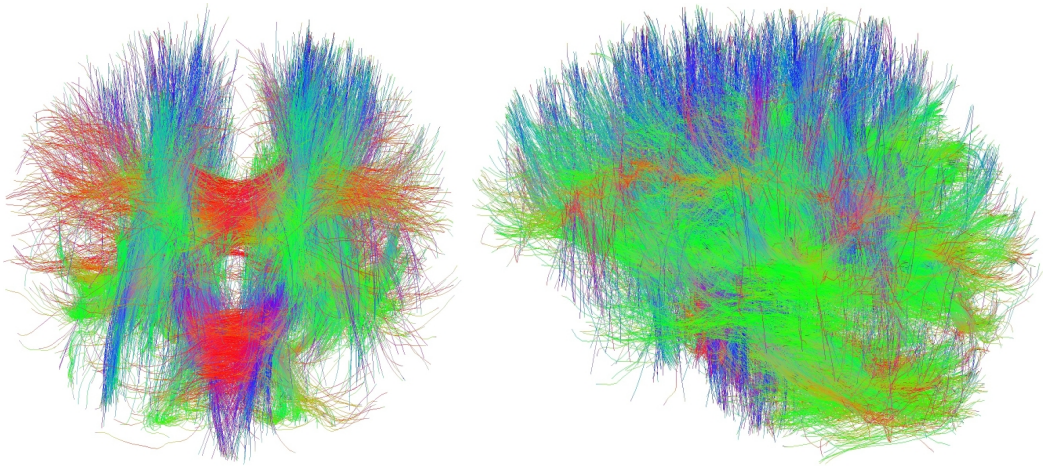


Figure 1.6: **Example of fiber tractography of the clinical dataset of Fig. 1.4.** Fibers do not look realistic, due to noise and the rather low resolution of the images, specially in the Z direction (dimensions: $128 \times 128 \times 30$, spatial resolution: $1.8 \times 1.8 \times 4mm$). Denoising and spatial interpolation become necessary in this case.

their variability. Another goal is to correlate this variability information with other functional, genetic or structural information (e.g., fiber bundles as obtained in diffusion tensor MRI). Computational anatomy is at the interface of geometry, statistics, and image analysis.

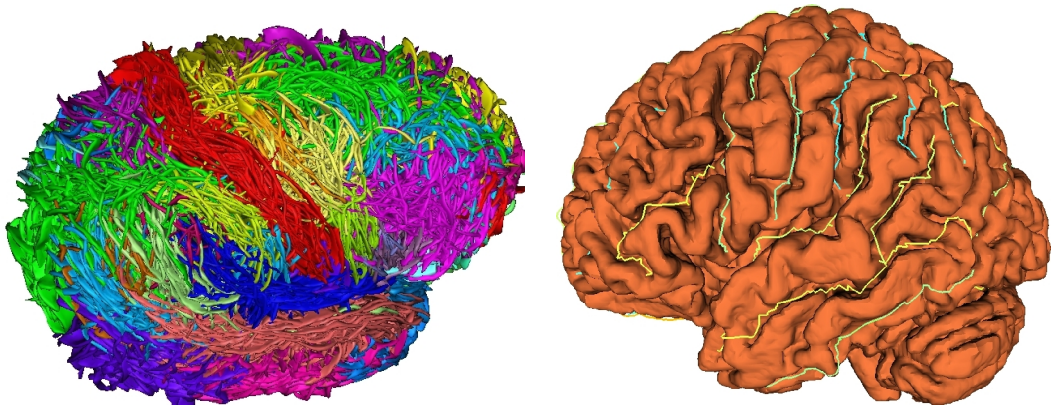


Figure 1.7: **Examples of brain anatomical landmarks that computational anatomy needs. Left:** Sulcal ribbons of 50 subjects were automatically extracted with BrainVisa (Courtesy of J.-F. Mangin, Neurospin, CEA) and overlapped in the same coordinate system. **Right:** Manual tracing of sulcal lines in a subject (the sulcal line is defined as the "bottom" line of the sulcal ribbon), courtesy of Dr. Paul M. Thompson.

Finding a point-by-point correspondence between organs of different subjects is a difficult task because of inter-subject variability. When one looks very roughly at two brains, we conclude that they look alike, e.g., they all have folds, a cerebellum and ventricles. But when we look closer, we begin to realize that many details differentiate them. For instance, the folding patterns are quite different from one individual to another: some major sulci appear in both (like the central sulcus), but others only appear in some individuals. Even common sulci and gyri vary drastically in shape among subjects. This raises the need for adapted physical models relating the anatomy of different subjects, in order to learn the geometrical relationship by statistical analysis. The general method is to identify anatomically representative geometric features (points, lines, surfaces, etc. - see Fig. 1.7) and to model their statistical distribution among a population (Fig. 1.8). This can be done, for instance, by computing the mean shape and the covariance structure after a group-wise alignment. In the case of the brain, one can rely on a wide range of anatomical and functional landmarks including AC-PC [Talairach 1988], curves (crest lines, sulcal lines) [Mangin 2004b, Le Goualher 1999, Le Goualher 1997], surfaces like sulcal ribbons [Cachia 2003, Rivière 2002], fiber bundles [Corouge 2006], or images when seen as 3D functions [Arsigny 2006a, Glaunès 2005, Vaillant 2005]. There is some evidence that variations of these entities among individuals is anisotropic. It means that the variability is depending on the spatial direction, and a simple scalar (typically the variance) cannot characterize such anisotropic variability. A parallel can be made with diffusion tensor MRI where the diffusion coefficient was replaced by a diffusion tensor: here we replace the variance by the covariance matrix. In this case, the covariance matrix characterizes the anisotropy of the variability taken at each individual position of the mean shape representation. We call this matrix a *variability tensor* by analogy to the diffusion tensor.

Common problems as with diffusion tensors arise: how can we manipulate these tensors? For instance, we may need to perform geometric operations, like interpolation or extrapolation, operations that are well-known for scalar images. Once again, is it possible to compute with tensors as we do with simple scalars? This is the question we investigate in the next section.

1.2.3 Limitations of Euclidean Calculus on Tensors

We call Euclidean calculus on tensors the fact of acting on the tensor coefficients directly. One can write a 3×3 tensor as:

$$D = \begin{pmatrix} d_{xx} & d_{xy} & d_{xz} \\ d_{xy} & d_{yy} & d_{yz} \\ d_{xz} & d_{yz} & d_{zz} \end{pmatrix}$$

Euclidean calculus consists in processing directly the tensor coefficients $(d_{xx}, d_{xy}, d_{xz}, d_{yy}, d_{yz}, d_{zz})$.

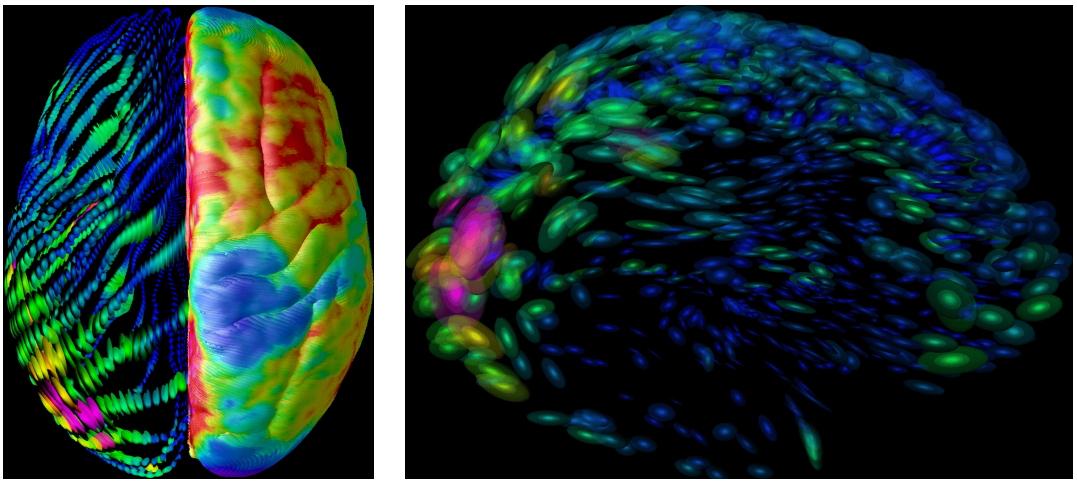


Figure 1.8: **Mapping brain variability in human populations.** This map displays the magnitude and principal directions of anatomical variability in the brain, based on a group of 40 normal subjects scanned with MRI. Pink colors indicate brain regions with large anatomical variability across subjects, blue colors low variability. The ellipsoidal glyphs are elongated along directions in which variation is greatest. These maps help distinguish pathological changes from normal variations. Images courtesy of Paul M. Thompson, Andrew Lee, Kiralee Hayashi, Agatha Lee, and Arthur Toga (LONI - UCLA School of Medicine).

To better understand how Euclidean calculus works, and what is actually going on, let us look at the tensor space. Tensors are symmetric matrices with a special condition on their eigenvalues. Thus, they live in a sub-space of the vector space of symmetric matrices Sym_n. We call the tensor space Sym_n^{+*}, with a '$+*$' representing the positive definite constraint. It is also denoted sometimes SPD (for Symmetric Positive Definite). For simplicity reasons, let us look at the space Sym_2^{+*} of 2×2 tensors. It is embedded in the space of 2×2 symmetric matrices Sym_2 which is actually isomorphic to $\mathbb{R}^3$ (a $n \times n$ symmetric matrix has $n(n+1)/2$ degrees of freedom). A 2×2 tensor has its two eigenvalues positive. Thus, the determinant of such tensor is positive:

$$D = \begin{pmatrix} x & y \\ y & z \end{pmatrix} \in Sym_2^{+*} \qquad \Longrightarrow \qquad \det(D) = x\,z - y^2 > 0$$

So this tensor space is entirely determined by this polynomial equation: $x z - y^2 > 0$, and its boundaries are given by: $x z - y^2 = 0$. Let us look at the function $f(x, y, z) = x\,z - y^2$. This is a quadric form that can be written:

$$f(x,y,z) = \begin{bmatrix} x & y & z \end{bmatrix} Q \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \text{with} \qquad Q = \begin{pmatrix} 0 & 0 & 0.5 \\ 0 & -0.5 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

The matrix Q can be decomposed into eigen elements to express the function f in a *normalized form*:

$$Q = PVP^\top \qquad \text{with} \qquad P = \begin{pmatrix} 0 & \sqrt{2}/2 & \sqrt{2}/2 \\ -1 & 0 & 0 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}, V = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}$$

Using the following change of coordinate: $[x', y', z']^\top = P\,[x, y, z]^\top$, we obtain a normalized form for f:

$$f(x', y', z') = -x'^2 - 0.5\,y'^2 + 0.5\,z'^2$$

Solving $f(x', y', z') = 0$ gives the following equation for the tensor space surface (in the coordinate system defined by P):

$$x'^2 + \frac{y'^2}{2} - \frac{z'^2}{2} = 0,$$

which is the equation of a *cone* with two halves joined at the apex.

This cone delimits the space of symmetric matrices with a null determinant (Fig. 1.9). Inside this cone, matrices have a positive determinant, which is given either by eigenvalues all positive or all negative. To determine which half corresponds to actual positive tensors, one can write the following. A tensor can be written in terms of eigenvectors/eigenvalues:

$$D = \lambda_1\, \overrightarrow{u}_1\, \overrightarrow{u}_1^\top + \lambda_2\, \overrightarrow{u}_2\, \overrightarrow{u}_2^\top = \begin{pmatrix} x & y \\ y & z \end{pmatrix}$$

By writing $\overrightarrow{u}_1 = [x_1, y_1]$ and $\overrightarrow{u}_2 = [x_2, y_2]$, we obtain $x = \lambda_1 x_1^2 + \lambda_2 x_2^2$. x is then the sum of positive numbers (eigenvalues are positive), which restrict the tensor space to the *positive quadrant* of $\mathbb{R}^3$. This is why the tensor space is sometimes called the *positive cone of symmetric matrices*.

So the tensor space is a cone. It is a convex space, which means that any convex combination of tensors is still a tensor. A convex combination is given by:

$$\alpha_1\, D_1 + \alpha_2\, D_2 + \ldots + \alpha_n\, D_n \qquad \text{with} \qquad \alpha_i \geq 0\, \forall i \qquad \text{and} \qquad \sum_{i=1}^{N} \alpha_i = 1$$

Examples of convex combinations include the (weighted) mean value and linear interpolation. However, any non-convex combination of tensors is not ensured to be a tensor. In general, if D_1 and D_2 are two tensors:

$$D_1 - D_2 \notin Sym_n^{+*}.$$

With the same spirit, the multiplication by a negative scalar leads out of the tensor space:

$$\forall \alpha \leq 0, \qquad \alpha.D \notin Sym_n^{+*}.$$

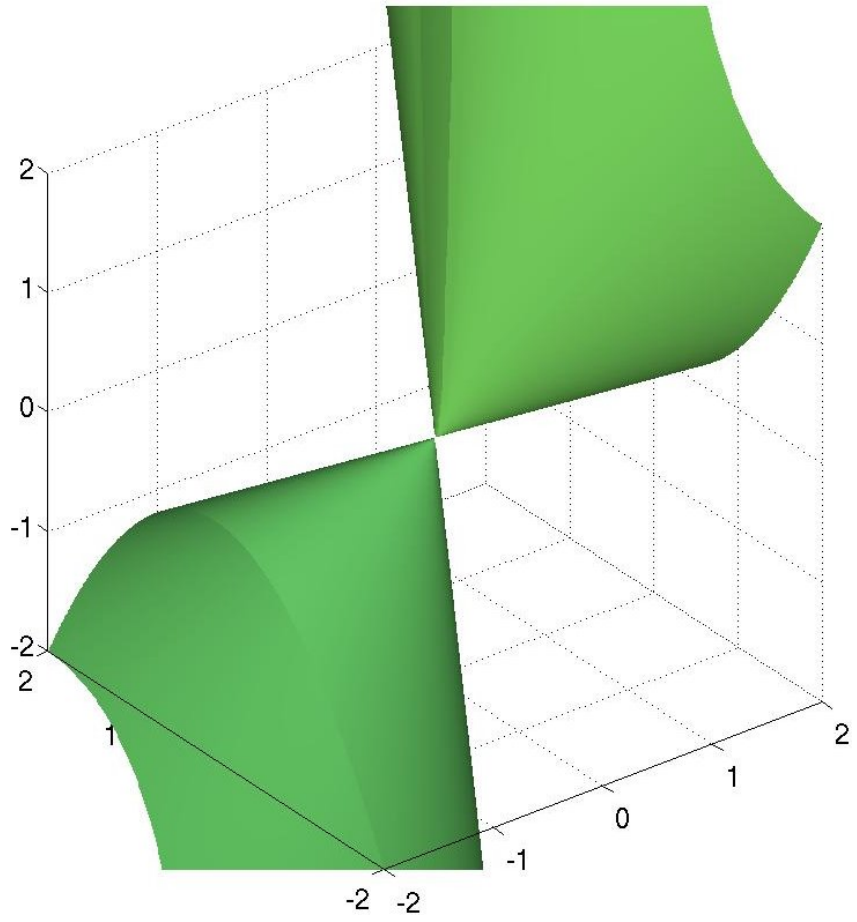


Figure 1.9: **The space of 2×2 tensors is the half of this cone located in the positive quadrant of $\mathbb{R}^3$.** The other half corresponds to matrices with all eigenvalues negative. Matrices on the surface have at least one null eigenvalue.

This lack of "vectorial" structure of the tensor space is problematic for many applications. For instance, extension of more elaborate vector-processing algorithms like image restoration using partial differential equations (PDEs) to tensors is not possible. It involves being able to do non-convex operations, which is not possible with Euclidean calculus. We must avoid at any cost to go through the boundaries of the tensor space. Is there any mathematical framework which allows to perform any operation on tensors while ensuring the result to be a tensor? If yes, what are the theoretical and practical properties of such framework(s)?

In summary, the three main themes that we investigate throughout this thesis are: tensor processing, diffusion MRI typical of clinical acquisition, and computational anatomy of the brain. Tensor processing is developed first, as this is the central tool that we use for the two ongoing applications. These two applications share a common ultimate goal, the modeling of the human brain, with a different

aspect for each. Below, we give an overview of the manuscript organization.

1.3 Manuscript Organization

This manuscript is organized in three parts reflecting the three main contributions of this thesis:

- The first part is a theoretical investigation on the alternative to Euclidean calculus to process tensors. In particular, we propose two original Riemannian frameworks for tensor computing which completely overcome the limitations of the Euclidean calculus. We demonstrate the feasibility of such frameworks with many examples on synthetic and real tensor data as those obtained in DT-MRI. This work has been published in three journal papers [Arsigny 2006b, Arsigny 2006c, Pennec 2006] (as well as two INRIA research reports [Pennec 2004b, Arsigny 2005b]), presented in one peer-reviewed conference [Arsigny 2005a], and one workshop [Fillard 2005a].

- The second part is a solution to a concrete problem: how can we optimally use and exploit diffusion tensor MRI typical from a clinical acquisition? Using such imaging modality is subject to special constraints: short scanning time leads to a limited number of acquisitions, and relatively old MRI scanners produces images with moderate and low signal-to-noise ratios. These constraints must be taken into account for a practical use of such modality in clinics. Here, we present specific methods for estimation and smoothing of DT-MRI of this type. We extensively use the frameworks developed in the first part. Finally, we show how the developed methodologies are actually integrated into a functional software called MedINRIA for an optimal transfer of the methods to end-users (medical experts). This work has been published in [Fillard 2007b, Ducreux 2007] (journal papers), [Fillard 2006a] (peer-reviewed conference), [Toussaint 2007b] (peer-reviewed workshop), and [Fillard 2005b] (INRIA research report).

- The third part is dedicated to the investigation of the modeling of brain anatomy, and in particular the cortex. We start by creating average models of certain sulcal lines that appear in all subjects by developing an original methods for curve matchings. Second, we model the individual variability of each sulcal positions by a variability tensor. We show how to extrapolate tensors that may be sparsely distributed to obtain a dense field, once again by using the frameworks developed in the first part. Third, we introduce new tools for measuring the anatomical correlations between any pairs of sulcal positions thanks to the total covariance matrix, a 6×6 tensor that is still compatible with our frameworks. We derive statistical

tests and end-up with promising new approaches for brain anatomy modeling. Publications relating this work are the following: [Fillard 2007c] (journal paper), [Fillard 2007a, Fillard 2005c] (conference papers), and [Fillard 2007d, Fillard 2006b] (INRIA research reports).

Below, we give more details about each part and a short description of each chapter.

Part I: A Riemannian Approach to Tensor Processing.

Chapter 3 is the core of the theoretical contribution of this thesis. We introduce the notions of Riemannian geometry necessary to follow the development that comes afterwards. We wanted to make this introduction as intuitive as possible to allow any one, even not familiar with this type of geometry, to be able to follow the rest of the chapter. Then, we propose two frameworks: the affine-invariant Riemannian metric and the Log-Euclidean metrics. Properties of each framework are analyzed and a comparison is made on their use in practice.

Chapter 4 is a demonstration of tensor processing using both frameworks. We go through a large number of image processing algorithms: (weighted) mean, multi-linear interpolation, Gaussian smoothing, anisotropic filtering, field restoration, all detailed with both metrics and compared to their Euclidean version. This chapter is a toolbox for tensor processing and contains all numerical tools to a practical implementation of the methods.

Chapter 5 concludes this part with a discussion on the choice of the metric. In particular, we show that in some cases (the structure tensor for example), the proposed frameworks are not adapted and fail. We left the reader with a recommendation: we strongly encourage him to figure out what type of properties he desires as a result of tensor operations, and then make the correct choice for the metric. This is summarized in this simple question: "Which metric for which application"?

Part II: An Optimal Workflow for Using DT-MRI in Clinical Applications.

Chapter 6 addresses the problem of diffusion tensor estimation and processing typical of clinical application. We demonstrate that thanks to the use of the proposed frameworks for tensor processing, and in particular the Log-Euclidean metrics, one can estimate tensors with a better prior on MRI noise. We show that MRI noise, which is Rician, induces a bias on the tensor estimates when the prior on the noise is wrong. Log-Euclidean metrics are extensively used here to solve this non-linear problem, and a comparison is made with two other noise models. Finally, we apply this method to synthetic and clinical data and show quantitative

improvements, as well a qualitatively better fiber reconstruction.

Chapter 7 describes the software MedINRIA. This project was started to transfer the methods developed at the chapters 4 and 6 to the clinic. We tried to meet the requirements of clinicians: they want a simple, ergonomic, reactive and efficient software. MedINRIA offers the possibility to process diffusion MRI, from estimation to fiber reconstruction, with a simple click, and contains high level algorithms, like Log-Euclidean tensor smoothing. This chapter describes the software architecture along with the set of applications which composes it. We show how MedINRIA has finally become a more general platform for medical image processing targetting the clinicians.

Chapter 8 is the accomplishment of the efforts put in methods and software development, as described in the two previous chapters. The context of the work presented here is clinical. Thanks to the help of a radiologist, we investigate the current and future indications of diffusion tensor MRI in spinal cord lesions. A large panel of pathologies are analyzed and DT-MRI, as well as tractography, appears to be more sensitive than classical T_2 imaging for detecting certain spine lesions.

This part is concluded in **Chapter 9** with a discussion on the future work and the future of diffusion tensors in the clinical world.

Part III: Statistical Analysis of the Human Brain Cortex Anatomy.

Chapter 10 introduces the notion of brain variability and some definitions needed by the other chapters. We also discuss the type of data we need to measure brain variability. In particular, we decide to rely on a set of sulcal lines that appear commonly in brain anatomy. We argue that using these lower dimensional landmarks helps recover inter-subject correspondences more easily than higher dimensional structures like surfaces or volumes. Second, as the sulcal lines are the locations of many functions, their study is of great interest to neuroscientists.

Chapter 11 describes, step by step, our strategy to build a second order model of brain variability. This model aims at giving the individual variation of any position of the cortex for a given population. We start by modeling each subject's sulcal line as the deformation of an individual average line. Once we obtained the mean lines, we extract along them variability tensors. Then, we show how to select a subset of meaningful tensors, and how to extrapolate them to obtain a dense field. This model is evaluated via two tests: the intra-sulcus recovery test (are we able to recover the variability of a full sulcal line just with a few tensors?), and the leave-one-out test (are we able to predict the variability of a missing line with neighboring information?).

Chapter 12 explores the potential correlations between any pairs of sulcal positions. We use the total covariance matrix between two points, and derive from it a correlation matrix. This type of covariance matrix is in fact a 6×6 tensor that fits in our tensor-processing framework. Thus, we are able to extrapolate it and obtain a dense map of correlations of one anatomical position to the rest of the brain. Then, we study the potential correlations between specific sulcal positions and the rest of the brain (namely three positions of the central and superior temporal sulci, which lead us to analyze the correlations between any point and its corresponding position in the opposite hemisphere. Very interesting new results are shown and discussed.

This part is concluded in **Chapter 13** with a discussion on the future work, especially on how these statistical information could be incorporated back to the registration as an a-priori.

Chapter 14 is the general conclusion section of the thesis. We summarize the contributions, list the publications of the author, and discuss about the future directions of this work. In particular, we focus on the need for validation of all methodologies described here, and for a systematic transfer of new technologies towards clinics.

Chapter 2

Introduction (français)

Contents

L'Homme est fasciné par le cerveau depuis des millénaires. Le rôle de cet organe mystérieux a cependant souvent été mal compris. En Égypte, vers la fin du Moyen Empire (deuxième siècle avant JC), lors de la momification, le cerveau était généralement ôté du corps par le nez, car on pensait que le siège de l'intelligence résidait dans le cœur. Selon Hérodote, un historien grecque (souvent considéré comme le premier historien de l'humanité), pendant les premières étapes de la momification : "La pratique la plus parfaite consiste à extraire le plus possible de cerveau avec un crochet en fer, et tout ce que le crochet ne peut pas atteindre est dissous à l'aide de substances". Bien évidemment, pendant les cinq mille années qui suivirent cette vue a été renversée et le cerveau a désormais le rôle qu'on lui connait. Cette première vision est néammoins restée dans le langage courant, notamment par le biais de l'expression : "Apprendre quelque chose par cœur".

Pendant le premier millénaire avant JC, les anciens grecques ont développé des points de vue différents sur la fonction du cerveau. Il est dit que c'est le Pythagorien Alcmaeon de Croton qui a été le premier à considérer que la pensée est localisée dans le cerveau. De manière anecdotique, pendant le 4^{me} siècle avant JC, Aristote pensait que le cerveau était un mécanisme de refroidissement du sang. Son argument était que les hommes sont plus raisonnés que les animaux parce qu'ils ont un plus grand cerveau pour refroidir leur "sang chaud".

Au cours des années 1800 eu lieu un débat entre scientifiques pour savoir si les régions du cerveau correspondent à des fonctions spécifiques, ou bien si le cerveau fonctionne comme un tout (théorie du "aggregate field"). Les travaux de Paul Broca,

Karl Wernicke et Korbinian Brodmann ont finalement contribué à prouver que les aires du cerveaux ont des fonctions bien spécifiques. Aujourd'hui, les neurosciences modernes sont en pleine expansion grâce à la disponibilité d'ordinateurs capables de traiter des informations d'une complexité sans précédent. Le dévelopement rapide de techniques d'imagerie comme l'imagerie par résonance magnétique (IRM), l'IRM fonctionelle (IRMf) et l'IRM du tenseur de diffusion (IRM-TD) permet aux scientifiques d'étudier le cerveau in-vivo d'une manière impossible qui était pour leurs prédécesseurs.

Jusqu'à présent, les connaissances sur le cerveau sont limitées par rapport à d'autres organes comme le cœur. Les erreurs qu'ont faits nos prédécesseurs paraissent sans doute risibles, mais sommes-nous sûrs que nos interprétations soient correctes? La majeure partie de nos connaissances sur l'anatomie du cerveau est un reccueil d'observations faites par des neuroanatomistes. Le cerveau tel que nous le connaissons est identique à la description qui en est faite dans le livre "Gray's anatomy" sorti en...1858 [Gray 1958]! Aujourd'hui, grâce à l'émergence de l'imagerie médicale pendant les 20 dernières années, un nouvel espoir de comprendre les mécanismes du cerveau est né. On peut l'observer in-vivo à des résolutions très fines avec les scanners modernes. On peut même l'observer en train de "fonctionner" avec l'IRMf ou l'imagerie PET. Il est temps de montrer à nos prédécesseurs que nous sommes capables de briser les secrets du cerveau. Ou bien nos successeurs riront à leur tour de nos erreurs. Seul l'avenir le dira.

2.1 La modélisation du cerveau humain : un challenge en neurosciences

Comprendre les mécanismes du cerveau, d'un point de vue informatique, est souvent synonyme de modélisation. Cependant, la complexité est telle qu'une modélisation algorithmique directe est quasi-impossible. En effet, le cerveau n'est pas un arrangement statique de circuits, mais plutôt un réseau de neurones vastement inter-connectés qui changent en permanence leur connectivité et leur sensibilité. Des travaux récents à la fois en neuroscience et en intelligence artificielle ont tenté de modéliser le cerveau en utilisant les outils mathématiques de la théorie du chaos et des systèmes dynamiques. Des recherches actuelles se sont également concentrées sur la création d'une structure neuronale du cerveau avec pour but la production de pensées humaines et d'intelligence artificielle, mais sans grand succès.

La construction de modèles algorithmiques du cerveau est le grand challenge des neurosciences. Ces modèles sont la clé pour une meilleure compréhension du fonctionnement du cerveau, et plus essentiellement, pourquoi certains cerveaux ne fonctionnent pas comme ils devraient. Notre connaissance est très minime sur de nombreuses pathologies comme Parkinson, Alzheimer, la schizophrénie,

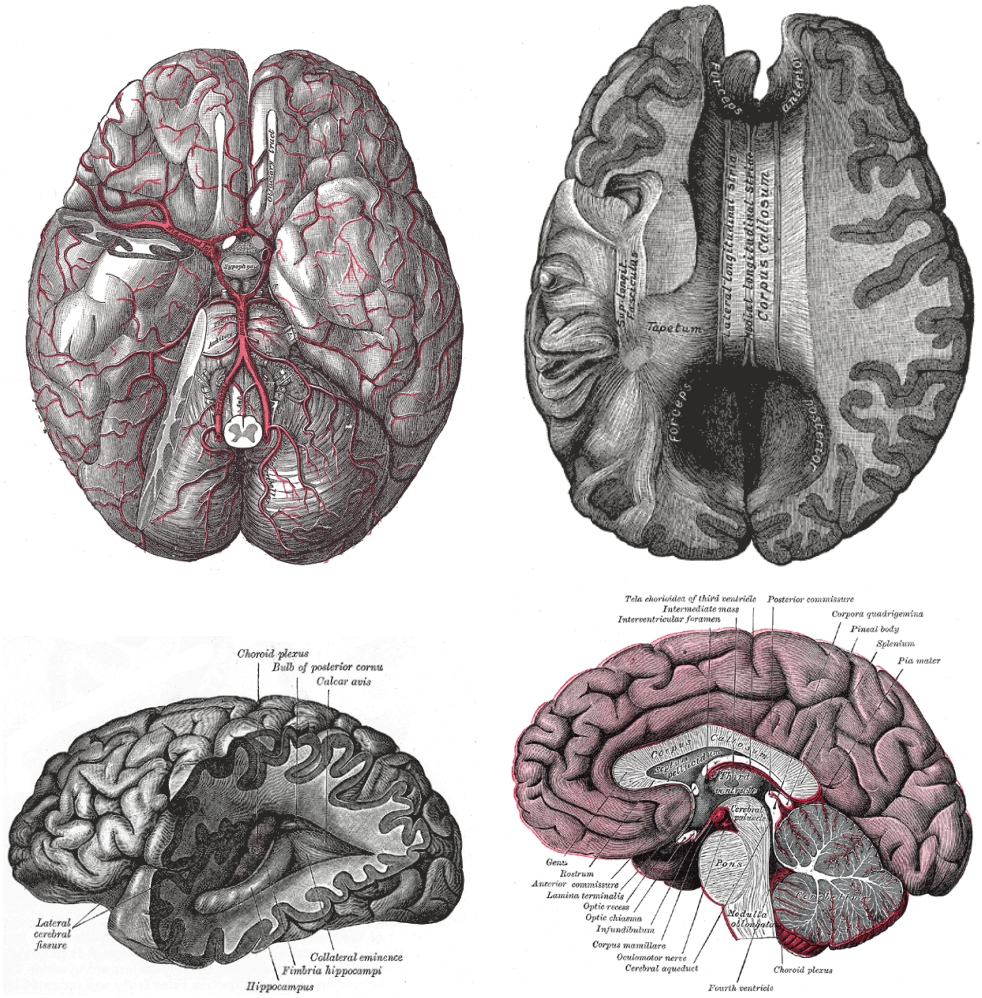


Figure 2.1: **Images du cerveau empruntées au livre "Gray's Anatomy" (1858).**

l'autisme, la sclérose en plaques, ou même certains cancers. Avoir une meilleure connaissance du fonctionnement cérébral normal nous aiderait à détecter et à traiter le plus tôt possible ces pathologies. Nos sociétés allant en vieillissant, les maladies neurologiques vont devenir de plus en plus fréquentes. L'impact de ces modèles, s'ils aboutissent à des traitements effectifs, serait énorme.

Comprendre le cerveau signifie également comprendre les changements du cerveau avec le temps, de la naissance à l'âge adulte, et de l'âge adulte au troisième âge. Une meilleure connaissance du processus normal de maturation cérébrale peut aider à détecter tôt et traiter des pathologies lourdes comme l'autisme ou la schizophrénie, qui apparaissent très souvent pendant la petite enfance ou l'adolescence. Un autre sujet très intéressant concerne la neuroplasticité, qui est la capacité qu'a le cerveau à changer son organisation (ce qui inclut la localisation des

zones fonctionnelles) avec l'expérience. Ce processus reste toujours très mystérieux et contient potentiellement la clé de la régénération fonctionnelle. On pourrait imaginer donner des médicaments à un patient souffrant d'amnésie, ou bien temporairement d'hémiplégie (sous-entendu que les nerfs sont toujours intacts), pour accélérer le processus de recouvrement de la mémoire et des fonctions motrices.

Il y a différents niveaux de modélisation qu'il est nécessaire de réaliser, et nous discuterons trois d'entre eux. Premièrement, il faut modéliser les fonctions, i.e., identifier avec précision quelles aires du cerveau sont responsables de quelles tâches fonctionnelles (comme l'aire de la parole). On notera que c'est une description très simplifiée du problème, puisque certaines fonctions plus complexes sont vraisemblablement distribuées entre plusieurs régions cérébrales (et non juste une). De plus, il existe une variabilité inter-individus sur la localisation de ces fonctions, ce qui rend le problème encore plus complexe. Néammoins, cette information fonctionelle est accessible en IRMf, qui est schématiquement capable de mesurer la réponse neuronale à un stimulus. Deuxièmement, nous devons modéliser l'anatomie, i.e., identifier (contourer) et modéliser la forme des structures du cerveau. La modélisation anatomique (et fonctionnelle) doit être réalisée sur une population de contrôles normaux pour caractériser ce qu'est la "normalité". L'IRM structurelle (ou anatomique) est le principal outil pour un tel niveau de modélisation. Finalement, il existe un troisième niveau de modélisation : la relation entre anatomie et fonction. Nous devons comprendre comment la fonction est liée à la structure et la forme. Cette modélisation fait le lien entre les deux premiers niveaux, et pourrait aider à comprendre comment les gènes, la neuroplasticité guidée par l'expérience, etc. affectent l'organisation du cerveau. L'émergence de l'IRM de diffusion, une modalité qui permet la reconstruction des faisceaux de fibres de la substance blanche, est un outil très prometteur pour mesurer de telles relations de dépendance.

2.2 Problèmes abordés dans cette thèse

La section précédente nous a permis d'introduire le thème principal de cette thèse qui est la modélisation du cerveau humain. Bien sûr, c'est un problème très complexe que l'on ne résolvera pas entièrement ici, et qui demande encore de nombreuses années de recherche collaborative entre différents domaines (médecine, neurosciences, physique, mathématiques et informatique) pour être résolu entièrement. Dans cette thèse, nous abordons deux aspects de ce problème : le traitement des IRMs de diffusion et l'anatomie algorithmique du cerveau, les deux jouant un rôle majeur dans le long trajet qui mène à une modélisation fine du cerveau.

Tout d'abord, l'IRM de diffusion, comme nous l'avons dit précédemment, est un outil prometteur pour étudier les relations entre fonctions et anatomie, car il donne accès à l'architecture des fibres de la substance blanche. Cependant, cette technique d'imagerie est relativement récente et pas assez mature pour être utilisée dans une

modélisation très précise. En effet, de nombreux problèmes inhérents à cette modalité doivent être résolus en premier, et en particulier dans un contexte clinique. Nous discutons ces limitations dans la section 2.2.1, et une première contribution de cette thèse est d'apporter une réponse à ces problèmes.

Ensuite, l'anatomie algorithmique est un nouveau domaine prometteur qui cible tout particulièrement l'étude d'un organe dans une population et dont le but est de créer un modèle de variabilité de cet organe. Nous introduisons l'anatomie algorithmique dans la section 2.2.2, et une seconde contribution de cette thèse est la construction de tels modèles pour le cerveau.

Ces deux domaines, qui peuvent sembler complètement indépendants à priori, ont quelque chose en commun : le tenseur. Les tenseurs sont des outils géométriques servant à étudier des processus statistiques et sont plus complexes à manipuler que de simples scalaires. Nous montrons dans les deux prochaines sections ce que sont les tenseurs, et quels rôles ils jouent en IRM de diffusion et en anatomie algorighmique. Dans la section 2.2.3, nous listons les limitations du calcul sur les tenseurs. Développer une solution pour pallier à ces limitations est le thème d'une autre contribution de cette thèse.

2.2.1 L'IRM du tenseur de diffusion : caractériser in-vivo les fibres de la substance blanche

L'émergence récente de l'IRM de diffusion, et en particulier de l'IRM du tenseur de diffusion [Basser 1994a], a été accueillie avec enthousiasme par la communauté médicale. En effet, l'IRM de diffusion est l'unique technique in-vivo qui renseigne quant à l'anisotropie des tissus. En pratique, l'IRM de diffusion consiste à mesurer, dans chaque direction de l'espace, la diffusion des molécules d'eau. Ce phénomène est plus connu sous le nom de *mouvement brownien* et a été formalisé par Einstein en 1905 [Einstein 1956]. À un niveau mascroscopique, ce phénomène est perçu comme une diffusion. Einstein a établi la relation entre le coefficient de diffusion et le déplacement des particules :

$$d = \frac{1}{2\tau} \left\langle R^\top R \right\rangle ,$$

où τ est le temps de diffusion, R est le vecteur déplacement ($R = r - r_0$, r_0 étant la position initiale d'une particule et r la position finale), et $<>$ dénote la moyenne sur toutes les particules. Le scalaire d est une mesure de la diffusion du milieu et dépend du type de particules ainsi que des propriétés du milieu. Cependant, il n'indique pas une diffusion potentiellement différente dans chacune des directions de l'espace.

Certains milieux ont des propriétés de diffusion différentes. En particulier, les fibres de la substance blanche du cerveau sont protégées par des gaines de myélines qui "obligent" les molécules d'eau à bouger dans une direction tangentielle à ces gaines plutôt que dans une direction perpendiculaire. En d'autres termes, les molécules d'eau bougent d'avantage le long des fibres que dans une direction trans-

verse. Dans ce cas, un simple scalaire n'est pas capable de caractériser la diffusion d'un tel tissu dit anisotrope. On le remplace alors par la matrice de covariance D du déplacement des particules :

$$D = \frac{1}{6\tau} \left\langle RR^\top \right\rangle .$$

D est une matrice 3×3, symétrique et définie positive. Le terme *tenseur de diffusion* a été utilisé pour la première fois par Stejskal et al. [Stejskal 1965] et provient du fait qu'il est formé du produit tensoriel du vecteur déplacement R avec lui-même. En notations mathématiques rigoureuses, un scalaire est un tenseur d'ordre 0, un vecteur est un tenseur d'ordre 1, et une matrice est un tenseur d'ordre 2 (sans restriction aucune sur sa symétrie et sur sa positivité). Ainsi, la dénomination *tenseur* est ambigüe car toute matrice est un tenseur d'ordre 2! Cependant, le terme a été adopté par la communauté et est utilisé comme tel depuis des années. A partir de maintenant, nous utiliserons le terme tenseur comme synonyme de matrice symétrique et définie positive.

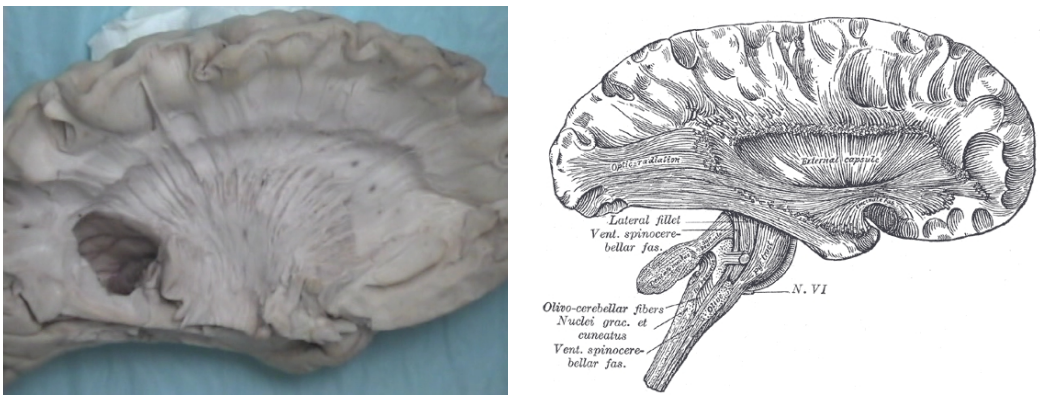


Figure 2.2: **Illustrations de fibres de la substance blanche du cerveau. Gauche :** Image de dissection d'un cerveau révélant certains des principaux faisceaux de fibres (courtoisie du Dr. Lennart Heimer, université de Virginie, USA). **Droite :** Image empruntée au livre Gray's Anatomy [Gray 1958].

Une matrice symétrique et définie positive est une matrice dont les valeurs propres sont toutes plus grandes que 0 strictement. Prenons une décomposition en éléments propres d'un tenseur D : $D = U\,V\,U^\top$. U est la matrice des vecteurs propres (c'est une matrice orthogonale, i.e., $U^{-1} = U^\top$) et V est la matrice des valeurs propres (c'est une matrice diagonale) :

$$U = \begin{pmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{pmatrix} \qquad \text{et} \qquad V = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

On dit que la valeur propre λ_i est associée au vecteur propre $\vec{u}_i$ pout tout i. De plus, lorsqu'on travaille avec des tenseurs de diffusion, on trie généralement

les valeurs propres par ordre décroissant, i.e., λ_i est plus grand que $\lambda_{i+1}\ \forall i$. Par conséquent, λ_1 est la plus grande valeur propre, et on appelle le vecteur propre associé $\overrightarrow{u}_1$ le vecteur propre principal (également appelé direction de diffusion principale - DDP - dans le cas des tenseurs de diffusion). Bien sûr, tout ceci est valable pour tout tenseur $n \times n$, pas seulement 3×3.

On utilise généralement un éllipsoïde pour représenter un tenseur (figure 2.3). En effet, les tenseurs ne sont rien d'autre que des formes quadriques. Une forme quadrique dont les valeurs propres sont toutes positives est un éllipsoïde. En réalité, la définition rigoureuse est qu'un éllipsoïde est une forme quadrique dont les valeurs propres sont toutes du même signe, positif ou négatif. Mais puisque nous nous intéressons au phénomène de diffusion, les valeurs propres sont homogènes à une distance carré par seconde, si bien que les valeurs propres négatives ne sont pas physiquement possibles. La forme et l'orientation de l'ellipsoïde sont déterminées par les éléments propres du tenseur : l'orientation de ses axes est donnée par les vecteurs propres, et l'amplitude le long de chaque axe est donnée par les valeurs propres correspondantes.

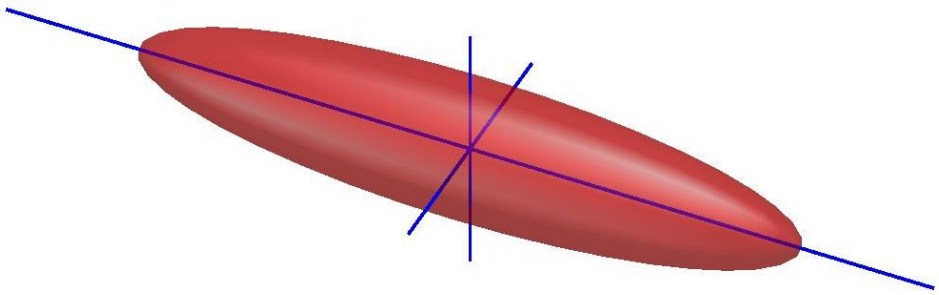


Figure 2.3: **Un tenseur représenté par un ellipsoïde.** Les directions des axes sont données par les vecteurs propres du tenseur. La longueur des axes est donnée par les valeurs propres.

Il convient d'insister sur le fait que le tenseur de diffusion est un *modèle* : il s'agit de la matrice de covariance du déplacement des particules. Ce n'est pas l'unique moyen de modéliser la diffusion. Premièrement, ce modèle assume une distribution gaussienne des déplacements des particules. La validité d'une telle hypothèse peut être remise en question, notamment dans des régions avec des croisements de fibres : dans ce cas, deux distributions gaussiennes se mélangent et rien n'assure le résultat d'être toujours gaussien. Un modèle plus élaboré serait nécessaire dans une telle situation et la littérature abonde d'alternatives au modèle tenseur : le modèle bi-tenseur [Alexander 2001a, Tuch 2002], le modèle CHARMED [Assaf 2005], l'imagerie Q-Ball [Tuch 2004, Descoteaux 2007b], la fonction de densité d'orientations des fibres [Tournier 2004a], ou même des tenseurs d'ordre supérieur [Barmpoutis 2007b]. Cependant, ces techniques nécessitent l'acquisition d'un très grand nombre d'IRMs (de plusieurs dizaines à plusieurs centaines). De

plus, si l'on veut une estimation robuste des paramètres de ces modèles, on doit très souvent répéter les acquisitions ou acquérir plus d'images que le nombre nécessaire. Cela peut être critique lorsque l'on travaille, par exemple, dans un environement clinique où le temps dácquisition doit être le plus court possible (en moyenne 15 minutes par patient). Nous apporterons une solution à ce problème plus tard dans ce manuscrit. Avec un modèle tenseur simple, six images seulement sont nécessaires (en fait sept puisqu'une image de référence est également indispensable) car le tenseur a six degrés de liberté et il y a donc six paramètres à estimer. Il est alors plus raisonnable d'augmenter la robustesse de l'estimation (i.e., multiplier les acquisitions) dans ce cas. Finalement, nous pensons que le modèle tenseur a encore beaucoup à offrir.

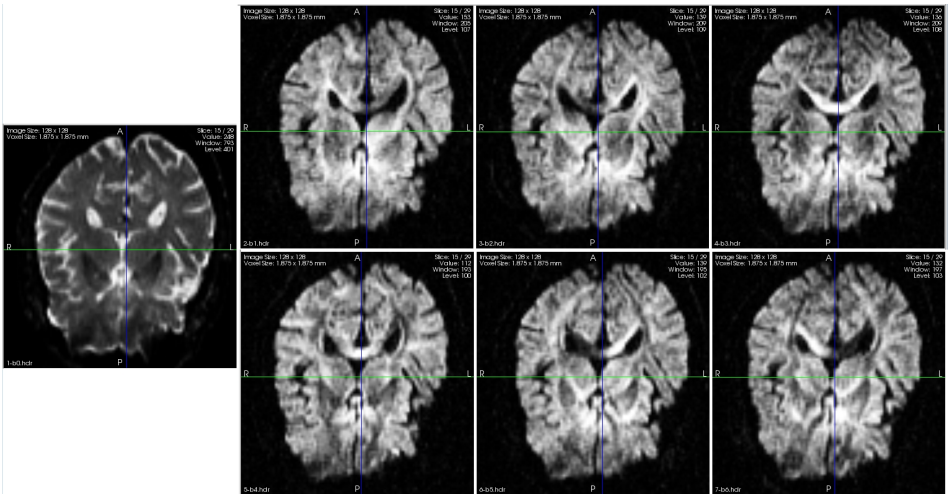


Figure 2.4: **Coupes d'une acquisition IRM-TD typique d'un environnement clinique.** (Courtoisie du Dr. Denis Ducreux, hôpital Bicêtre, Paris.) Dimensions : $128 \times 128 \times 30$, resolution spatiale : $1.8 \times 1.8 \times 4mm$. Le jeu de données est composé d'une image de référence (la plus à gauche) et de 6 images pondérées en diffusion (la séquence de Basser [Basser 1994a] est utilisée).

En pratique, le tenseur de diffusion D est relié à l'image pondérée en diffusion par l'équation de diffusion de Stejskal-Tanner :

$$S_i = S_0 \exp\left(-b \overrightarrow{g}_i^{\top} D \overrightarrow{g}_i\right), \tag{2.1}$$

où S_i est l'image pondérée en diffusion (IPD), S_0 est l'image de référence (i.e. une image sans gradient de diffusion, typiquement une image pondérée T_2), $\overrightarrow{g}_i$ le gradient de diffusion, et b la *valeur b* exprimée en $s.mm^{-2}$ qui dépend des paramètres d'acquisition (elle est généralement constante pour toute les IPD). Pour une description plus complète de l'acquistion d'IRM de diffusion, nous renvoyons le lecteur vers [Poupon 1999]. Schématiquement, le coefficient $\overrightarrow{g}_i^{\top} D \overrightarrow{g}_i$ est un scalaire donnant une mesure de la diffusion dans la direction spatiale $\overrightarrow{g}_i$. Il est

identique au coefficient de diffusion isotrope d que nous avons décrit précédemment avec l'exception qu'il est mesuré dans une direction spatiale donnée. $\overrightarrow{g}_i$ est appelé un *gradient de diffusion* car il correspond réellement à un gradient de champ magnétique utilisé dans la séquence IRM. S_i est le signal produit par le scanner. Le but de l'IRM du tenseur de diffusion est de reconstruire le champ de tenseurs de diffusion D à partir des mesures S_i. Une fois encore, puisque D a six degrés de liberté, au moins six images pondérées en diffusion sont nécessaires. L'estimation est réalisée généralement en linéarisant l'équation de diffusion 2.1 pour chacun des gradients de diffusion, et en résolvant ce système aux moindres carrés grâce à des méthodes algébriques. Une fois l'estimation effectuée, il est possible de représenter l'architecture de la substance blanche (dans le cas d'ITD du cerveau) par un processus appelé *tractographie*.

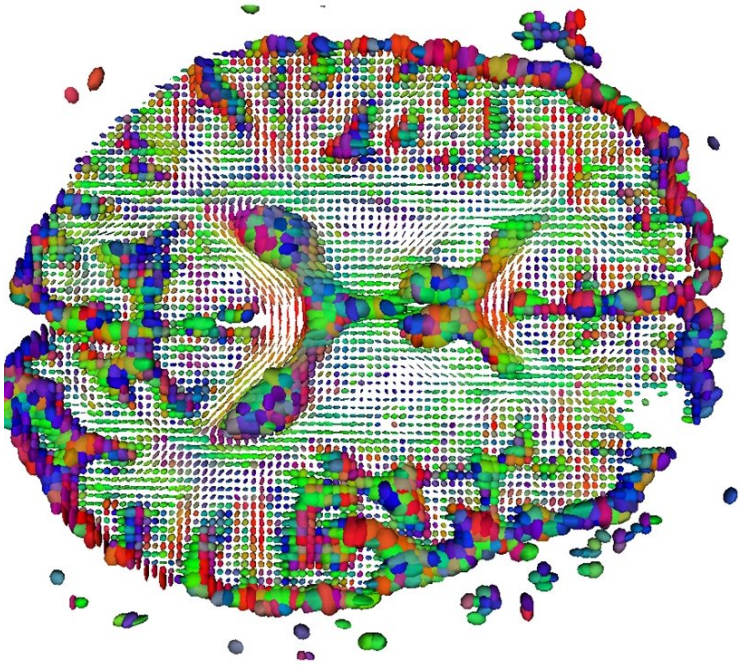


Figure 2.5: **Coupe d'un champ de tenseurs estimé à partir des IRM-TD de la figure 2.4 avec une technique classique.** On remarque que le champ apparait plutôt bruité, et que des éléments sont manquants (coin inférieur droit de la coupe) à cause des distorsions et du bruit d'acquisition.

La reconstruction de fibres (ou suivi de fibres, ou encore tractographie) consiste à construire une représentation géométrique des fibres de la substance blanche. On fait l'hypothèse que les tenseurs de diffusion sont alignés avec la direction des tissus orientés sous-jacents comme les fibres nerveuses dans le case de l'ITD du cerveau (bien que cette hypothèse soit questionable dans le cas de croisements de fibres). La littérature est abondante sur le suivi de fibres et on

peut classer les algorithmes en deux catégories : les algorithmes intégratifs et probabilistes. Schématiquement, dans les approches intégratives, le champ composé des directions de diffusion principale (i.e. le champ de vecteurs composé des vecteurs propres principaux des tenseurs) est intégré à partir de points sources et les chemins obtenus sont alors considérés comme des fibres. Les approches probabilistes simulent quant-à-elles de nombreuses fois le phénomène de diffusion à partir d'un point source et considèrent le chemin le plus probable comme une fibre.

Dans cette thèse, nous avons choisi de nous placer dans le contexte de l'IRM de diffusion clinique. La principale raison est qu'en France, très peu de scanners sont dédiés à la recherche et que la plupart des acquisitions ITD que nous obtenons proviennent d'un environnement clinique. De plus, les experts médicaux sont très intéressés par cette technique relativement récente car c'est la seule qui permette d'étudier in-vivo l'intégrité des fibres de la substance blanche. Il existe néammoins des contraintes inhérentes aux acquisitions cliniques qui doivent être prises en compte. Premièrement, le temps d'acquisition est plutôt court : 15 minutes maximum par partient en général. Pendant ce laps de temps, il est nécessaire d'acquérir les modalités d'image liées à la pathologie du patient (T_1, T_2, densité de protons, etc.). Et s'il reste du temps, la séquence d'IRM-TD peut être lancée. Cela laisse la plupart du temps entre 5 à 6 minutes pour les séquences de diffusion. La conséquence sur l'acquisition est directe : seul un faible nombre de gradients peut être acquis avec un rapport signal-sur-bruit moyen voir faible. Par exemple, sur un scanner Siemens Sonata à 1.5T, les acquisitions sont limitées à 25 gradients de diffusion avec une seule répétition. Le tenseur de diffusion est alors le meilleur modèle pour ce type d'acquisition, et passer à des modèles plus élaborés serait hasardeux. Finalement, comme toute acquisition d'image par résonance magnétique, l'IRM de diffusion est sujette au bruit d'acquistion. Ce bruit peut sérieusement affecter la qualité des tenseurs de diffusion (figure 2.5), et par conséquent le résultat de la tractographie (figure 2.6) : l'étape d'estimation peut produire des matrices qui ne sont pas des tenseurs (rien n'assure le résultat d'être défini positif, et le bruit, les distorsions de l'image, ou tout autre artefact lié à l'acquisition - mouvement, etc. - affecte la qualité des tenseurs), et le bruit ajoute une incertitude sur les directions et l'amplitude des tenseurs.

En conclusion, l'ITD clinique doit faire face à un certain nombre de problèmes : acquisitions avec un faible nombre de gradients (environ 25) et un rapport signal-sur-bruit relativement faible. Néammoins les cliniciens doivent être en mesure de reconstruire les fibres à partir des images d'un patient. Comment pourrions-nous exploiter de manière optimale les ITD typiques d'acquisitions cliniques et donner aux cliniciens la possibilité de reconstruire ces fibres?

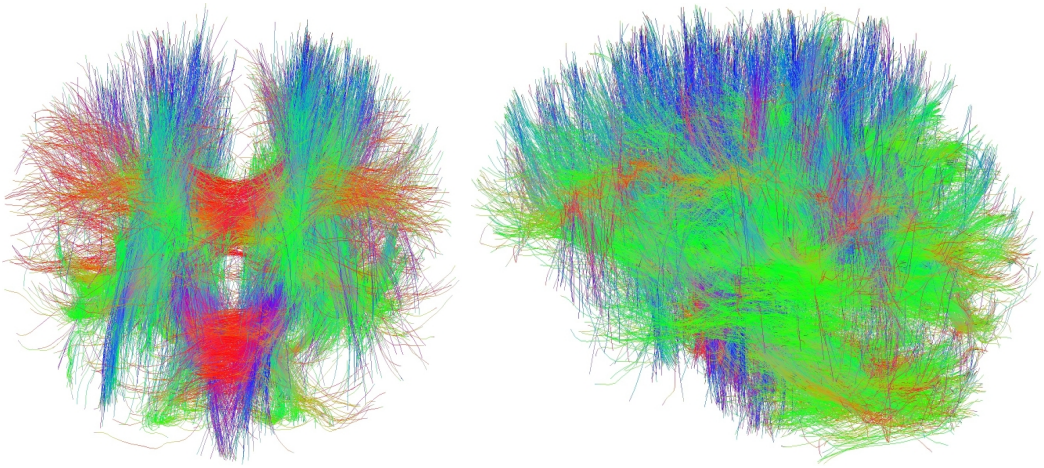


Figure 2.6: **Exemple de tractographie des données cliniques de la figure 2.4.** Les fibres ne sont pas réalistes à cause du bruit et de la résolution assez faible des images, en particulier dans la direction Z (dimensions : $128 \times 128 \times 30$, résolution spatiale : $1.8 \times 1.8 \times 4mm$). Le débruitage et l'interpolation spatiale deviennent nécessaires dans ce cas.

2.2.2 L'anatomie algorithmique : un domaine émergent prometteur

Le but de l'anatomie algorithmique est de développer des algorithmes pour modéliser et analyser les formes biologiques des tissus et des organes. Le but est non seulement d'estimer l'anatomie représentative d'un organe selon l'espèce, la population, la maladie, l'âge, etc. mais aussi de modéliser le développement de l'organe avec le temps (croissance) et d'établir sa variabilité. Un autre but est de corréler cette information de variabilité avec d'autres informations, qu'elles soient fonctionnelles, génétiques ou structurelles (par exemple les faisceaux de fibres obtenus en IRM du tenseur de diffusion). L'anatomie algorithmique est au croisement de la géométrie, des statistiques et de l'analyse d'images.

Trouver les correspondances point-à-point entre les organes de différents sujets est une tâche difficile du fait de la variabilité inter-individuelle. Lorsqu'on compare grossièrement deux cerveaux, on conclut qu'ils se ressemblent, i.e., qu'ils possèdent des replis corticaux, un cervelet et des ventricules par exemple. Mais si on s'attache à les comparer avec précision, on réalise que de nombreux détails les séparent. Par exemple, les schémas des replis sont très différents d'un individu à l'autre : des sillons majeurs apparaissent chez chacun (comme le sillon central), mais d'autres n'apparaissent que chez certains individus. Même des sillons et gyris communs ont une forme qui peut varier de manière drastique entre individus. Cela souligne le besoin de modèles physiques adaptés pour relier l'anatomie de différents sujets entre elles, afin d'apprendre les relations géométriques par des analyses statistiques. La méthode générale est de d'identifier des repères géométriques

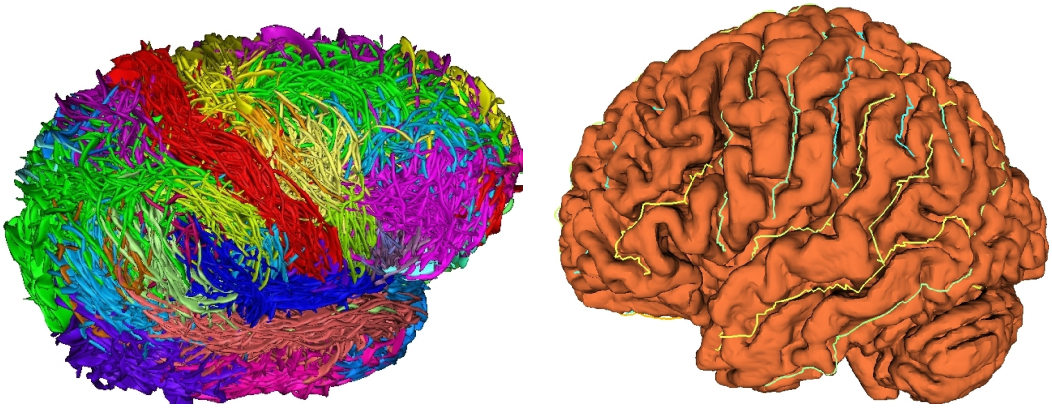


Figure 2.7: **Exemples de repères anatomiques utilisés en anatomie algorithmique. Gauche :** Les rubans corticaux de 50 sujets ont été extraits automatiquement avec le logiciel BrainVisa (courtoisie de J.-F. Mangin, Neurospin, CEA) et superposés dans le même système de coordonnées. **Droite :** Tracés manuels de lignes sulcales chez un sujet (la ligne sulcale est définie comme étant la ligne de "fond" du ruban cortical), courtoisie du Dr. Paul M. Thompson.

qui sont représentatifs de l'anatomie (points, lignes, surfaces, etc. - un exemple est donné à la figure 2.7) et de modéliser leur distribution statistique dans une population (figure 2.8). Cela peut être fait, par exemple, en calculant la forme moyenne et la structure de covariance après un alignement de groupe. Dans le cas du cerveau, on peut se reposer sur un large panel de repères anatomiques et fonctionels comprenant l'AC-PC (points) [Talairach 1988], les courbes (lignes de crêtes, lignes sulcales) [Mangin 2004b, Le Goualher 1999, Le Goualher 1997], les surfaces comme les rubans corticaux [Cachia 2003, Rivière 2002], les faisceaux de fibres [Corouge 2006], ou les images lorsqu'elles sont considérées comme des fonctions 3D [Arsigny 2006a, Glaunès 2005, Vaillant 2005].
Il existe des signes évidents que la variabilité inter-individuelle de ces repères est anisotrope. Cela signifie que la variabilité dépend de la direction spatiale, et qu'un simple scalaire (typiquement la variance) ne peut pas caractériser une telle anisotropie. On peut faire un parallèle avec l'IRM du tenseur de diffusion où le coefficient de diffusion est remplacé par un tenseur de diffusion : ici la matrice de covariance remplace la variance. Dans ce cas, la matrice de covariance caractérise l'anisotropie de la variabilité prise à chaque position individuelle de la représentation moyenne d'une forme. Nous appelons cette matrice un *tenseur de variabilité* par analogie avec le tenseur de diffusion.

Des problèmes communs aux tenseurs de diffusion se posent : comment peut-on manipuler ces tenseurs? Par exemple, on peut être amené à effectuer des opérations géométriques comme l'interpolation et l'extrapolation de tenseurs, opérations bien connues pour des images scalaires. Une fois encore, est-il possible de travailler avec

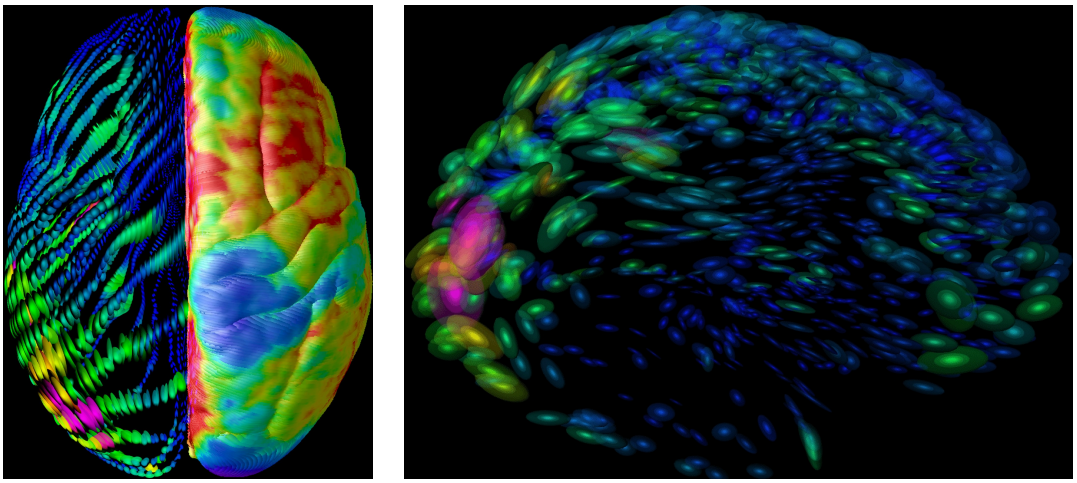


Figure 2.8: **Cartographie de la variabilité anatomique chez l'Homme.** Cette carte montre l'amplitude et les directions principales de la variabilité anatomique du cerveau en se basant sur un groupe de 40 IRMs de sujets normaux. La couleur rose indique les régions du cerveau où la variabilité anatomique inter-sujets est la plus grande, tandis que la couleur bleue indique les régions peu variables. Ces cartes aident à distinguer des changements dus à des pathologies des variations normales. Les images sont la courtoisie de Paul M. Thompson, Andrew Lee, Kiralee Hayashi, Agatha Lee, and Arthur Toga (LONI - UCLA School of Medicine).

les tenseurs comme nous le faisons avec de simples scalaires? C'est la question que nous nous posons dans la prochaine section.

2.2.3 Limitations du calcul euclidien sur les tenseurs

Nous appelons calcul euclidien sur les tenseurs le fait d'agir sur les coefficients du tenseur directement. On peut écrire un tenseur 3×3 de la manière suivante :

$$D = \begin{pmatrix} d_{xx} & d_{xy} & d_{xz} \\ d_{xy} & d_{yy} & d_{yz} \\ d_{xz} & d_{yz} & d_{zz} \end{pmatrix}$$

Le calcul euclidien consiste à traiter directement les coefficients du tenseur $(d_{xx}, d_{xy}, d_{xz}, d_{yy}, d_{yz}, d_{zz})$.

Afin de mieux comprendre comment le calcul euclidien fonctionne, et ce qu'il se passe réellement, intéressons-nous à l'espace des tenseurs. Les tenseurs sont des matrices symétriques avec une condition spéciale sur leurs valeurs propres. Ainsi, ils vivent dans un sous-espace de l'espace vectoriel des matrices symétriques Sym_n. Nous dénotons l'espace des tenseurs Sym_n^{+*}, avec un '$+*$' qui représente la contrainte définie positive. Il est également dénoté quelques fois SDP (pour Symétrique Défini

Positif). Pour des raisons de simplicité, nous nous reposons sur l'espace Sym_2^{+*} des tenseurs 2×2. Il est inclus dans l'espace des matrices symétriques 2×2 Sym_2, qui est en fait isomorphe à $\mathbb{R}^3$ (une matrice symétrique $n \times n$ a $n(n+1)/2$ degrés de liberté). Un tenseur 2×2 a ses deux valeurs propres strictement positives. Ainsi, le déterminant d'un tel tenseur est strictement positif :

$$D = \begin{pmatrix} x & y \\ y & z \end{pmatrix} \in Sym_2^{+*} \qquad \Longrightarrow \qquad \det(D) = x\,z - y^2 > 0$$

Donc, l'espace des tenseurs est entièrement défini par l'équation polynomiale : $x\,z - y^2 > 0$, et ses frontières sont données par : $x\,z - y^2 = 0$. Regardons maintenant la fonction $f(x, y, z) = x\,z - y^2$. Il s'agit d'une forme quadrique qui peut s'écrire de la manière suivante :

$$f(x, y, z) = \begin{bmatrix} x & y & z \end{bmatrix} Q \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \text{avec} \qquad Q = \begin{pmatrix} 0 & 0 & 0.5 \\ 0 & -0.5 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

La matrice Q peut être décomposée en éléments simples pour exprimer la fonction f sous une *forme normale* :

$$Q = PVP^\top \qquad \text{with} \qquad P = \begin{pmatrix} 0 & \sqrt{2}/2 & \sqrt{2}/2 \\ -1 & 0 & 0 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}, V = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}$$

En utilisant le changement de coordonnées suivant : $[x', y', z']^\top = P\,[x, y, z]^\top$, nous obtenons une forme normale pour f :

$$f(x', y', z') = -x'^2 - 0.5\,y'^2 + 0.5\,z'^2$$

Résoudre $f(x', y', z') = 0$ donne l'équation suivante de la surface de l'espace des tenseurs (dans le repère défini par P) :

$$x'^2 + \frac{y'^2}{2} - \frac{z'^2}{2} = 0,$$

qui est l'équation d'un *cône* à deux nappes.

Ce cône délimite l'espace des matrices symétriques dont le déterminant est nul (Fig. 2.9). A l'intérieur du cône, les matrices ont un déterminant positif, qui est obtenu par des valeurs propres toutes positives ou toutes négatives. Pour déterminer quelle partie correspond réellement aux tenseurs, on considère ce qui suit. Un tenseur peut s'écrire en termes de ses vecteurs propres/valeurs propres :

$$D = \lambda_1\,\overrightarrow{u}_1\,\overrightarrow{u}_1^\top + \lambda_2\,\overrightarrow{u}_2\,\overrightarrow{u}_2^\top = \begin{pmatrix} x & y \\ y & z \end{pmatrix}$$

En écrivant $\overrightarrow{u}_1 = [x_1, y_1]$ et $\overrightarrow{u}_2 = [x_2, y_2]$, on obtient $x = \lambda_1 x_1^2 + \lambda_2 x_2^2$. x est alors la somme de nombres positifs (les valeurs propres sont positives), ce qui restreint

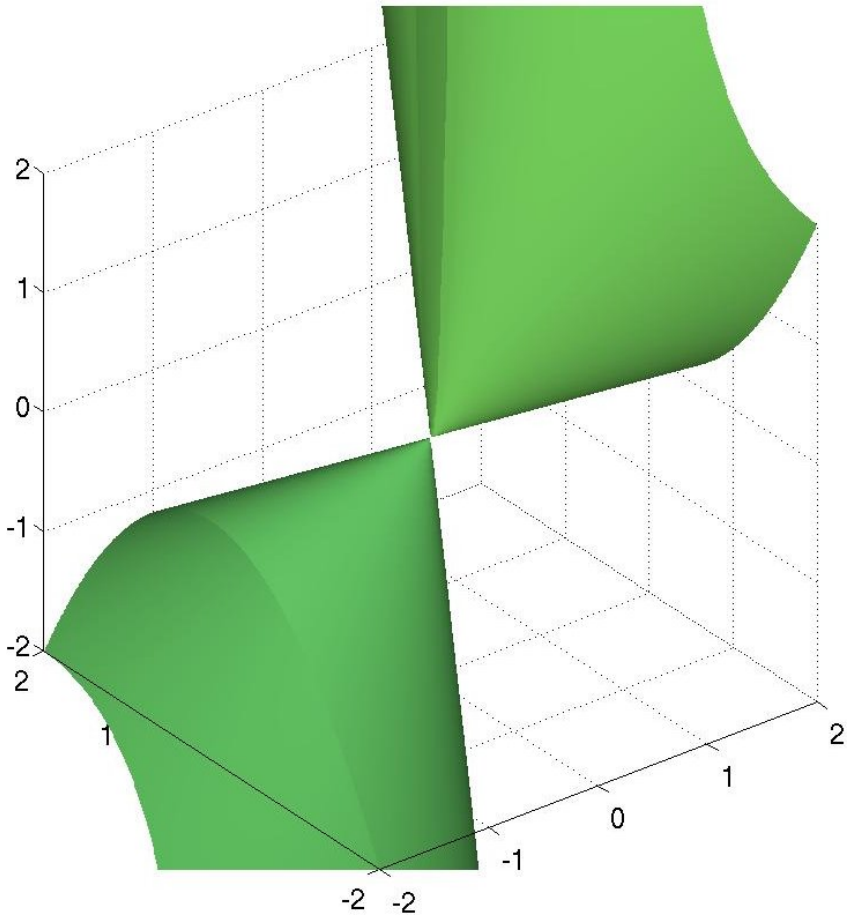


Figure 2.9: **L'espace des tenseurs 2×2 est le cône situé dans le quadrant positif de $\mathbb{R}^3$.** L'autre cône correspond aux matrices dont les valeurs propres sont toutes négatives. La surface représente les matrices dont au moins une des valeurs propres est nulle.

l'espace des tenseurs au *quadrant positif* de $\mathbb{R}^3$. C'est la raison pour laquelle l'espace des tenseurs est parfois appelé le *cône positif des matrices symétriques*.
Donc l'espace des tenseurs est un cône. C'est un espace convexe, ce qui signifie que toute combinaison convexe de tenseurs est toujours un tenseur. Une combinaison convexe est donnée par :

$$\alpha_1 D_1 + \alpha_2 D_2 + \ldots + \alpha_n D_n \qquad \text{avec} \qquad \alpha_i \geq 0 \, \forall i \qquad \text{et} \qquad \sum_{i=1}^{N} \alpha_i = 1$$

Parmi les opérations convexes, nous noterons le calcul de moyenne (pondérée) et l'interpolation linéaire. Cependant, toute opération non-convexe de tenseurs n'est pas assurée de produire un tenseur. En général, si D_1 et D_2 sont deux tenseurs :

$$D_1 - D_2 \notin Sym_n^{+*}.$$

Dans le même esprit, la multiplication par un scalaire négatif conduit en dehors de l'espace des tenseurs :

$$\forall \alpha \leq 0, \qquad \alpha.D \notin Sym_n^{+*}.$$

Ce manque de structure "vectorielle" est problématique dans de nombreuses applications. Par exemple, l'extension aux tenseurs d'algorithmes de traitement vectoriel plus élaborés comme la restoration d'image en utilisant des équations aux dérivées partielles (EDPs) n'est pas possible. Cela implique d'être capable de faire des opérations non-convexes, ce qui est impossible avec le calcul euclidien. On doit éviter à tout prix de traverser les frontières de l'espace des tenseurs. Existe-t-il un cadre mathématique qui autorise d'effectuer n'importe quelle opération sur les tenseurs tout en assurant le résultat d'être un tenseur? Si oui, quelles sont les propriétés théoriques et pratiques d'un tel cadre?

En résumé, les trois principaux thèmes que nous investiguons au travers de cette thèse sont : le traitement des tenseurs, l'IRM de diffusion typique des acquisitions cliniques, et l'anatomie algorithmique du cerveau. Le traitement des tenseurs est développé en premier, car il est le thème central autour duquel s'articulent les deux applications qui suivent. Ces deux applications partagent un but ultime, la modélisation du cerveau humain, en abordant un aspect différent chacun. Dans le chapitre suivant, nous donnons un aperçu de l'organisation du manuscrit.

2.3 Organisation du manuscrit

Ce manuscrit est organisé en trois parties reflétant les trois principales contributions de cette thèse :

- La première partie est une investigation théorique sur les alternatives au calcul euclidien pour traiter les tenseurs. En particulier, nous proposons deux cadres riemanniens originaux pour le traitement des tenseurs qui gomment les limitations du calcul euclidien. Nous démontrons l'utilité de tels cadres avec de nombreux exemples sur des données synthétiques et réelles comme celles obtenues en ITD. Ce travail a été l'objet de trois publications dans des revues internationales [Arsigny 2006b, Arsigny 2006c, Pennec 2006] (deux rapports de recherche sont également disponibles : [Pennec 2004b, Arsigny 2005b]), et a été présenté lors d'une conférence avec comité de relecture [Arsigny 2005a], et un workshop [Fillard 2005a].

- La seconde partie apporte une solution à un problème concret : comment peut-on utiliser et exploiter de manière optimale les IRM du tenseur de diffusion acquises en milieu clinique. Utiliser une telle modalité d'image est sujet à des contraintes particulières : un temps d'acquisition court qui mène à un nombre limité d'acquisitions, et des scanners relativement anciens produisant des images avec un rapport signal-sur-bruit moyen voir faible. Ces

contraintes doivent être prises en compte pour une réelle utilisation clinique d'une telle modalité d'image. Ici, nous présentons des méthodes dédiées d'estimation et de lissage des ITD de ce type. Nous utilisons de manière intensive les cadres de travail développés dans la première partie. Finalement, nous montrons comment les méthodologies proposées sont intégrées dans un logiciel fonctionnel appelé MedINRIA pour assurer un transfert optimal de ces méthodes vers les utilisateurs finaux (experts médicaux). Ce travail a été publié dans [Fillard 2007b, Ducreux 2007] (articles de revues), [Fillard 2006a] (conférence avec comité de relecture), [Toussaint 2007b] (workshop avec comité de relecture), et [Fillard 2005b] (rapport de recherche INRIA).

- La troisième partie est dédiée à la modélisation de l'anatomie du cerveau, et en particulier du cortex. Nous commençons par créer des modèles moyens de certaines lignes sulcales qui sont présentes chez tous les sujets en développant une méthode originale de mise en correspondance de courbes. Ensuite, nous modélisons la variabilité de chaque position sulcale par un tenseur de variabilité. Nous montrons comment extrapoler des tenseurs qui peuvent être distribués de manière épars pour obtenir un champ dense, une fois encore en utilisant les cadres de calcul développés à la première partie. Finalement, nous introduisons de nouveaux outils pour mesurer les corrélations anatomiques entre n'importe quelle paire de positions sulcales grâce à la matrice de covariance totale, un tenseur 6×6 qui reste compatible avec nos cadres de calcul. Nous concluons avec de nouvelles approches prometteuses pour la modélisation anatomique du cerveau. Les publications reliées à ces travaux sont les suivantes : [Fillard 2007c] (article de revue), [Fillard 2007a, Fillard 2005c] (conference avec comité de relecture), et [Fillard 2007d, Fillard 2006b] (rapport de recherche INRIA).

Nous donnons maintenant plus de détails sur chaque partie ainsi qu'une brève description de chaque chapitre.

Partie I : Un cadre riemannien pour le traitement des tenseurs.

Le chapitre 3 est le noyau de la contribution théorique de cette thèse. Nous introduisons les notions de géométrie riemannienne nécessaires pour suivre le dévelopement qui vient ensuite. Nous voulions rendre cette introduction la plus intuitive possible pour que chacun puisse suivre le reste du chapitre, y compris ceux qui découvrent ce type de géométrie pour la première fois. Ensuite, nous proposons deux cadres de calcul : la métrique affine-invariante et les métriques Log-Euclidienne. Les propriétés de chaque cadre sont analysées et une comparaison est faite sur leurs utilisations en pratique.

Le chapitre 4 est une démonstration du traitement des tenseurs utilisant

les deux cadres. Nous abordons un grand nombre d'algorithmes de traitement d'images : la moyenne (pondérée ou non), l'interpolation multi-linéaire, le lissage gaussien, le filtrage anisotrope, la restoration de champs, que nous déclinons selon les deux métriques et que nous comparons à la version euclidienne. Ce chapitre est en quelque sorte une boite à outils pour le traitement des tenseurs et contient les clés pour une implémentation pratique des méthodes.

Le chapitre 5 conclut cette partie avec une discussion sur le choix de la métrique. En particulier, nous montrons que dans certains cas (le tenseur de structure par exemple), les cadres que nous proposons ne sont pas adaptés. Nous donnons une recommandation au lecteur : nous l'encourageons fortement à déterminer quel type de propriétés il désire obtenir pour les résultat d'opérations sur les tenseurs, et de faire ainsi le bon choix de la métrique. Ceci est résumé par la simple question : "Quelle métrique pour quelle application"?

Partie II : Une chaîne de traitements optimale pour utiliser les IRM-TD en applications cliniques.

Le chapitre 6 apporte une solution au problème de l'estimation et du traitement des tenseurs de diffusion typique des acquisitions cliniques. Nous démontrons que grâce aux cadres de calcul sur les tenseurs que nous proposons, et en particulier des métriques Log-Euclidienne, il est possible d'estimer les tenseurs avec un meilleur a-priori sur le bruit IRM. Nous montrons que le bruit IRM, qui est ricien, induit un biais sur les tenseurs lorsque l'a-priori sur le bruit est faux. Les métriques Log-Euclidiennes sont utilisées de manière intensive pour résoudre ce problème non-linéaire, et nous faisons une comparaison avec deux autres modèles de bruit. Finalement, nous testons cette méthode sur des données synthétiques et cliniques et montrons une amélioration quantitative et qualitative de la reconstruction de fibres.

Le chapitre 7 décrit le logiciel MedINRIA. Ce projet a été initié pour transférer les méthodes développées aux chapitres 4 et 6 au monde clinique. Nous avons tenté de satisfaire aux mieux les besoins des cliniciens : ils veulent un logiciel qui soit simple, ergonomique, réactif et efficace. MedINRIA offre la possibilité de traiter les IRM de diffusion, de l'estimation au suivi de fibres, par le biais d'une simple clic de souris, et contient des algorithmes de haut niveau, comme le lissage des tenseurs utilisant les métriques Log-Euclidiennes. Ce chapitre décrit également l'architecture du logiciel, avec l'ensemble des applications qui le compose. Nous montrons comment MedINRIA est devenu maintenant une plateforme générique de traitement des images médicales ciblant les cliniciens.

Le chapitre 8 est l'aboutissement des efforts entrepris lors des développements méthodologiques et logiciel, comme décris dans les deux chapitres précédents. Le travail présenté dans ce chapitre s'inscrit dans un contexte clinique. Grâce à

l'aide d'un radiologue, nous avons investigué les indications courantes et futures de l'IRM du tenseur de diffusion dans les lésions de la moëlle épinière. Nous passons en revue un large panel de pathologies et l'IRM-TD ainsi que la tractographie apparaissent comme étant plus sensibles que la classique image T_2 pour détecter certaines pathologies de la moëlle.

Nous concluons cette partie avec le **chapitre 9** sur une discussion sur les travaux futurs et l'avenir du tenseur de diffusion dans le monde clinique.

Partie III : Analyse statistique de l'anatomie du cortex cérébral humain.

Le chapitre 10 introduit la motion de variabilité cérébrale et les définitions nécessaires dans les autres chapitres. Nous faisons également une discussion sur le type de données dont nous avons besoin pour mesurer la variabilité du cerveau. En particulier, nous décidons de nous reposer sur un ensemble de lignes sulcales qui sont présentes dans toutes les anatomies. Nous argumentons sur le fait qu'utiliser ces repères de faible dimension aide à recouvrer plus facilement les correspondances inter-sujets que d'utiliser des structures de plus grande dimension comme des surfaces ou des volumes. Ensuite, comme les lignes sulcales sont les lieux de nombreuses fonctions, les neuroscientifiques sont également très intéressés pour les étudier.

Le chapitre 11 décrit, pas à pas, notre stratégie pour construire un modèle du deuxième ordre de la variabilité du cerveau. Ce modèle a pour but de fournir la variation individuelle de chaque position du cortex chez une population donnée. Nous commençons par modéliser les lignes de chaque sujet comme la déformation de lignes moyennes. Une fois que les lignes moyennes sont obtenues, nous extrayons le long de ces lignes des tenseurs de variabilité. Ensuite, nous montrons comment sélectionner un sous-ensemble de tenseurs porteurs d'information, et comment les extrapoler pour obtenir un champ dense. Ce modèle est évalué via deux tests : le test de recouvrement intra-sillon (sommes-nous capables de retrouver la variabilité entière d'une ligne sulcale avec seulement quelques tenseurs?), et le test du leave-one-out (sommes-nous capables de prédire la variabilité d'une ligne manquante avec seulement des informations voisines?).

Le chapitre 12 explore les corrélations potentielles entre n'importe quelle paire de positions sulcales. Nous utilisons la matrice de covariance totale entre deux points, et nous en dérivons une matrice de corrélation. Ce type de matrice de covariance est en fait un tenseur 6×6 qui est tout à fait compatible avec notre cadre de traitement des tenseurs. Par conséquent, nous sommes capables de l'extrapoler et d'obtenir une carte dense des corrélations entre une position anatomique et le reste du cerveau. Ensuite, nous étudions les corrélations potentielles entre des positions sulcales bien spécifiques et le reste du cerveau (trois positions des sillons

central et temporel supérieur), ce qui nous amène à analyser les corrélations entre n'importe quel point et sa position correspondante dans l'hémisphère opposé. Nous présentons et discuronts des résultats novateurs.

Nous concluons cette partie dans le **chapitre 13** avec une discussion sur les travaux futurs, et en particulier comment ces informations statistiques pourraient être réinjectées comme a-priori dans le recalage.

Le chapitre 14 est la conclusion générale de cette thèse. Nous résumons les contributions, nous listons les publications de l'auteur, et nous discutons les futures directions de ce travail. En particulier, nous insistons sur le besoin de validation des méthodes décrites ici, et sur le besoin d'un transfert systématique de nouvelles technologies vers l'hôpital.

Part I

A Riemannian Approach to Tensor Processing

"Do not worry about your difficulties in Mathematics.
I can assure you mine are still greater."
Albert Einstein.

Medical image processing and analysis, as in the context of the computational anatomy, has to deal with data of a complex nature: tensors, rigid and affine transforms, or even diffeomorphisms to cite only a few examples. These data have the serious drawback not to live in vector spaces, like simple scalars, and should be manipulated with extreme care. For instance, adding two diffeomorphisms does not return in general a diffeomorphism. Instead, these data rather live in finite or infinite *manifolds*. Manifolds can be seen as curved spaces, locally very close to a vector space. It means that if we get infinitesimally close to the manifold, the curvature can be neglected and the manifold can be approximated by a vector space. A famous example of this approximation is...the surface of the earth! We know that our planet is curved, however we keep drawing planar maps. Doing this implies that we neglect its curvature, which is reasonable in a sufficiently small region (like France). Riemannian geometry has gained a lot of interest within the medical imaging community as it provides the necessary tools to work with manifolds. It turns out that the properties and tools of such non-linear geometry offers a very elegant, efficient, and mathematically grounded framework to manipulate data living in manifolds. Of course, playing with Riemannian geometry is more complex than with the classical Euclidean geometry, but we will show that reasoning in terms of this geometry is not more complex than drawing a map of the earth.

In Chapter 3, we start by exploring the state-of-the-art in tensor processing, and the basis of Riemannian geometry. Then, we propose two Riemannian frameworks as alternatives to the Euclidean calculus on tensors. In Chapter 4, we demonstrate the practical feasibility of these two Riemannian frameworks with geometric operations on tensors like interpolation or filtering (on both synthetic and real datasets). We conclude in Chapter 5 with a discussion on the choice of the metric with respect to the application.

Chapter 3

Two Riemannian Frameworks for Tensor Processing

Contents

3.1 Related Work

Quite an impressive literature has now been issued on processing of tensor fields especially in the context of Diffusion Tensor Imaging (DTI) regularization [Basser 1994a, Le Bihan 2001, Westin 2002]. Most of the works dealing with the geometric nature of tensors has been performed for the discontinuity-preserving regularization of tensor fields using Partial Differential Equations (PDEs). For instance, [Coulon 2004] anisotropically restores the principal direction of the tensor, and uses this regularized directions map as input for the anisotropic regularization of the eigenvalues. A quite similar idea is adopted in [Tschumperlé 2002], where a spectral

decomposition $\Sigma(x) = U(x)\, D(x)\, U(x)^{\mathrm{T}}$ of the tensor field is performed at each points to independently regularize the eigenvalues and eigenvectors (orientations). This approach requires an additional reorientation step of the rotation matrices due to the non-uniqueness of the decomposition (each eigenvector is defined up its sign and there may be joint permutations of the eigenvectors and eigenvalues) in order to avoid the creation of artificial discontinuities (see Fig. 3.1). Another problem arises

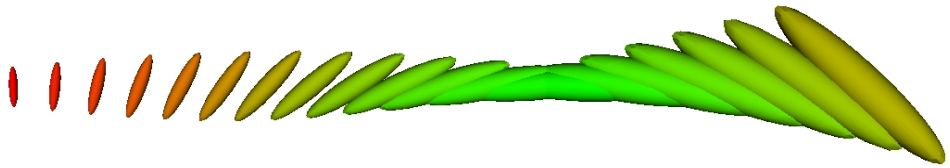


Figure 3.1: **Independent interpolation of tensor eigenvectors and eigenvalues.** A re-orientation step is necessary to prevent this behavior to occur: the interpolated tensors seem not to take the most direct path between the two extremities.

when two or more eigenvalues become equal: a whole subspace of unit eigenvectors is possible, and even a re-orientation becomes difficult.

Other possible frameworks for tensor processing include the J-divergence [Wang 2004b], which is a symmetrized version of the Kullback-Leibler divergence between distributions. While the J-divergence gives an affine-invariant dissimilarity measure between tensors, it does not provide a complete framework to perform interpolation, gradient descent and so on. Another interesting framework is the Cholesky decomposition of tensors [Wang 2004a]. The Cholesky decomposition of symmetric positive-definite matrices states that a tensor can be written in terms of the product of a lower triangular matrix L: $\Sigma(x) = L(x)\, L(x)^{\mathrm{T}}$. As the space of lower triangular matrices is a vector space, one can run computations on the field $L(x)$, and go back to the tensor space using $L(x)\, L(x)^{\mathrm{T}}$. However, we will show later that it does not overcome all limitations of the Euclidean calculus.

3.2 Basic Tools of Riemannian Geometry

We summarize in this Section the theory of statistics on Riemannian manifolds developed in [Pennec 1999, Pennec 2004a]. The aim is to demonstrate the fact that choosing a Riemannian metric "automatically" determines a powerful framework to work on the manifold through the use of a few tools from differential geometry.

In the geometric framework, one can specify the structure of a manifold $\mathcal{M}$ by a *Riemannian metric*. This is a continuous collection of scalar products on the tangent space at each point of the manifold. Thus, if we consider a curve on the manifold, we can compute at each point its instantaneous speed vector and its norm, the instantaneous speed. To compute the length of the curve, we can proceed as usual by integrating this value along the curve. The distance between two points of a connected Riemannian manifold is the minimum length among the curves joining these points. The curves realizing this minimum for any two points

of the manifold are called geodesics. The calculus of variations shows that geodesics are the solutions of a system of second order differential equations depending on the Riemannian metric. In the following, we assume that the manifold is *geodesically complete*, i.e. that the definition domain of all geodesics can be extended to $\mathbb{R}$. This means that the manifold has no boundary nor any singular point that we can reach in a finite time. As an important consequence, the Hopf-Rinow-De Rham theorem states that there always exists at least one minimizing geodesic between any two points of the manifold (i.e. whose length is the distance between the two points).

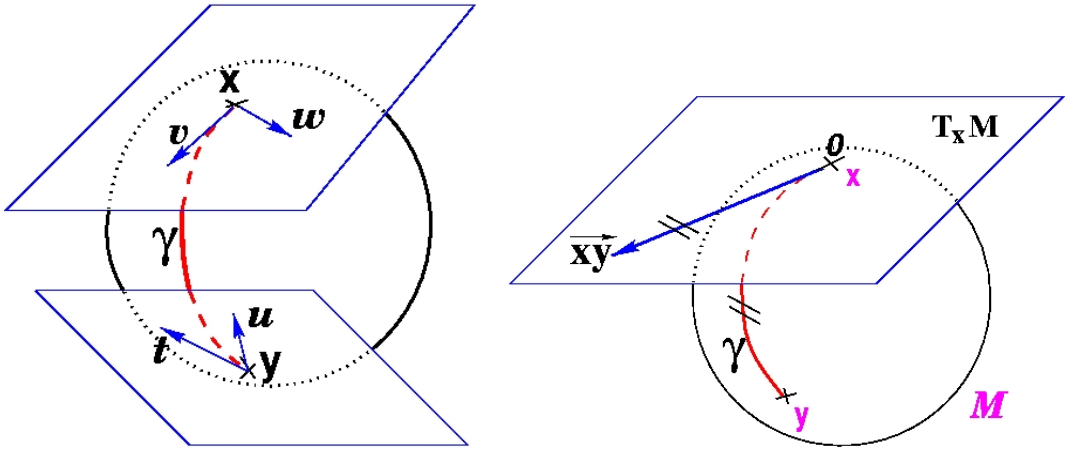


Figure 3.2: **Left:** The tangent planes at points x and y of the sphere $\mathcal{S}_2$ are different: the vectors v and w of $T_x\mathcal{M}$ cannot be compared to the vectors t and u of $T_y\mathcal{M}$. Thus, it is natural to define the scalar product on each tangent plane. **Right:** The geodesics starting at x are straight lines in the exponential map and the distance along them is conserved.

3.2.1 Exponential chart

Let x be a point of the manifold that we consider as a local reference and $\overrightarrow{xy}$ a vector of the tangent space $T_x\mathcal{M}$ at that point. From the theory of second order differential equations, we know that there exists one and only one geodesic starting from that point with this tangent vector. This allows us to develop the manifold in the tangent space along the geodesics (think of rolling a sphere along its tangent plane at a given point). The geodesics going through the reference point are transformed into straight lines and the distance along these geodesics is conserved (at least in a neighborhood of x).

The function that maps to each vector $\overrightarrow{xy} \in T_x\mathcal{M}$ the point y of the manifold that is reached after a unit time by the geodesic starting at x with this tangent vector is called the *exponential map*. This map is defined in the whole tangent space $T_x\mathcal{M}$ (since the manifold is geodesically complete) but it is generally one-to-one only locally around 0 in the tangent space (i.e. around x in the manifold). In the sequel, we denote by $\overrightarrow{xy} = \log_x(y)$ the inverse of the exponential map: this is the smallest vector such that $y = \exp_x(\overrightarrow{xy})$. If we look for the maximal definition domain, we

find out that it is a star-shaped domain delimited by a continuous curve C_{x} called the *tangential cut-locus*. The image of C_{x} by the exponential map is the cut locus $\mathcal{C}_{\mathrm{x}}$ of point x. This is the closure of the set of points where several minimizing geodesics starting from x meet. On the sphere $\mathcal{S}_2(1)$ for instance, the cut locus of a point x is its antipodal point and the tangential cut locus is the circle of radius π.

The exponential map within this domain realizes a chart called *the exponential chart*. It covers all the manifold except the cut locus of the reference point x, which has a null measure. In this chart, geodesics starting from x are straight lines, and the distance from the reference point are conserved. This chart is somehow the "most linear" chart of the manifold with respect to the reference point x.

3.2.2 Practical implementation

In fact, most of the usual operations using additions and subtractions may be reinterpreted in a Riemannian framework using the notion of *bipoint*, an antecedent of vector introduced during the 19th century. Indeed, one defines vectors as equivalent classes of bipoints (oriented couples of points) in a Euclidean space. This is possible because we have a canonical way (the translation) to compare what happens at two different points. In a Riemannian manifold, we can still compare things locally (by parallel transportation), but not any more globally. This means that each "vector" has to remember at which point of the manifold it is attached, which comes back to a bipoint.

However, one can also see a vector $\overrightarrow{\mathrm{xy}}$ (attached at point x) as a vector of the tangent space at that point. Such a vector may be identified to a point on the manifold using the geodesic starting at x with tangent vector $\overrightarrow{\mathrm{xy}}$, i.e. using the exponential map: $\mathrm{y} = \exp_{\mathrm{x}}(\overrightarrow{\mathrm{xy}})$. Conversely, the logarithmic map may be used to map almost any bipoint (x, y) into a vector $\overrightarrow{\mathrm{xy}} = \log_{\mathrm{x}}(\mathrm{y})$ of $T_{\mathrm{x}}\mathcal{M}$. This reinterpretation of addition and subtraction using logarithmic and exponential maps is very powerful to generalize algorithms working on vector spaces to algorithms on Riemannian manifolds, as illustrated by Table 3.1. It is also very powerful in terms of implementation since we can practically express all the geometric operations in these terms: the implementation of $\log_{\mathrm{x}}$ and $\exp_{\mathrm{x}}$ is the basis of any programming on Riemannian manifolds, as we will see in the following.

3.2.3 Basic statistical tools

The Riemannian metric induces an infinitesimal volume element on each tangent space, and thus a measure $d\mathcal{M}$ on the manifold that can be used to measure random events on the manifold and to define the probability density function (if it exists) of these random elements. With the probability measure of a random element, we can integrate functions from the manifold to any vector space, thus defining the expected value of this function. However, we generally cannot integrate manifold-valued functions. Thus, one cannot define the mean or expected "value" of a random manifold element using a weighted sum or an integral as usual. One solution is to

rely on a distance-based variational formulation: the Fréchet or Karcher expected features basically minimize globally (or locally) the variance. As the mean is now defined through a minimization procedure, its existence and uniqueness are not ensured any more (except for distributions with a sufficiently small compact support). In practice, one mean value almost always exists, and it is unique as soon as the distribution is sufficiently peaked. The properties of the mean are very similar to those of the modes (that can be defined as central Karcher values of order 0) in the vectorial case.

To compute the mean value, [Pennec 1999, Pennec 2004a] designed an original Gauss-Newton gradient descent algorithm that essentially alternates the computation of the barycenter in the exponential chart centered at the current estimation of the mean value, and a re-centering step of the chart at the point of the manifold that corresponds to the computed barycenter (geodesic marching step). To define higher moments of the distribution, they used the exponential chart at the mean point: the random feature is thus represented as a random vector with null mean in a star-shaped domain. With this representation, there is no difficulty to define the covariance matrix and potentially higher order moments. Based on this covariance matrix, they defined a Mahalanobis distance between a random and a deterministic feature that basically weights the distance between the deterministic feature and the mean feature using the inverse of the covariance matrix. Interestingly, the expected Mahalanobis distance of a random element with itself is independent of the distribution and is equal to the dimension of the manifold, as in the vectorial case.

As for the mean, [Pennec 1996, Pennec 1999, Pennec 2004a] chose a variational approach to generalize the Normal Law: they define it as the distribution that minimizes the information knowing the mean and the covariance. This amounts to consider a Gaussian distribution on the exponential chart centered at the mean point that is truncated at the cut locus (if there is one). However, the relation between the concentration matrix (the "metric" used in the exponential of the probability density function) and the covariance matrix is slightly more complex than the simple inversion of the vectorial case, as it has to be corrected for the curvature of the manifold. Last but not least, using the Mahalanobis distance of a normally distributed random feature, the χ^2 law can be generalized: they were able to show

	Vector space	Riemannian manifold
Subtraction	$\overrightarrow{xy} = y - x$	$\overrightarrow{\mathrm{xy}} = \log_{\mathrm{x}}(\mathrm{y})$
Addition	$y = x + \overrightarrow{xy}$	$\mathrm{y} = \exp_{\mathrm{x}}(\overrightarrow{\mathrm{xy}})$
Distance	$\mathrm{dist}(x,y) = \|y - x\|$	$\mathrm{dist}(\mathrm{x},\mathrm{y}) = \|\overrightarrow{\mathrm{xy}}\|_{\mathrm{x}}$
Mean value (implicit)	$\sum_i \overrightarrow{\bar{x}x_i} = 0$	$\sum_i \log_{\bar{\mathrm{x}}}(\mathrm{x}_i) = 0$
Gradient descent	$x_{t+\varepsilon} = x_t - \varepsilon \nabla C(x_t)$	$\mathrm{x}_{t+\varepsilon} = \exp_{\mathrm{x}_t}(-\varepsilon \nabla C(x_t))$
Linear (geodesic) interpolation	$x(t) = x_1 + t\,\overrightarrow{x_1 x_2}$	$\mathrm{x}(t) = \exp_{\mathrm{x}_1}(t\,\overrightarrow{\mathrm{x}_1\mathrm{x}_2})$

Table 3.1: Re-interpretation of basic standard operations in a Riemannian manifold.

that is has the same density as in the vectorial case up to an order 3 in σ. This opens the way to the generalization of many other statistical tests, as we may expect similarly simple approximations for sufficiently centered distributions.

3.3 An Affine-Invariant Riemannian Metric for Tensors

The affine-invariant Riemannian metric we detail in this section may be traced back to the work of [Nomizu 1954] on affine invariant connections on homogeneous spaces. It is implicitly hidden under very general theorems on symmetric spaces in many differential geometry textbooks [Kobayashi 1969, Helgason 1978, Gamkrelidze 1991] and sometimes considered as a well known result as in [Bhatia 2003]. In statistics, it has been introduced as the Fisher information metric [Skovgaard 1984] to model the geometry of the multivariate normal family. Interestingly, the invariant metric has been independently proposed by [Förstner 1999] to deal with covariance matrices, and recently by [Fletcher 2004] for the analysis of principal modes of sets of diffusion tensors. By looking for a suitable metric on the space of Gaussian distributions for the segmentation of diffusion tensor images, [Lenglet 2006b] also ended up with the same metric. They further used it to denoise tensor field using an anisotropic regularization scheme [Castaño-Moraga 2006, Castaño-Moraga 2007], which is similar in essence to the one we proposed in [Pennec 2006] and which will be detailed later in this manuscript (Sec. 4.3). Nethertheless, it is interesting to see that completely different approaches, relying on an affine-invariant requirement on the one hand, and relying on an information measure to evaluate the distance between distributions on the other hand, lead to the same metric on the tensor space. However, to our knowledge, this Riemannian metric has not been promoted as a complete computing framework, as we propose here.

Let us now focus on the space $\mathcal{S}ym_n^+$ of positive definite symmetric matrices (tensors). The goal is to find a Riemannian metric with interesting enough properties. It turns out that it is possible to require an invariance by the full linear group. This leads to a very regular manifold structure where tensors with null and infinite eigenvalues are both at an infinite distance of any positive definite symmetric matrix: the cone of positive definite symmetric matrices is replaced by a space which has an infinite development in each of its $n(n+1)/2$ directions. Moreover, there is one and only one geodesic joining any two tensors, and we can even define globally consistent orthonormal coordinate systems of tangent spaces. Thus, the structure we obtain is very close to a vector space, except that the space is curved.

3.3.1 Exponential, logarithm and square root of tensors

In the following, we will make extensive use of a few functions on symmetric matrices. The exponential of any matrix can be defined using the series $\exp(M) = \sum_{k=0}^{\infty} \frac{M^k}{k!}$. In the case of symmetric matrices, we have some important simplifications. Let $\Sigma = U D U^T$ be a diagonalization, where U is an orthonormal matrix, and $D = \mathrm{DIAG}(d_i)$ is the diagonal matrix of the eigenvalues. We can write any power of Σ in the same

basis: $\Sigma^k = U\, D^k\, U^{\mathrm{T}}$. This means that we may factor out the rotation matrices in the series and map the exponential individually to each eigenvalue:

$$\exp(\Sigma) = \sum_{k=0}^{+\infty} \frac{W^k}{k!} = U\ \mathrm{DIAG}(\exp(d_i))\ U^{\mathrm{T}}.$$

The series defining the exponential function converges for any (symmetric) matrix argument, but this is generally not the case for the series defining its inverse function: the logarithm. However, any tensor can be diagonalized into $\Sigma = U\,\mathrm{DIAG}(d_i)\,U^{\mathrm{T}}$ with strictly positive eigenvalues d_i. Thus, the function

$$\log(\Sigma) = U\,(\mathrm{DIAG}(\log(d_i)))\,U^{\mathrm{T}}$$

is always well defined on tensors. Moreover, if all the eigenvalues are small enough ($|d_i - 1| < 1$), then the series defining the usual log converges and we have (using the identity matrix I_d):

$$\log(\Sigma) = U\left(\mathrm{DIAG}\left(\sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k}(d_i-1)^k\right)\right)U^{\mathrm{T}} = \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k}(\Sigma - \mathrm{Id})^k. \quad (3.1)$$

The logarithm we defined is obviously the inverse function of exp. Thus, the matrix exponential realizes a one-to-one mapping between the space of symmetric matrices to the space of tensors.

Classically, one defines the (left) square root of a matrix B as the set $\{B_L^{1/2}\} = \{A \in GL_n \,/\, A\,A^{\mathrm{T}} = B\}$. One could also define the right square root: $\{B_R^{1/2}\} = \{A \in GL_n \,/\, A^{\mathrm{T}}\,A = B\}$. For tensors, we define the square root as:

$$\Sigma^{1/2} = \{\Lambda \in \mathcal{S}ym_n^+ \,/\, \Lambda^2 = \Sigma\}.$$

The square root is always defined and moreover unique: let $\Sigma = U\,D^2\,U^{\mathrm{T}}$ be a diagonalization (with positives values for the d_i's). Then $\Lambda = U\,D\,U^{\mathrm{T}}$ is of course a square root of Σ, which proves the existence. For the uniqueness, let us consider two symmetric and positive square roots Λ_1 and Λ_2 of Σ. Then, $\Lambda_1^2 = \Sigma$ and $\Lambda_2^2 = \Sigma$ obviously commute and thus they can be diagonalized in the same basis: this means that the diagonal matrices D_1^2 and D_2^2 are equal. As the elements of D_1 and D_2 are positive, they are also equal and $\Lambda_1 = \Lambda_2$. Last but not least, we have the property that

$$\Sigma^{1/2} = \exp\left(\frac{1}{2}(\log\Sigma)\right).$$

3.3.2 An affine invariant distance

Let us consider the following action of the linear group GL_n on the tensor space $\mathcal{S}ym_n^+$:

$$A \star \Sigma = A\Sigma A^{\mathrm{T}} \qquad \forall A \in GL_n \qquad \text{and} \qquad \Sigma \in \mathcal{S}ym_n^+.$$

This group action corresponds for instance to the standard action of the affine group on the covariance matrix Σ_{xx} of a random variables x in $\mathbb{R}^n$: if $y = Ax + t$, then $\bar{y} = A\bar{x} + t$ and $\Sigma_{yy} = \boldsymbol{E}[(y - \bar{y})\,(y - \bar{y})^{\mathrm{T}}] = A\Sigma_{xx}A^{\mathrm{T}}$.

This action is naturally extended to tangent vectors is the same way: if $\Gamma(t) = \Sigma + t\,W + O(t^2)$ is a curve passing at Σ with tangent vector W, then the curve $A \star \Gamma(t) = A\,\Sigma\,A^{\mathrm{T}} + t\,A\,W A^{\mathrm{T}} + O(t^2)$ passes through $A \star \Sigma$ with tangent vector $A \star W$.

Following [Pennec 1998], any invariant distance on $\mathcal{S}ym_n^+$ satisfies $\mathrm{dist}(A\star\Sigma_1, A\star\Sigma_2) = \mathrm{dist}(\Sigma_1, \Sigma_2)$. Choosing $A = \Sigma_1^{-1/2}$, we can reduce this to a pseudo-norm, or distance to the identity:

$$\mathrm{dist}(\Sigma_1, \Sigma_2) = \mathrm{dist}\left(\mathrm{Id}, \Sigma_1^{-\frac{1}{2}}\Sigma_2\Sigma_1^{-\frac{1}{2}}\right) = N\left(\Sigma_1^{-\frac{1}{2}}\Sigma_2\Sigma_1^{-\frac{1}{2}}\right).$$

Moreover, as the invariance has to hold for any transformation, N should be invariant under the action of the isotropy group $\mathcal{H}(\mathrm{Id}) = O_n = \{U \in GL_n \,/\, UU^{\mathrm{T}} = \mathrm{Id}\}$:

$$\forall U \in O_n, \qquad N(U\,\Sigma\,U^{\mathrm{T}}) = N(\Sigma).$$

Using the spectral decomposition $\Sigma = UD^2U^{\mathrm{T}}$, it is easy to see that $N(\Sigma)$ has to be a symmetric function of the eigenvalues. Moreover, the symmetry of the distance $\mathrm{dist}(\Sigma, \mathrm{Id}) = \mathrm{dist}(\mathrm{Id}, \Sigma)$ imposes that $N(\Sigma) = N(\Sigma^{(-1)})$. Thus, a good candidate is the sum of the squared logarithms of the eigenvalues:

$$N(\Sigma)^2 = \|\log(\Sigma)\|^2 = \sum_{i=1}^{n} (\log(\sigma_i))^2. \tag{3.2}$$

This "norm" verifies by construction the properties of symmetry and positiveness. $N(\Sigma) = 0$ implies that $\sigma_i = 1$ (and conversely), so that the separation axiom is satisfied. However, we do no know any simple proof of the triangle inequality, which should read $N(\Sigma_1) + N(\Sigma_2) \geq N(\Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2})$, even if we can verify it experimentally (see e.g. [Förstner 1999]).

3.3.3 An invariant Riemannian metric

Another way to determine the invariant distance is through the Riemannian metric. Let us take the most simple scalar product on the tangent space at the identity matrix: if W_1 and W_2 are tangent vectors (i.e. symmetric matrices, not necessarily definite nor positive), we define the scalar product to be the standard matrix scalar product $\langle\, W_1 \mid W_2 \,\rangle = \mathrm{Tr}(W_1^{\mathrm{T}}\, W_2)$. This scalar product is obviously invariant by the isotropy group O_n. Now, if W_1 and W_2 are two tangent vectors at Σ, we require their scalar product to be invariant by the action of any transformation: $\langle\, W_1 \mid W_2 \,\rangle_\Sigma = \langle\, A \star W_1 \mid A \star W_2 \,\rangle_{A\star\Sigma}$. This should be true in particular for $A = \Sigma^{-1/2}$, which allows us to define the scalar product at any Σ from the scalar product at the tangent space at the identity matrix:

$$\langle\, W_1 \mid W_2 \,\rangle_\Sigma = \left\langle\, \Sigma^{-\frac{1}{2}} W_1 \Sigma^{-\frac{1}{2}} \;\middle|\; \Sigma^{-\frac{1}{2}} W_2 \Sigma^{-\frac{1}{2}} \,\right\rangle_{\mathrm{Id}} = \mathrm{Tr}\left(\Sigma^{-\frac{1}{2}} W_1 \Sigma^{-1} W_2 \Sigma^{-\frac{1}{2}}\right).$$

One can easily verify that this definition is left unchanged if we use any other transformation $A = U\,\Sigma^{-1/2}$ (where U is a free orthonormal matrix) that transports Σ to the identity: $A \star \Sigma = A\,\Sigma\,A^{\mathrm{T}} = U\,U^{\mathrm{T}} = \mathrm{Id}$.

To find the geodesic without going though the computation of Christoffel symbols, we may rely on a result from differential geometry [Gamkrelidze 1991, Helgason 1978, Kobayashi 1969] which says that the geodesics for the invariant metrics on affine symmetric spaces are generated by the action of the one-parameter subgroups of the acting Lie group. Since the one-parameter subgroups of the linear group are given by the matrix exponential $\exp(t\,A)$, geodesics on our tensor manifold going through Σ with tangent vector W should have the following form:

$$\Gamma_{(\Sigma,W)}(t) = \exp(t\,A)\;\Sigma\;\exp(t\,A)^{\mathrm{T}} \qquad \text{with} \qquad W = A\,\Sigma + \Sigma\,A^{\mathrm{T}}. \tag{3.3}$$

For our purpose, we need to relate explicitly the geodesic to the tangent vector in order to define the exponential chart. Since Σ is a symmetric matrix, there is hopefully an explicit solution to the Sylvester equation $W = A\,\Sigma + \Sigma\,A^{\mathrm{T}}$. We get $A = \frac{1}{2}\left(W\,\Sigma^{(-1)} + \Sigma^{1/2}\,Z\,\Sigma^{-1/2}\right)$, where Z is a free skew-symmetric matrix. However, introducing this solution into the equation of geodesics (Eq. 3.3) does not lead to a very tractable expression. Let us look at an alternative solution.

Since our metric (and thus the geodesics) is invariant under the action of the group, we can focus on the geodesics going through the origin (the identity). In that case, a symmetric solution of the Sylvester equation is $A = \frac{1}{2}W$, which gives the following equation for the geodesic going through the identity with tangent vector W:

$$\Gamma_{(\,\mathrm{Id},W)}(t) = \exp\left(\frac{t}{2}\,W\right)\exp\left(\frac{t}{2}\,W\right)^{\mathrm{T}} = \exp(t\,W).$$

We may observe that the tangent vector along this curve is the parallel transportation of the initial tangent vector. If $W = U\,\mathrm{DIAG}(w_i)\,U^{\mathrm{T}}$,

$$\frac{d\Gamma(t)}{dt} = U\,\mathrm{DIAG}\left(w_i \exp(t\,w_i)\right)\,U^{\mathrm{T}} = \Gamma(t)^{\frac{1}{2}}\,W\,\Gamma(t)^{\frac{1}{2}} = \Gamma(t)^{\frac{1}{2}} \star W.$$

By definition of our invariant metric, the norm of this vector is constant: $\|\Gamma(t)^{1/2} \star W\|^2_{\Gamma(t)^{1/2}\star\,\mathrm{Id}} = \|W\|^2_{\mathrm{Id}} = \|W\|_2^2$. This was expected since geodesics are parameterized by arc-length. Thus, the length of the curve between time 0 and 1 is

$$\mathcal{L} = \int_0^1 \left\| \frac{d\Gamma(t)}{dt} \right\|^2_{\Gamma(t)} dt = \|W\|^2_{\mathrm{Id}}.$$

Solving for $\Gamma_{(\,\mathrm{Id},W)}(1) = \Sigma$, we obtain the "norm" $N(\Sigma)$ of Eq.(3.2). Using the invariance of our metric, we easily obtain the geodesic starting from any other point of the manifold using our group action:

$$\Gamma_{(\Sigma,W)}(t) = \Sigma^{\frac{1}{2}} \star \Gamma_{\left(\,\mathrm{Id},\Sigma^{-1/2}\star W\right)}(t) = \Sigma^{\frac{1}{2}} \exp\left(t\,\Sigma^{-\frac{1}{2}} W \Sigma^{-\frac{1}{2}}\right)\Sigma^{\frac{1}{2}}. \tag{3.4}$$

Coming back to the distance $\mathrm{dist}^2(\Sigma,\,\mathrm{Id}) = \sum_i (\log\sigma_i)^2$, it is worth noticing that tensors with null eigenvalues are located as far from the identity as tensors

with infinite eigenvalues: at the infinity. Thanks to the invariance by the linear group, this property holds for the distance to any (positive definite) tensor of the manifold. Thus, the original cone of positive definite symmetric matrices (a linear manifold with a flat metric but which is incomplete: there is a boundary at a finite distance) has been changed into a regular and complete (but curved) manifold with an infinite development in each of its $n(n+1)/2$ directions.

3.3.4 Exponential and logarithm maps

As a general property of Riemannian manifolds, geodesics realize a local diffeomorphism from the tangent space at a given point of the manifold to the manifold itself: $\Gamma_{(\Sigma,W)}(1) = \exp_\Sigma(W)$ associates to each tangent vector $W \in T_\Sigma \mathcal{S}ym_n^+$ a point of the manifold. This mapping is called the exponential map, because it corresponds to the usual exponential in some matrix groups. This is exactly our case for the exponential map around the identity:

$$\exp_{\mathrm{Id}}(UDU^{\mathrm{T}}) = \exp(UDU^{\mathrm{T}}) = U\ \mathrm{DIAG}\left(\exp(d_i)\right)\ U^{\mathrm{T}}.$$

However, the Riemannian exponential map associated to our invariant metric has a more complex expression at other tensors:

$$\exp_\Sigma(W) = \Sigma^{\frac{1}{2}} \exp\left(\Sigma^{-\frac{1}{2}} W \Sigma^{-\frac{1}{2}}\right) \Sigma^{\frac{1}{2}}.$$

In our case, this diffeomorphism is global, and we can uniquely define the inverse mapping everywhere:

$$\log_\Sigma(\Lambda) = \Sigma^{\frac{1}{2}} \log\left(\Sigma^{-\frac{1}{2}} \Lambda \Sigma^{-\frac{1}{2}}\right) \Sigma^{\frac{1}{2}}.$$

Thus, $\exp_\Sigma$ gives us a collection of one-to-one and complete maps of the manifold, centered at any point Σ. As explained in Section 3.2, these charts can be viewed as the development of the manifold onto the tangent space along the geodesics. Moreover, as the manifold has a non-positive curvature [Skovgaard 1984], there is no cut-locus and the statistical properties detailed in [Pennec 2004a] hold in their most general form. For instance, we have the existence and uniqueness of the mean of any distribution with a compact support [Kendall 1990].

For completeness, we give here the affine-invariant distance between two tensors:

$$\begin{aligned} \mathrm{dist}^2(\Sigma_1, \Sigma_2) &= \left\langle \log_{\Sigma_1}(\Sigma_2) \,|\, \log_{\Sigma_1}(\Sigma_2) \right\rangle_{\Sigma_1} \\ &= \mathrm{Tr}\left(\log\left(\Sigma_1^{-1/2} \Sigma_2 \Sigma_1^{-1/2}\right)^2\right) \end{aligned} \tag{3.5}$$

3.3.5 Induced and orthonormal coordinate systems

One has to be careful because the coordinate system of all these charts is not orthonormal. Indeed, the coordinate system of each chart is induced by the standard coordinate system (here the matrix coefficients), so that the vector $\overrightarrow{\Sigma\Lambda}$ corresponds to the standard derivative in the vector space of matrices: we have $\Lambda = \Sigma + \overrightarrow{\Sigma\Lambda} + O(\|\overrightarrow{\Sigma\Lambda}\|^2)$. Even if this basis is orthonormal at some points of the manifold (such as at the identity for our tensors), it has to be corrected for the Riemannian metric at other places due to the manifold curvature.

From the expression of the metric, one can observe that

$$\|\overrightarrow{\Sigma\Lambda}\|_\Sigma^2 = \|\log_\Sigma(\Lambda)\|_\Sigma^2 = \|\Sigma^{-\frac{1}{2}}\log_\Sigma(\Lambda)\Sigma^{-\frac{1}{2}}\|_{\mathrm{Id}}^2 = \|\log(\Sigma^{-\frac{1}{2}} \star \Lambda)\|_2^2.$$

This shows that $\overrightarrow{\Sigma\Lambda}_\perp = \log(\Sigma^{-\frac{1}{2}} \star \Lambda) \in T_{I_d}\mathcal{S}ym_n^+$ is the expression of the vector $\overrightarrow{\Sigma\Lambda}$ in an orthonormal basis. In our case, the transformation $\Sigma^{1/2} \in GL_n$ is moreover uniquely defined (as a positive square root) and is a smooth function of Σ over the complete tensor manifold. Thus, $\overrightarrow{\Sigma\Lambda}_\perp$ results in an atlas of orthonormal exponential charts which is globally smooth with respect to the reference point[1] Σ. This group action approach was chosen in earlier works [Pennec 1996, Pennec 1997, Pennec 1998] with what was called the placement function.

For some statistical operations, we need to use a minimal representation (e.g. 6 parameters for 3×3 tensors) in a (locally) orthonormal basis. This can be realized through the classical "Vec" operator that maps the element $a_{i,j}$ of a $n \times n$ matrix A to the $(i\,n+j)$th element $\mathrm{Vec}(A)_{i\,n+j}$ of a $n \times n$ dimensional vector $\mathrm{Vec}(A)$. Since we are working with symmetric matrices, we have only $n(n+1)/2$ independent coefficients, say the upper triangular part. However, the off-diagonal coefficients are counted twice in the L_2 norm at the identity: $\|W\|_2^2 = \sum_{i=1}^n w_{i,i}^2 + 2\sum_{i<j\leq n} w_{i,j}^2$. Thus, to express our minimal representation in an orthonormal basis, we need to multiply the off diagonal terms by $\sqrt{2}$:

$$\mathrm{Vec}_{\mathrm{Id}}(W) = \left(w_{1,1}, \sqrt{2}\,w_{1,2}, w_{2,2}, \ldots, \sqrt{2}\,w_{(n-1),n}, w_{n,n}\right)^{\mathrm{T}}.$$

Now, for a vector $\overrightarrow{\Sigma\Lambda} \in T_\Sigma\mathcal{S}ym_n^+$, we define its minimal representation in the orthonormal coordinate system as:

$$\mathrm{Vec}_\Sigma(\overrightarrow{\Sigma\Lambda}) = \mathrm{Vec}_{\mathrm{Id}}(\overrightarrow{\Sigma\Lambda}_\perp) = \mathrm{Vec}_{\mathrm{Id}}\left(\log(\Sigma^{-\frac{1}{2}} \star \Lambda)\right). \tag{3.6}$$

The mapping Vec_Σ realizes an explicit isomorphism between $T_\Sigma\mathcal{S}ym_n^+$ and $\mathbb{R}^{n(n+1)/2}$ with the canonical metric.

[1] On most homogeneous manifolds, this can only be realized locally. For instance, on the sphere, there is a singularity at the antipodal point of the chosen origin for any otherwise smooth placement function.

3.3.6 Gradient descent and PDEs: an intrinsic geodesic marching scheme

Let $f(\Sigma)$ be an objective function to minimize, Σ_t the current estimation of Σ, and $W_t = \partial_\Sigma f = [\partial f / \partial \sigma_{ij}]$ its matrix derivative at that point, which is of course symmetric. The principle of a first order gradient descent is to go toward the steepest descent, in the direction opposite to the gradient for a short time-step ε, and iterate the process. However, the standard operator $\Sigma_{t+1} = \Sigma_t - \varepsilon W_t$ is only valid for very short time-steps in the flat Euclidean matrix space, and we could easily go out of the cone of positive definite tensors. A much more interesting numerical operator is given by following the geodesic backward starting at Σ with tangent vector W_t during a time ε. This intrinsic gradient descent ensures that we cannot leave the manifold. It can easily be expressed using the exponential map:

$$\Sigma_{t+1} = \Gamma_{(\Sigma_t, W_t)}(-\varepsilon) = \exp_{\Sigma_t}(-\varepsilon W_t) = \Sigma^{\frac{1}{2}} \exp(-\varepsilon \Sigma^{-\frac{1}{2}} W_t \Sigma^{-\frac{1}{2}}) \Sigma^{\frac{1}{2}}.$$

This intrinsic scheme is trivially generalized to partial differential evolution equations (PDEs) on tensor fields such as $\partial_t \Sigma(x,t) = -W(x,t)$: we obtain $\Sigma(x, t+dt) = \exp_{\Sigma(x,t)}(-dt\, W(x,t))$.

In the next section, we present a second family of metrics, the Log-Euclidean metrics.

3.4 Log-Euclidean Metrics

Log-Euclidean metrics are another type or Riemannian metrics that we have developed. They rely on a very simple fact: the matrix exponential and logarithm are two diffeomorphisms between the tensor space and the space of symmetric matrices. As a consequence, computation on tensors will be drastically simplified while keeping almost the same properties as the affine-invariant family. In the following, we first present the theory of Log-Euclidean metrics, and then show how to use them in practice.

3.4.1 Definition of the Log-Euclidean Metrics

The notion of matrix logarithms and exponentials are central in the theoretical framework presented here. We recall that the *matrix exponential* exp is the generalization to square matrices of the classical scalar exponential. For any square matrix M, it is given by: $\exp(M) = \sum_{k=0}^{\infty} M^k / k!$. Basically, the *matrix logarithm*, denoted "log", is defined as in the scalar case as the inverse of the matrix exponential. But in the general matrix case, this is complicated by the fact that the logarithm of a given invertible matrix can fail to exist [Culver 1966], and that whenever it exists, an infinity of logarithms may also exists. This complex situation is typical of general Lie groups [Bourbaki 1989].

By contrast, the situation is very simple in the tensor case. A tensor Σ has a unique matrix logarithm $L = \log(\Sigma)$. It is uniquely defined by the equation

$\Sigma = \exp(L)$ and L is a simple symmetric matrix. Conversely, each symmetric matrix is associated with a unique tensor by the exponential. L is obtained from Σ by changing its eigenvalues into their natural logarithms, which can be done easily in an orthonormal basis in which Σ (and L) is diagonal. This simple *spectral decomposition* of tensors allows for the generalization of the notion of powers (and in particular square roots) to tensors: for any real number α and any tensor Σ, Σ^α is well-defined. It is obtained by transforming its eigenvalues into their α^{th} power. For example, the square root of a tensor is obtained by replacing its eigenvalues by their square roots, and its inverse (corresponding to the "-1^{th}" power) by replacing its eigenvalues by their inverses.

Based on the specific properties of the matrix exponential on tensors, we can now define a novel structure of *vector space* on tensors. Since there is a one-to-one mapping between the tensor space and the *vector* space of symmetric matrices, one can transfer to tensors the addition "+" and the scalar multiplication "." with the matrix exponential. This defines on tensors the *logarithmic product* $\odot$ and the *logarithmic scalar multiplication* $\circledast$, given by:

$$\begin{cases} \Sigma_1 \odot \Sigma_2 \stackrel{def}{=} \exp\left(\log(\Sigma_1) + \log(\Sigma_2)\right) \\ \lambda \circledast \Sigma \stackrel{def}{=} \exp\left(\lambda . \log(\Sigma)\right) = \Sigma^\lambda. \end{cases} \tag{3.7}$$

The logarithmic product is *commutative* and coincides with matrix multiplication whenever the two tensors Σ_1 and Σ_2 commute in the matrix sense. With $\odot$ and $\circledast$, the tensor space has by construction a *vector space structure*, which is *not* the usual structure directly derived from the addition and scalar multiplication on square matrices.

When one considers only the multiplication $\odot$ on the tensor space, one has a *Lie group* structure [Gallot 1993], i.e. a space which is both a smooth manifold and a group in which algebraic operations are *smooth mappings*. This comes from the fact that both the exponential and the logarithm mappings are smooth, as shown in [Arsigny 2005b]. Among Riemannian metrics in Lie groups, the most convenient in practice are *biinvariant metrics*, i.e. distances that are invariant by multiplication and inversion. When they exist, these metrics are used in differential geometry to generalize properly the notion of mean to Lie groups as in the case of rotations [Pennec 1996, Moakher 2002]. However, such metrics are not always available, as in the case of the groups of Euclidean motions [Pennec 2004a] and affine transformations. It is remarkable that biinvariant metrics exist in the tensor Lie group. Moreover, they are particularly simple. This existence results from the *commutativity* of the logarithmic product between tensors. We have named such metrics *Log-Euclidean* metrics, since they correspond to Euclidean metrics in the domain of logarithms. From a Euclidean norm $\|.\|$ on symmetric matrices, they can be written:

$$dist(\Sigma_1, \Sigma_2) = \|\log(\Sigma_1) - \log(\Sigma_2)\|. \tag{3.8}$$

As one can see, the Log-Euclidean distance is much simpler than the equivalent affine-invariant distance given by Eq. (3.5), where the two tensors are mixed with

square roots and inverses before taking norm of the logarithm. Log-Euclidean metrics take this form because they are simply Euclidean distances for the vector space structure we have defined earlier. We did not define them directly from the latter algebraic structure to emphasize the fact that they are Riemannian metrics, like affine-invariant ones.

The greater simplicity of Log-Euclidean metrics can also be seen from Log-Euclidean geodesics in the tensor space. The shortest path between two tensors Σ_1 and Σ_2 parametrized proportionally to its arc length by a parameter t between 0 and 1 is a straight line in the domain of logarithms in the Log-Euclidean case. It is given by:

$$\exp\left((1-t)\log(\Sigma_1) + (t)\log(\Sigma_2)\right). \tag{3.9}$$

The equivalent affine-invariant geodesic involves the use of square roots and inverses (Eq. 3.4). Contrary to the classical Euclidean framework on tensors, and similarly to the affine-invariant framework, one can see from Eq. [3.8] that matrices with null or negative eigenvalues are at an infinite distance from tensors and will not appear in practical computations. Indeed, among symmetric matrices, only tensors are at a finite distance from any tensor, whereas the distance between the null matrix (which is not a tensor) and the identity (which is a tensor) is finite in the Euclidean case.

3.4.2 Invariance Properties of the Log-Euclidean Metrics

A metric on tensors will be all the more relevant as it exhibits as many natural invariance properties as possible. First, distances are not changed by inversion, since taking the inverse of a system of matrices only results in the multiplication by -1 of their logarithms, which does not change the value of the distance given by Eq. [3.8].

Also, Log-Euclidean metrics are by construction invariant with respect to any logarithmic multiplication, i.e. are invariant by any translation in the domain of logarithms. But there is more. Although Log-Euclidean metrics will not yield full affine-invariance as the affine-invariant metrics defined previously, a number of them are invariant by similarity (orthogonal transformation and scaling). This means that computations on tensors using these metrics will be invariant with respect to a change of coordinates obtained by a similarity. The similarity-invariant Log-Euclidean metric used throughout this thesis is given by:

$$\mathrm{dist}(\Sigma_1, \Sigma_2) = \left(\mathrm{Trace}\left(\{\log(\Sigma_1) - \log(\Sigma_2)\}^2\right)\right)^{\frac{1}{2}}. \tag{3.10}$$

The invariance by similarity of this metric can be simply proved by using the following elementary algebraic property for two symmetric matrices: $\mathrm{Trace}(A.B) = \mathrm{Trace}(B.A)$ and the fact that a scaling in the domain of tensors corresponds to a translation by a multiple of the identity in the domain of logarithms.

3.4.3 Log-Euclidean Computations on Tensors

From a practical point of view, one would ideally like operations such as averaging, filtering, etc. to be as simple as possible. In the affine-invariant case, such operations rely on an intensive use of matrix exponentials, logarithms, inverses and square roots. But here, the vector space of tensors with a Log-Euclidean metric is in fact isomorphic (the algebraic structure of vector space is conserved) and *isometric* (distances are conserved) with the corresponding Euclidean space of symmetric matrices. As a consequence, the Riemannian framework for statistics and analysis is extremely simplified [Arsigny 2005b].

In general and contrary to the affine-invariant case, the processing of tensors in the Log-Euclidean framework is simply Euclidean in the logarithmic domain. Tensors are first converted into symmetric matrices (i.e. vectors) using the matrix logarithm. Then vector processing tools are directly used on logarithms. Finally, results obtained on logarithms are mapped back to the tensor domain with the exponential. Hence, statistical tools or PDEs are very readily generalized to tensors in this framework.

3.4.4 Link with the Affine-Invariant Metric

The Log-Euclidean metric is strongly linked to the affine-invariant one. Indeed, it can be seen as the special case where all tangent vectors are expressed in the tangent space at the identity matrix: $\log_{I_d}(\Sigma) = \log(\Sigma)$. In Table 3.2, we express the basic operations on Riemannian manifolds with the affine-invariant and Log-Euclidean metrics.

Operation	Affine-Invariant metric	Log-Euclidean metric
$\overrightarrow{\Sigma_1\Sigma_2} = \log_{\Sigma_1}(\Sigma_2)$	$\Sigma_1^{1/2}\log(\Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2})\Sigma_1^{1/2}$	$\log(\Sigma_2) - \log(\Sigma_1)$
$\Sigma_2 = \exp_{\Sigma_1}(\overrightarrow{\Sigma_1\Sigma_2})$	$\Sigma_2 = \Sigma_1^{1/2}\exp(\overrightarrow{\Sigma_1\Sigma_2})\Sigma_1^{1/2}$	$\Sigma_2 = \exp(\log(\Sigma_1) + \overrightarrow{\Sigma_1\Sigma_2})$
$\mathrm{dist}^2(\Sigma_1, \Sigma_2)$	$\mathrm{Tr}(\log(\Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2})^2)$	$\mathrm{Tr}((\log(\Sigma_2) - \log(\Sigma_1))^2)$

Table 3.2: Basic Riemannian operations in the affine-invariant and Log-Euclidean frameworks.

We have developed throughout this chapter two alternatives to the flawed Euclidean calculus. Both of them use tools from Riemannian geometry, and reject any non-positive matrix at an infinite distance of any tensor. Schematically, they "unfold" the tensor space to turn it into a vector space, where operations are simple, and then "fold" it again to "project" the result back onto the tensor space. Moreover, they exhibit interesting invariance properties: invariance by any affine transform (case of the affine-invariant Riemannian metric), or by a similitude transform (case of the Log-Euclidean metric). However, their practical implementation is different, though they both use the matrix logarithm and exponential that appear as the key functions to act on the tensor space. One may wonder whether one framework is more valuable practically than another, in terms of quality of the result, compu-

tational efficiency, and ease of implementation. In the next chapter, we illustrate the practical use of these two Riemannian frameworks with geometric and image processing operations on tensors.

Chapter 4

Riemannian Processing of Tensor-Valued Images: Application to DT-MRI

Contents

We have presented the theory of two Riemannian frameworks for tensor processing, which overcomes the limitations of the standard Euclidean calculus. In this chapter, we demonstrate the use of these metrics in practice. In particular, we show how to apply some statistical tools to tensors (mean, covariance), to perform multi-linear interpolation, and to smooth or restore tensor images. To keep a close eye on the potential of these tools on real data, some examples will be done on tensor fields obtained from diffusion weighted MRI, thus illustrating their practical use on this type of data. For each processing, we will detail the affine-invariant and the Log-Euclidean implementations, and confront them.

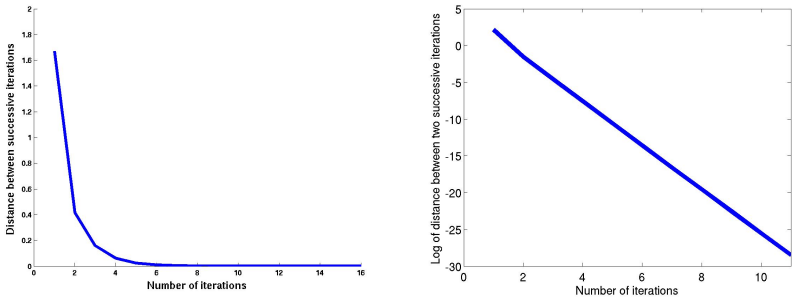


Figure 4.1: Convergence of the affine-invariant mean calculation of 1000 random Gaussian tensors. **Left:** The graph shows the evolution of the distance between successive iterations (i.e., estimations of the mean). The convergence is clearly very fast. **Right:** Log of the previous graph. The slope was evaluated at -3, which shows that the convergence is cubic.

4.1 Statistical Analysis of Tensors

4.1.1 The Fréchet Mean

Let $\Sigma_1 \dots \Sigma_N$ be a set of measures of the same tensor. Let $\omega_1 \dots \omega_N$ be a set of positive weights such that $\sum_{i=1}^{N} \omega_i = 1$. The Fréchet mean is the set of tensors (it does not have to be unique in general) minimizing the sum of squared distances: $C(\Sigma) = \sum_{i=1}^{N} \omega_i \, \mathrm{dist}^2(\Sigma, \Sigma_i)$.

4.1.1.1 The Affine-Invariant Case:

Endowed with the affine-invariant metric, the tensor manifold has a non-positive curvature [Skovgaard 1984], so that there is one and only one mean value $\bar{\Sigma}$ [Kendall 1990]. Moreover, a necessary and sufficient condition for an optimum is a null gradient of the criterion. Thus, the intrinsic Newton gradient descent algorithm gives the following mean value at estimation step $t+1$:

$$\bar{\Sigma}_{t+1} = \exp_{\bar{\Sigma}_t}\left(\sum_{i=1}^{N} \omega_i \log_{\bar{\Sigma}_t}(\Sigma_i)\right) = \bar{\Sigma}_t^{\frac{1}{2}} \exp\left(\sum_{i=1}^{N} \omega_i \log\left(\bar{\Sigma}_t^{-\frac{1}{2}} \Sigma_i \bar{\Sigma}_t^{-\frac{1}{2}}\right)\right) \bar{\Sigma}_t^{\frac{1}{2}}. \quad (4.1)$$

Note that we cannot easily simplify this expression further as in general the data Σ_i and the mean value $\bar{\Sigma}_t$ cannot be diagonalized in a common basis. However, this gradient descent algorithm usually converges very fast (about 10 iterations, see Fig. 4.1).

4.1.1.2 The Log-Euclidean Case:

The Log-Euclidean Fréchet mean is straightforward to calculate. It is a direct generalization of the geometric mean of positive numbers and is given explicitly by:

$$\bar{\Sigma} = \exp\left(\sum_{i=1}^{N} w_i \log(\Sigma_i)\right). \tag{4.2}$$

The closed form given by Eq. [4.2] makes the computation of Log-Euclidean means immediate. To illustrate the differences between the Euclidean, affine-invariant and Log-Euclidean frameworks for the mean tensor computation, we did the following experiment. First, we created 1000 random Gaussian tensors using the following procedure: we generated $n(n+1)/2$ independent and normalized real Gaussian samples, multiplied the corresponding vector by the square root of the desired covariance matrix (expressed in our Vec coordinate system), and come back to the tensor manifold using the inverse Vec mapping. Using this procedure, we can easily generate noisy measurements of known tensors (see Fig. 4.2). Second, we computed the Euclidean, affine-invariant and Log-Euclidean means. We compared the computational time, and the Frobenius distance between the estimated mean and the true mean which tells us how close the estimations are from the true mean. Finally, we iterated this process for several variances (from 0.1 to 1.0) and plot a graph of the computational time w.r.t. variance, and accuracy of the mean estimation w.r.t. the variance (Fig. 4.3). The reason why we increase the variance is because the affine-invariant computation is an iterative process, and the convergence is decided when the Frobenius distance between two consecutive iterations is less than 1.10^{-6}. The increase of the variance spreads the samples far from the mean value, which forces the affine-invariant computation to perform more iterations to converge to the solution. Another remark is that the Euclidean estimated mean gets worse when the variance increases. This is mainly due to the swelling effect: average tensors tend to get bigger than the original samples. We will discuss this effect more in details in Sec. 4.2.1. Notice that the type of noise added is well adapted to our Riemannian metrics (Gaussian on a tangent space), and not to the Euclidean framework, which explains also why the results with the Euclidean metric are so bad. However, one should keep in mind that adding a Gaussian noise directly on the mean tensor coefficients (which would be the Euclidean way of creating noisy tensors) does not necessarily produce tensors, i.e., positive definite symmetric matrices. Indeed, when the noise variance is too high w.r.t. the coefficients, negative eigenvalues appear and one cannot correctly generate noisy tensors.

4.1.2 Covariance Matrix and Mahalanobis Distance

4.1.2.1 The Affine-Invariant Case:

As described in [Pennec 2004a], we may generalize most of the usual statistical methods by using the exponential chart at the mean point. For instance, the empirical covariance matrix of a set of N tensors Σ_i of mean $\bar{\Sigma}$ is defined using the tensor product: $\frac{1}{N-1}\sum_{i=1}^{n} \overrightarrow{\bar{\Sigma}\Sigma_i} \otimes \overrightarrow{\bar{\Sigma}\Sigma_i}$. Using our *Vec* mapping, we may come back to more usual matrix notations and write its expression in a minimal representation with an

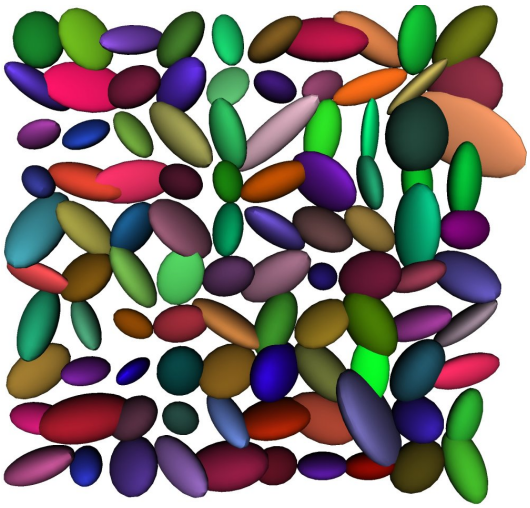


Figure 4.2: **A 10×10 field of Gaussian random tensors.** Mean: I_d, variance: 0.1.

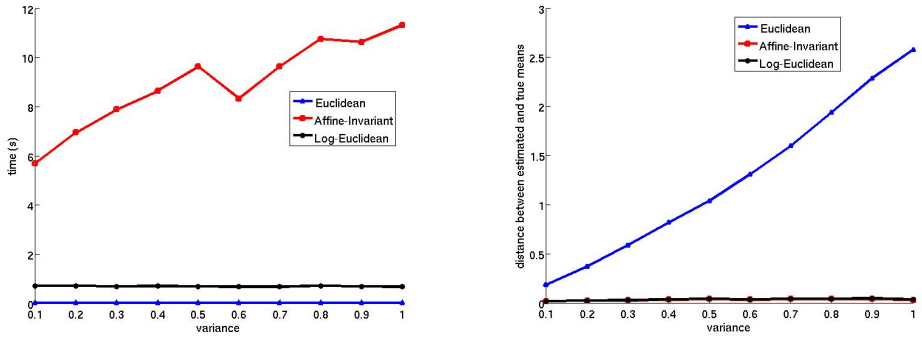


Figure 4.3: **Mean tensor calculation: Euclidean vs. Affine-Invariant vs. Log-Euclidean.** 1000 random tensors were used. Implementation was made with Matlab on a Pentium M at 2GHz with 1Gb of RAM. **Left:** Time versus variance of random tensors. While Euclidean (blue line) and Log-Euclidean (black line) computational times are constant w.r.t. the variance, the affine-invariant computational time (red line) increases with the variance. Indeed, more iterations are necessary when the spread of the samples around the mean is greater. **Right:** Euclidean distance between estimated and true means. Affine-Invariant and Log-Euclidean estimations have a constant accuracy w.r.t. the variance. The Euclidean estimation, however, gets worse when the variance increases, mainly because of the swelling effect.

orthonormal coordinate system:

$$\Gamma = \frac{1}{N-1} \sum_{i=1}^{N} \mathrm{Vec}_{\bar{\Sigma}} \left(\overrightarrow{\bar{\Sigma}\Sigma_i} \right) \; \mathrm{Vec}_{\bar{\Sigma}} \left(\overrightarrow{\bar{\Sigma}\Sigma_i} \right)^{\mathrm{T}}.$$

One may also define the Mahalanobis distance:

$$\mu^2_{(\bar{\Sigma},\Gamma)}(\Sigma) = \mathrm{Vec}_{\bar{\Sigma}} \left(\overrightarrow{\bar{\Sigma}\Sigma} \right)^{\mathrm{T}} \Gamma^{(-1)} \; \mathrm{Vec}_{\bar{\Sigma}} \left(\overrightarrow{\bar{\Sigma}\Sigma} \right).$$

Looking for the probability density function that minimizes the information with a constrained mean and covariance, we obtain a generalization of the Gaussian distribution of the form:

$$N_{\bar{\Sigma},\Gamma}(\Sigma) = k \, \exp\left(-\frac{1}{2}\mu^2_{\bar{\Sigma},\Gamma}(\Sigma)\right).$$

The main difference with a Euclidean space is that we have a curvature to take into account: the invariant measure induced on the manifold by our metric is linked to the usual matrix measure by $d\mathcal{M}(\Sigma) = d\Sigma/\det(\Sigma)$. Likewise, the curvature slightly modifies the usual relation between the covariance matrix, the concentration matrix Γ and the normalization parameter k of the Gaussian distribution [Pennec 2004a]. These differences have an impact on the calculations using continuous probability density functions. However, from a practical point of view, the curvature correction term is here very small [Lenglet 2006a] and could be neglected.

To check the implementation of our charts and geodesic marching algorithms, we verified experimentally the central limit theorem. This theorem states that the empirical mean of N independently and identically distributed (IID) random variables with a variance γ^2 asymptotically follows a Gaussian law of variance γ^2/N, centered at the exact mean value. The principle of our experiments is now as follows. We randomly generated N Gaussian tensors around a random average tensor $\bar{\Sigma}$ with a variance of $\gamma^2 = 1$, using the same strategy as in Sec. 4.1.1.2. We computed the mean $\hat{\Sigma}$ using the algorithm of Eq. 4.1. Now, if the error between the computed and the exact mean really follows a Gaussian law of variance γ^2/N, then the normalized Mahalanobis distance $\mu^2 = N \, \mathrm{dist}(\bar{\Sigma}, \hat{\Sigma})^2/\gamma^2$ should follow a χ^2_6 distribution. However, this simple experiment only gives us one measurement. Thus, to verify the distribution, we repeated this experiment with N varying from 10 to 1000. Figure 4.4 presents the histogram of the normalized Mahalanobis distances we obtain. The empirical distribution follows quite well the theoretical χ^2_6 distribution, as expected, with an empirical mean of 5.85 and a variance of 12.17 (expected values are 6 and 12). Moreover, a Kolmogorov-Smirnov test confirms that the distance between the empirical and theoretical cumulative pdf is not significant (p-value of 0.19).

4.1.2.2 The Log-Euclidean Case:

Similarly to the Fréchet mean, the covariance matrix Γ of a set of N tensors Σ_i with mean $\bar{\Sigma}$ in the Log-Euclidean case is simply expressed in terms of matrix logarithms:

$$\Gamma = \frac{1}{N-1}\sum_{i=1}^{N} \mathrm{Vec}_{I_d}\left(\log(\Sigma_i) - \log(\bar{\Sigma})\right) \, \mathrm{Vec}_{I_d}\left(\log(\Sigma_i) - \log(\bar{\Sigma})\right)^\top$$

The Mahalanobis distance is given by:

$$\mu^2_{(\bar{\Sigma},\Gamma)}(\Sigma) = \mathrm{Vec}_{I_d}\left(\log(\Sigma) - \log(\bar{\Sigma})\right) \Gamma^{-1} \, \mathrm{Vec}_{I_d}\left(\log(\Sigma) - \log(\bar{\Sigma})\right)^\top$$

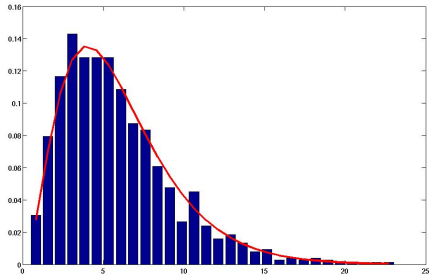


Figure 4.4: Histogram of the renormalized Mahalanobis distance $\mu^2 = N\ \mathrm{dist}(\bar{\Sigma}, \hat{\Sigma})^2/\gamma^2$ between the computed and the exact mean tensors. The red curve is the pdf of the χ^2_6 distribution.

As in the affine-invariant case, the process to generate random Gaussian tensors is the following: we generate a set of $n(n+1)/2$ independent real Gaussian samples whose mean is the logarithm of the desired value and whose variance is 1.0. Then, we multiply the vector (as obtained with Vec) by the square root of the desired covariance matrix, and come back to the tensor space by using the matrix exponential. We verified experimentally that the normalized Mahalanobis distance between the expected mean and the estimation has a χ^2_6 distribution (mean: 6.01, variance: 11.98).

4.2 Tensor Interpolation

One of the important operations in geometric data processing is to interpolate values between known measurements. In 3D image processing, (tri-) linear interpolation is often used thanks to its very low computational load and comparatively much better results than nearest neighbor interpolation. Other popular methods include the cubic and, more generally, spline interpolations [Thévenaz 2000, Meijering 2002].

The standard way to define an interpolation on a regular lattice of dimension d is to consider that the interpolated function $f(x)$ is a linear combination of samples f_k at integer (lattice) coordinates $k \in \mathbb{Z}^d$: $f(x) = \sum_k \omega(x-k)\, f_k$. To realize an interpolation, the "sample weight" function ω has to vanish at all integer coordinates except 0 where it has to be one. A typical example where the convolution kernel has an infinite support is the sinus cardinal interpolation. With the nearest-neighbor, linear (or tri-linear in 3D), and higher order spline interpolations, the kernel is piecewise polynomial, and limited to a few neighboring points in the lattice.

When it comes to an irregular sampling (i.e. a set of measurements f_k at positions x_k), interpolation may still be defined using a weighted mean: $f(x) = \sum_{k=1}^{N} \omega_k(x)\, f_k$. To ensure that this is an interpolating function, one has to require that $\omega_i(x_j) = \delta_{ij}$ (where δ_{ij} is the Kronecker symbol). Moreover, the coordinates are usually normalized so that $\sum_{k=1}^{N} \omega_k(x) = 1$ for all position x within the domain of interest. Typical examples in triangulations or tetrahedrizations are barycentric

and natural neighbor coordinates [Sibson 1981].

4.2.1 Example of the linear interpolation

The linear interpolation is simple as this is a walk along the geodesic joining the two tensors. For instance, the interpolation in the standard Euclidean matrix space would give, for $t \in [0,1]$:

$$\Sigma(t) = (1-t)\,\Sigma_1 + t\,\Sigma_2.$$

This operation is convex (i.e., the weights are positive and the sum is equal to one), thus the result is still a tensor. With the Riemannian metrics, we have the closed-form expression:

$$\Sigma(t) = \exp_{\Sigma_1}\left(t\ \log_{\Sigma_1}(\Sigma_2)\right) = \exp_{\Sigma_2}\left((1-t)\ \log_{\Sigma_2}(\Sigma_1)\right), t \in [0,1],$$

which gives with the affine-invariant metric:

$$\Sigma(t) = \Sigma_1^{1/2} \exp\left(t \log\left(\Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2}\right)\right)\Sigma_1^{1/2}.$$

In the Log-Euclidean framework, it is even simpler:

$$\Sigma(t) = \exp((1-t)\log(\Sigma_1) + t\log(\Sigma_2)).$$

We displayed in Fig. 4.5 the Euclidean, affine-invariant and Log-Euclidean interpolations between 2D tensors of eigenvalues (5,1) horizontally and (1,50) at 45 degrees, along with the evolution of their trace, product (i.e., determinant of the matrix or area of the ellipses) and fractional anisotropy (i.e., distance from the circle).

With the standard matrix coefficient interpolation, the evolution of the trace is perfectly linear (which was expected since this is a linear function of the coefficients). What is much more annoying is that the determinant (i.e., area of the ellipses in 2D, or volume of the ellipsoids in 3D) does not grow regularly in between the two tensors, but goes through a maximum. If we interpret tensors as covariance matrices of Gaussian distributions, this means that the probability of a random point to be accepted as a realization of the distributions is larger in between than at the measurement points themselves! We call this the *swelling effect* (first described by Tschumperlé in [Tschumperlé 2002]), and we will see in the following that it is encountered quite often in tensor processing. On the contrary, one can clearly see a regular evolution of the eigenvalues and their product with the affine-invariant and Log-Euclidean interpolations. Moreover, there is a much smoother rotation of the eigenvectors than with the Euclidean interpolation.

Interestingly, there is almost no difference between the affine-invariant and Log-Euclidean interpolations: both behave quite similarly (smooth interpolation of the eigenvectors, monotonic evolution of the determinant), with slightly more anisotropy in the LE case. In fact, one can show that they are exactly the same when the two tensors commute. Let us assume that Σ_1 and Σ_2 commute. Then, they can be decomposed in the same basis, i.e., there exists an orthogonal matrix U such that

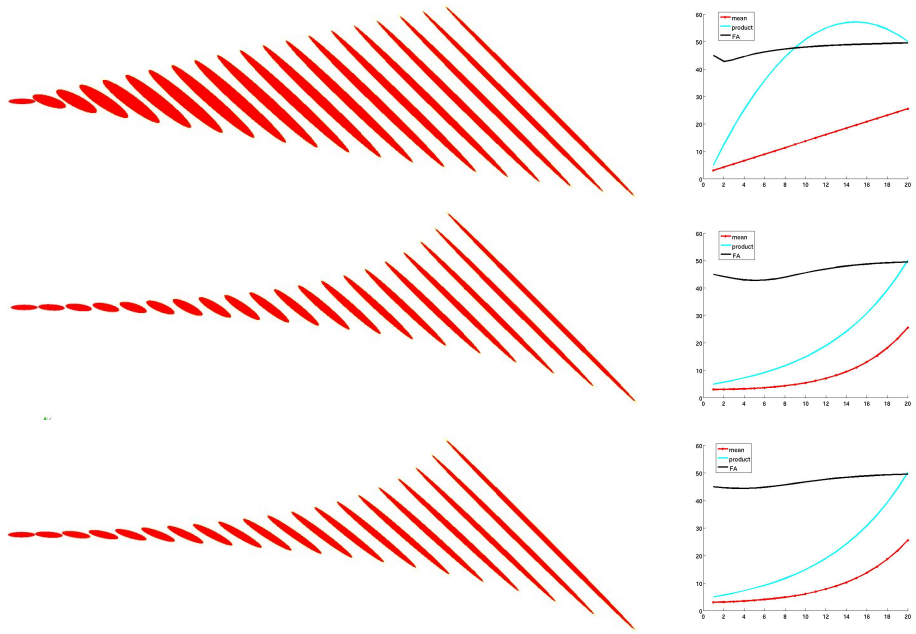


Figure 4.5: **Left:** Linear interpolation between 2D tensors of eigenvalues (5,1) horizontally and (1,50) at 45 degrees. **Top:** interpolation in the standard matrix space (interpolation of the coefficients). **Middle:** geodesic interpolation with the affine-invariant metric. **Bottom:** geodesic interpolation with a Log-Euclidean metric. **Right:** Evolution of the trace (red), product (blue) and fractional anisotropy (black).

$\Sigma_1 = UV_1U^\top$ and $\Sigma_2 = UV_2U^\top$. As V_1 and V_2 are diagonal matrices, they commute and then any power of Σ_1 commute with any power of Σ_2. One can write, in the affine-invariant case:

$$\begin{aligned}
\Sigma(t) &= \exp_{\Sigma_1}\left(t \log_{\Sigma_1}(\Sigma_2)\right) \\
&= \Sigma_1^{1/2} \exp\left(t \log\left(\Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2}\right)\right)\Sigma_1^{1/2} \\
&= \Sigma_1^{1/2} \exp\left(t \log\left(\Sigma_2\Sigma_1^{-1}\right)\right)\Sigma_1^{1/2} \\
&= \Sigma_1^{1/2} \exp\left(t \log(\Sigma_2) - t \log(\Sigma_1)\right)\Sigma_1^{1/2} \\
&= \Sigma_1^{1/2}\Sigma_1^{-t}\Sigma_2^{t}\Sigma_1^{1/2} \\
&= \Sigma_1\Sigma_1^{-t}\Sigma_2^{t} \\
\Sigma(t) &= \Sigma_1^{1-t}\Sigma_2^{t}
\end{aligned}$$

In the Log-Euclidean case, on has:

$$\begin{aligned}
\Sigma(t) &= \exp\left((1-t)\log(\Sigma_1) + t\log(\Sigma_2)\right) \\
\Sigma(t) &= \Sigma_1^{1-t}\Sigma_2^{t}
\end{aligned}$$

Both are the geometric interpolation of the two tensors.

4.2.2 Multi-linear interpolation

The bi- and tri-linear interpolation of tensors on a regular grid in 2D or 3D are almost as simple, except that we do not have any longer an explicit solution in the affine-invariant case since there are more than two points. After computing the (bi-) tri-linear weights with respect to the neighboring sites of the point we want to evaluate, we now have to go through the iterative optimization of the weighted mean (Eq. 4.1) in the affine-invariant case to compute the interpolated tensor. The Euclidean and Log-Euclidean cases have both a closed form solution for the weighted mean calculation (see e.g. Eq. 4.2 for the Log-Euclidean mean).

We display an example of bi-linear interpolation in Figure 4.6. One can see that the surface of the 2D tensors (or volume in 3D) is much more important with the Euclidean than with the Riemannian interpolations. This fact is exemplified with the interpolation of the determinant which is not monotonic in the Euclidean case (see Fig. 4.6 third row), while it is with the Affine-Invariant and Log-Euclidean metrics. Like for the geodesic interpolation, the fractional anisotropy goes trough a minimum value for the three interpolations, but the effect is less pronounced with the Riemannian interpolations (Fig. 4.6 last row). We also get a much smoother interpolation of the principal directions with these Riemannian methods. Once again, both Affine-Invariant and Log-Euclidean frameworks give very close results, with a slightly higher FA for the last one.

4.2.3 Interpolation of non regular measurements

When tensors are not measured on a regular grid but "randomly" localized in space, defining neighbors becomes an issue. One solution, proposed by [Sibson 1981] and later used for surfaces by [Cazals 2001], is the natural neighbor interpolation. For any point x, its natural neighbors are the points of $\{x_i\}$ whose Voronoi cells are chopped off upon insertion of x into the Voronoi diagram. The weight w_i of each natural neighbor x_i is the proportion of the new cell that is taken away by x to x_i in the new Voronoï diagram. One important restriction of these interesting coordinates is that they are limited to the convex hull of the point set (otherwise the volume or surface of the cell is infinite).

Another idea is to rely on radial-basis functions to define the relative influence of each measurement point. A radial basis function (RBF) is a a real-valued function whose value depends only on the distance from the origin, so that $\phi(\mathbf{x}) = \phi(||\mathbf{x}||)$; or alternatively on the distance from some other point c, called a center, so that $\phi(\mathbf{x}, \mathbf{c}) = \phi(||\mathbf{x}-\mathbf{c}||)$. Interpolating a set of discrete measures y_i defined at positions $x_i, i \in [1, N]$ via a RBF ϕ consists in finding the coefficients λ_i that define a function $y(x)$ such that:

$$y(x) = \sum_{j=1}^{N} \lambda_j \phi(x - x_j), \qquad \text{with} \qquad y(x_i) = y_i.$$

The weights λ_i are determined by the linear system: $\forall i \in [1, N], y(x_i) = y_i$. In

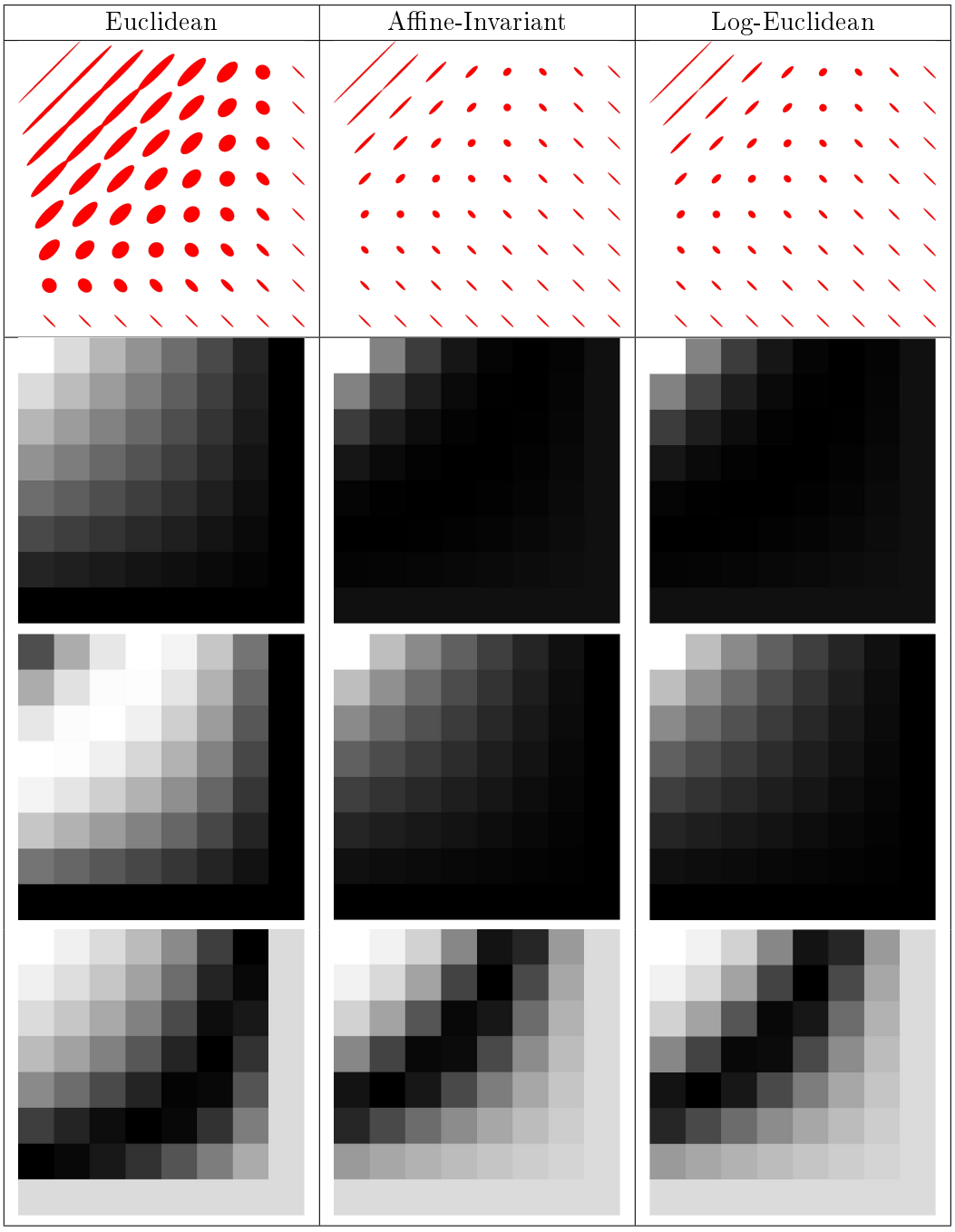


Figure 4.6: Bi-linear interpolation between the four 2D tensors at the corners, using the standard Euclidean calculus (left column), the affine-invariant metric (middle column), and the Log-Euclidean framework (right column). To better illustrate the behavior these different interpolations, we computed the corresponding images of the trace (second row), of the surface (third row), and of the FA (last row).

practice, one has to solve the following system:

$$H.\begin{bmatrix}\lambda_1\\ \vdots \\ \lambda_N\end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N\end{bmatrix}, \qquad \text{with} \qquad [H]_{(i,j)} = \phi(x_i - x_j).$$

$[H]_{(i,j)}$ corresponds to the element of H at the i^{th} row and j^{th} column. A solution to this system is ensured to exist if the matrix H is positive definite. This is the case, for instance, when using RBF like:

- The Gaussian: $\phi(x) = \exp(-\alpha\|x\|^2)$;
- Functions of the family: $\phi(x) = \frac{1}{1+(\|x\|^2/\alpha^2)^\gamma}$;
- The Thin Plate Spline: $\phi(x) = \|x\|^2 \log(\|x\|)$.

So far, this interpolation is only valid for scalar measures. For general vector spaces, and in particular matrix (and tensors)-valued functions, one needs to: 1. convert matrices into vectors using the Vec mapping (Eq. 3.6), 2. interpolate independently each vector coordinate using the RBF, and 3. map the result back to the tensor space using the inverse Vec mapping. Note that it is necessary to solve one system (i.e., compute a set of λ_i) per coefficient. For instance, interpolating 3D tensors with RBF requires solving 6 linear systems (tensors are turned into 6-dimensional vectors). However, the systems can be solved once for all as they are the same for a given set of measures.

4.2.3.1 The Affine-Invariant Case:

This strategy works in Euclidean calculus. In the affine-invariant framework, one needs to map all tensors to a common tangent plane so that tensors become elements of the same vector space in order to be able to combine them. It implies a choice on the point of reference. A reasonable choice is the (affine-invariant) mean of all measures, which is by nature close to each measure. Let $(\Sigma_i, x_i), i \in [1, N]$ be a set of N tensor measures at positions x_i. The first step is to compute the reference tensor $\bar{\Sigma}$, and turn all measures into vectors in the tangent space defined at this reference:

$$\forall i \in [1, N], W_i = \log_{\bar{\Sigma}}(\Sigma_i).$$

Second, these tangent vectors are turned into vectors using the Vec mapping (Eq. 3.6):

$$V_i = Vec_{\bar{\Sigma}}(W_i) = \begin{bmatrix} v_i^0 \\ \vdots \\ v_i^p \end{bmatrix}$$

We denote by v_i^p the p^{th} component of vector V_i. Third, for each component p, one compute the weights λ_j^p, such that:

$$\forall i \in [1, N], v_i^p = \sum_{j=1}^{N} \lambda_j^p \phi(x_i - x_j).$$

Finally, we obtain the interpolated tensor-valued function by mapping the interpolated value back to the tensor space using the inverse Vec function and the

exponential map:

$$\Sigma(x) = \exp_{\bar{\Sigma}}\left(Vec_{\bar{\Sigma}}^{-1}\left(\begin{bmatrix} v^1(x) \\ \vdots \\ v^p(x) \end{bmatrix}\right)\right), \text{ where } v^p(x) = \sum_{j=1}^{N} \lambda_j^p \phi(x - x_j) \forall p.$$

4.2.3.2 The Log-Euclidean Case:

In the Log-Euclidean framework, one simply needs to take the logarithm of each measure, turn them into vectors using the Vec mapping, and perform RBF interpolation on these vectors by computing the weights per vector component exactly as in the affine-invariant case. To map the result back to the tensor space, one needs to turn the interpolated vectors back into matrices using the inverse Vec mapping and exponentiate them as usual. Notice that doing so implicitly uses the identity matrix as tensor of reference. Indeed, one can write:

$$W_i = \log(\Sigma_i) = \log(\Sigma_i) - \log(I_d)$$

In fact, this is similar to the affine-invariant case when the identity tensor is chosen as reference. One can choose another value of reference, like the mean of all measures:

$$W_i = \log(\Sigma_i) - \log(\bar{\Sigma})$$

One should not forget to add the reference value when exponentiating the result:

$$\Sigma(x) = \exp\left(\log(\bar{\Sigma}) + W(x)\right)$$

Results of an RBF interpolation with the thin plate spline function is given in Fig. 4.7. As expected, the Euclidean interpolation tends to make tensor swell, while affine-invariant and Log-Euclidean interpolations preserve the volumes. Affine-invariant and Log-Euclidean results are so closed that they cannot be visually differentiated (we display only one image).

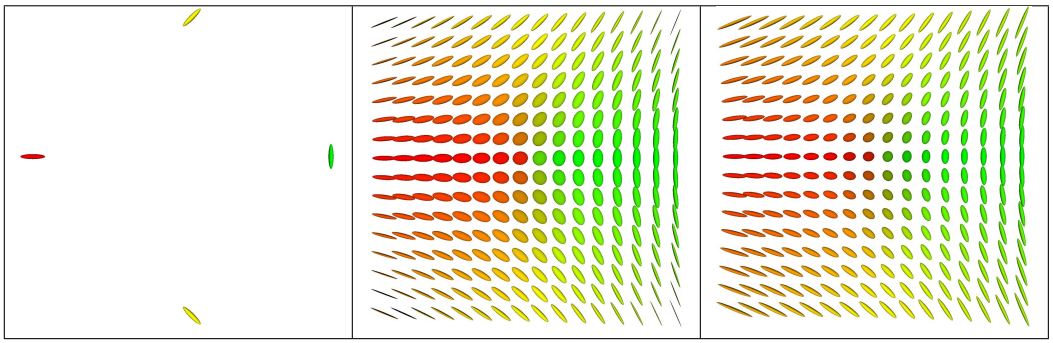


Figure 4.7: RBF tensor interpolation using the Thin Plate Spline function. **Left:** The four initial measures. **Middle:** The Euclidean interpolation. **Right:** The Affine-Invariant and Log-Euclidean interpolations (results are too close to be differentiated).

We will describe in Section 4.4.2 a last alternative that performs the interpolation and extrapolation of sparsely distributed tensor measurements using diffusion.

4.3 Filtering of Tensor Fields

Let us now consider that we have a tensor field, for instance like in Diffusion Tensor Imaging (DTI). One of the goals of DTI is to retrieve the main fiber tracts from the diffusion tensor field. However, it is obtained from noisy images typical of MR acquisition (we will see in the next chapter an optimal workflow to estimate tensors from raw MR data) and needs to be regularized before being further analyzed. A naive but simple and often efficient regularization on signal or images is the Gaussian filter. Its generalization to tensor fields is quite straightforward using once again weighted means (Section 4.3.1 below). An alternative is to consider a regularization using diffusion. This will be the subject of Sections 4.3.3 and 4.3.4.

4.3.1 Gaussian Filtering

In the continuous setting, the convolution of a vector field $F_0(x)$ by a Gaussian is:

$$F(x) = \int_y G_\sigma(y-x)\, F_0(y)\, dy.$$

In the discrete setting, coefficients are renormalized since the neighborhood $\mathcal{V}$ is usually limited to points within one to three times the standard deviation:

$$F(x) = \frac{\sum_{u\in\mathcal{V}(x)} G_\sigma(u)\, F_0(x+u)}{\sum_{u\in\mathcal{V}(x)} G_\sigma(u)} = \arg\min_F \sum_{u\in\mathcal{V}(x)} G_\sigma(u)\, \|F_0(x+u) - F\|^2.$$

Like in Section 4.2.2, this weighted mean can be solved in the affine-invariant framework using our intrinsic gradient descent scheme. The Log-Euclidean formulation is once again a weighted mean on the tensor logarithms mapped back to the tensor space via the matrix exponential.

We illustrate in Fig. 4.8 the comparative Gaussian filtering of a slice of a DT MR image using the Euclidean metric on the coefficients (since weights are positive, a weighted sum of positive definite matrices is still positive definite) and our Riemannian metrics. One can see that tensors around the ventricles tend to swell and blur the nearby smaller anisotropic ones (e.g., as inside the genu tract). Indeed, the large isotropic tensors within the ventricles largely influence the smoothing of the nearby region due to the fact that in general, the Euclidean average tensor of a set of tensors have a greater determinant (i.e., tensor volume) than each sample. This swelling effect is problematic and gives the impression that the ventricles have swelled after the smoothing. Riemannian metrics prevent this undesirable behavior and gives results intuitively close to what we are expecting of a Gaussian smoothing on tensors.

However, Gaussian smoothing is in general not the best filtering approach, in particular when one wants to perform fiber reconstruction as a final processing. In that case, one would like to preserve the discontinuities of the tensor fields using an anisotropic regularization procedure. For instance, the smoothed field of Fig. 4.8 right shows a corpus callosum less anisotropic than it should and dimished in size

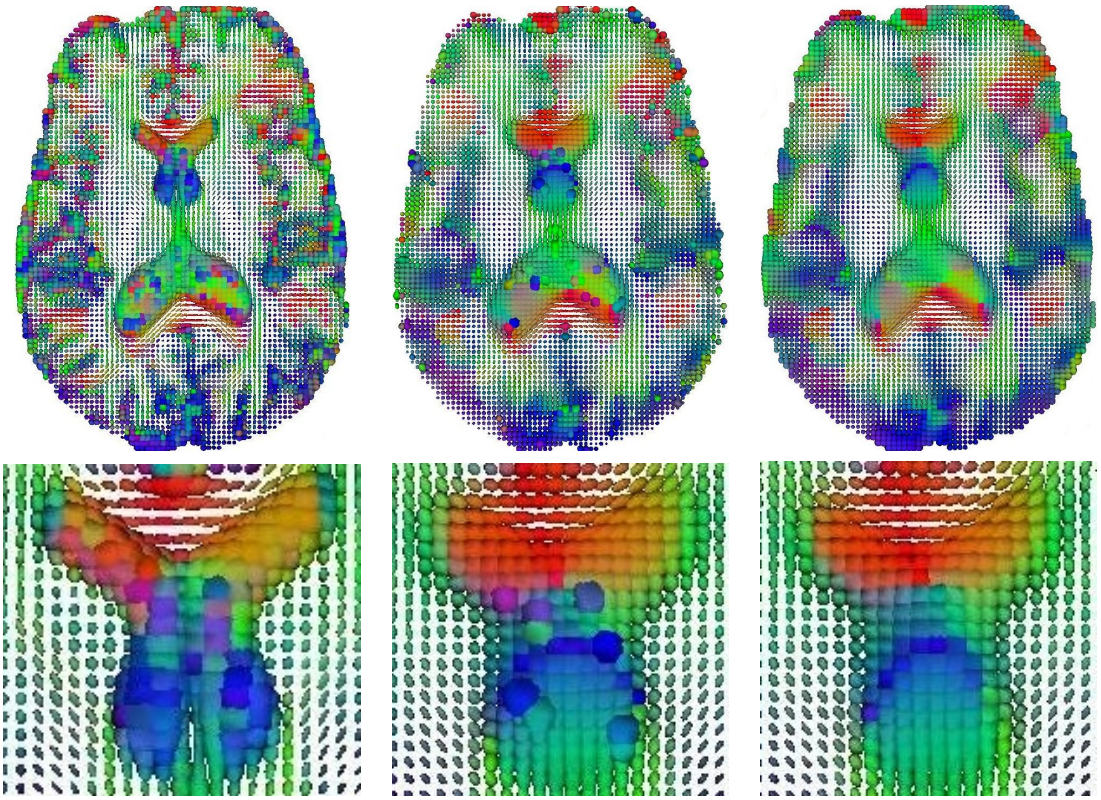


Figure 4.8: Regularization of a DTI slice by isotropic Gaussian filtering. **Top:** The original slice (left), after Euclidean filtering (middle), and after affine-invariant or Log-Euclidean filtering (right). **Bottom:** Close-up on the junction between the ventricles and the genu of the corpus callosum. Euclidean filtering produces outliers (large tensors inside the ventricles for example), due to the fact that this metric advantages large values. Notice how the ventricles "swelled" compared to the Riemannian filtering, which is closer to what we would expect of a Gaussian smoothing on tensors.

due to the influence of the tensors inside the ventricles: we would rather prefer to have the boundaries between those to regions preserved by the smoothing. But first of all, one needs to define what the "discontinuities" of a tensor field are. As we are in the discrete setting, we are not looking for discontinuities in the topological sense, but rather for places where the tensor field varies much more rapidly within a given neighborhood. In other words, we need to express the spatial gradient of a tensor image, and to measure its strength. This last point is intrinsically related to the metric used on tensors.

4.3.2 Spatial gradient of Tensor fields

On a n-dimensional vector field $F(x) = (f_1(x_1, \dots x_d), \dots f_n(x_1, \dots x_d))^{\mathrm{T}}$ over $\mathbb{R}^d$, one may express the spatial gradient in an orthonormal basis as:

$$\nabla F^{\mathrm{T}} = \left(\frac{\partial F}{\partial x}\right) = [\partial_1 F, \dots \partial_d F] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}, & \cdots & \frac{\partial f_1}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}, & \cdots & \frac{\partial f_n}{\partial x_d} \end{bmatrix}.$$

The linearity of the derivatives implies that we could use directional derivatives in more than the d orthogonal directions. This is especially well adapted to stabilize the discrete computations: the finite difference estimation of the directional derivative is $\partial_u F(x) = F(x+u) - F(x)$. By definition, the spatial gradient is related to the directional derivatives through $< \nabla F | u >= \partial_u F(x)$. Thus, we may compute ∇F as the matrix that best approximates (in the least-square sense) the directional derivatives in the neighborhood $\mathcal{V}$ (e.g., 6 or 26 connectivity in 3D):

$$\begin{aligned} \nabla F(x) &= \arg\min_G \sum_{u \in \mathcal{V}} \|G^{\mathrm{T}}\, u - \partial_u F(x)\|^2 = \left(\sum_{u \in \mathcal{V}} u\, u^{\mathrm{T}}\right)^{(-1)} \left(\sum_{u \in \mathcal{V}} u\, \partial_u F(x)^{\mathrm{T}}\right) \\ &\simeq \left(\sum_{u \in \mathcal{V}} u\, u^{\mathrm{T}}\right)^{(-1)} \left(\sum_{u \in \mathcal{V}} u\, (F(x+u) - F(x))^{\mathrm{T}}\right). \end{aligned}$$

It was experimentally shown in other applications (e.g. to compute the Jacobian of a deformation field in non-rigid registration [Rey 2002, p. 169]) that this gradient approximation scheme was more stable and much faster than computing all derivatives using convolutions, for instance by the derivative of the Gaussian.

To quantify the local amount of variability independently of the space direction, one usually takes the norm of the gradient: $\|\nabla F(x)\|^2 = \sum_{i=1}^{d} \|\partial_i F(x)\|^2$. Once again, this can be approximated using all directional derivatives in the neighborhood

$$\|\nabla F(x)\|^2 \simeq \frac{d}{Card(\mathcal{V})} \sum_{u \in \mathcal{V}} \frac{\|F(x+u) - F(x)\|^2}{\|u\|^2}, \tag{4.3}$$

where $Card(\mathcal{V})$ represents the total number of neighbors at the considered point (typical values in 3D are 6 or 26). Notice that this approximation is consistent with the previous one only if the directions u are normalized to unity.

4.3.2.1 Affine-Invariant Formulation:

For the tensor manifold endowed with the affine-invariant metric, we can proceed similarly, except that the directional derivatives $\partial_i \Sigma(x)$ are now tangent vectors of $T_{\Sigma(x)}\mathcal{M}$. They can be approximated just like above using finite "differences" in the exponential chart:

$$\partial_u \Sigma(x) \simeq \overrightarrow{\Sigma(\mathrm{x})\, \Sigma(\mathrm{x}+\mathrm{u})} = \Sigma(x)^{\frac{1}{2}}\, \log\left(\Sigma(x)^{-\frac{1}{2}}\, \Sigma(x+u)\, \Sigma(x)^{-\frac{1}{2}}\right)\, \Sigma(x)^{\frac{1}{2}}. \tag{4.4}$$

As observed in Section 3.3.5, we must be careful that this directional derivative is expressed in the standard matrix coordinate system (coefficients). Thus, the basis is not orthonormal: to quantify the local amount of variation, we have to take the metric at the point $\Sigma(x)$ into account, so that:

$$\|\nabla\Sigma(x)\|^2_{\Sigma(x)} = \sum_{i=1}^{d} \|\partial_i\Sigma(x)\|^2_{\Sigma(x)} \simeq \frac{d}{Card(\mathcal{V})} \sum_{u\in\mathcal{V}} \frac{\left\|\log\left(\Sigma(x)^{-\frac{1}{2}}\,\Sigma(x+u)\,\Sigma(x)^{-\frac{1}{2}}\right)\right\|^2_2}{\|u\|^2}. \tag{4.5}$$

4.3.2.2 Log-Euclidean Formulation:

Log-Euclidean formulation of the directional derivatives is straightforward. Using the same finite difference scheme, one can write:

$$\partial_u\Sigma(x) \simeq \log\left(\Sigma(x+u)\right) - \log\left(\Sigma(x)\right)$$

The norm of the gradient is:

$$\|\nabla\Sigma(x)\|^2_{\Sigma(x)} \simeq \frac{d}{Card(\mathcal{V})} \sum_{u\in\mathcal{V}} \frac{\|\log\left(\Sigma(x+u)\right) - \log\left(\Sigma(x)\right)\|^2_2}{\|u\|^2}. \tag{4.6}$$

Example of the gradient norm of a DTI slice is shown in Fig. 4.9. The same windowing was used for both Euclidean and Riemannian gradients, so that images can be visually compared. The affine-invariant and Log-Euclidean metrics gave results almost similar, we display only one of them. What is striking is that the Riemannian gradient exhibits a more detailed structure inside the brain white matter compared to the Euclidean gradient. In particular, the Euclidean gradient shows a high contrast around the ventricles and at the periphery of the brain: these regions are made with large isotropic tensors (cortico-spinal fluid inside the ventricles, and grey matter at the periphery), which once again influence a lot the result. By contrast, the Riemannian gradient acts like a "normalizer" for the gradient: interface between large and smaller tensors (like around the ventricles) have the same contrast as interfaces between two tensor regions of the same size but with different orientations (like within the white matter). It is especially interesting since such interfaces are generally delimiting two distinct fiber tracts, and should be preserved during an anisotropic smoothing for example. This is the topic of the next sections.

4.3.3 Filtering using PDEs

Regularizing a scalar, vector or tensor field F aims at reducing the amount of its spatial variations. The first order measure of such variations is the spatial gradient ∇F that we dealt with in the previous section. To obtain a regularity criterion over the domain Ω, we just have to integrate: $Reg(F) = \int_\Omega \|\nabla F(x)\|^2\,dx$. Starting from an initial field $F_0(x)$, the goal is to find at each step a field $F_t(x)$ that minimizes the regularity criterion by gradient descent in the space of (sufficiently smooth and square integrable) functions.

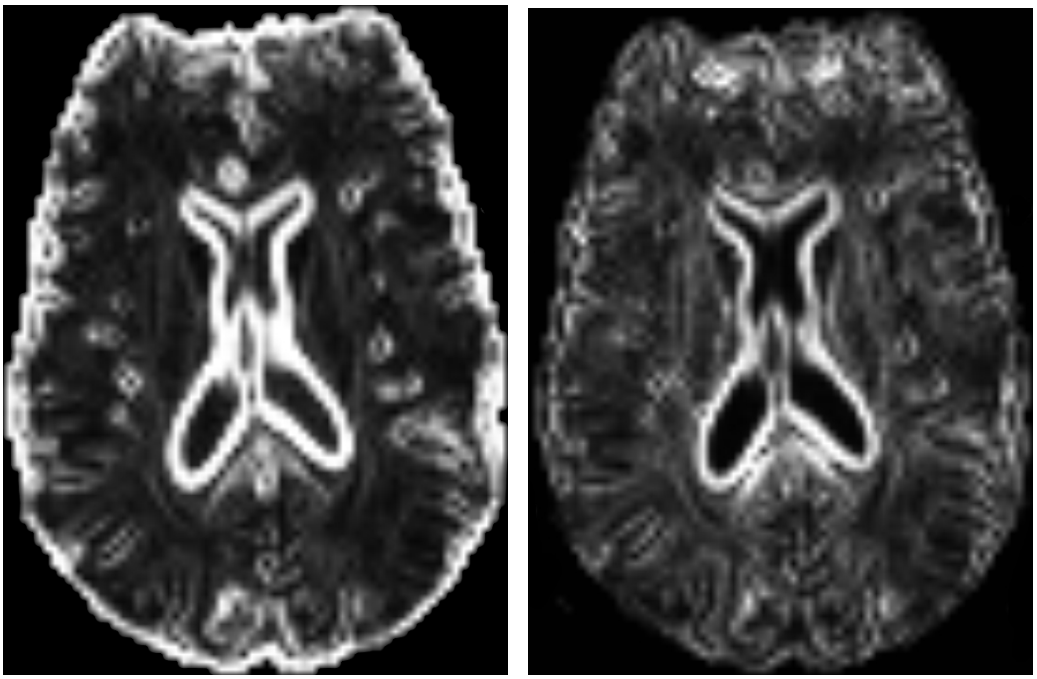


Figure 4.9: Norm of the gradient of a tensor field (same slice as Fig. 4.8), computed on the coefficients with Eq. 4.3 (left), and computed with either the affine-invariant (Eq. 4.5) or Log-Euclidean metrics (Eq. 4.6) (right). One can see much more structures within the brain, which will be better preserved during an anisotropic regularization step.

To compute the first order variation, we write a Taylor expansion for an incremental step in the direction of the field H.

$$Reg(F + \varepsilon\, H) = Reg(F) + 2\,\varepsilon \int_\Omega \langle\, \nabla F(x) \mid \nabla H(x) \rangle\; dx + O(\varepsilon^2).$$

We get the directional derivative: $\partial_H Reg(F) = 2 \int_\Omega \langle\, \nabla F(x) \mid \nabla H(x) \rangle\; dx$. To compute the steepest descent, we now have to find the gradient $\nabla Reg(F)$ such that for all variation H, we have $\partial_H Reg(F) = \int_\Omega \langle\, \nabla Reg(F)(x) \mid H(x) \rangle_{F(x)}\; dx$. Notice that $\nabla Reg(F)(x)$ and $H(x)$ are elements of the tangent space at $F(x)$, so that the scalar product should be taken at $F(x)$ for a tensor field.

4.3.3.1 The case of a scalar field

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a scalar field. Our regularization criterion is $Reg(f) = \int_\Omega \|\nabla f(x)\|^2\; dx$. Let us introduce the divergence $\mathrm{div}(.) = \langle\, \nabla \mid . \rangle$ and the Laplacian operator $\Delta f = \mathrm{div}(\nabla f)$. The divergence is usually written $\nabla^{\mathrm{T}} = (\partial_1, \ldots, \partial_d)$, so that in an orthonormal coordinate system we have $\Delta f = \langle\, \nabla \mid \nabla f \rangle = \sum_{i=1}^d \partial_i^2 f$. Let now $G(x)$ be a vector field. Typically, we will use $G(x) = \nabla f(x)$. Using the standard differentiation rules, we have:

$$\mathrm{div}(h\, G) = \langle\, \nabla \mid h\, G \rangle = h\, \mathrm{div}(G) + \langle\, \nabla h \mid G \rangle\,.$$

Now, thanks to the Green's formula (see e.g. [Gallot 1993]), we know that the flux going out of the boundaries of a (sufficiently smooth) region Ω is equal to the integral of the divergence inside this region. If we denote by n the normal pointing outward at a boundary point, we have:

$$\int_{\partial\Omega} \langle\, h\, G \mid n \rangle \; dn = \int_{\Omega} \mathrm{div}(h\, G) = \int_{\Omega} h\, \mathrm{div}(G) + \int_{\Omega} \langle\, \nabla h \mid G \rangle\,.$$

This result can also be interpreted as an integration by part in $\mathbb{R}^d$. Assuming homogeneous Neumann boundary conditions (gradient orthogonal to the normal on $\partial\Omega$: $\langle\, G \mid n \rangle = 0$), the flow across the boundary vanishes, and we are left with: $\int_\Omega \langle\, G \mid \nabla h \rangle = -\int_\Omega h\, \mathrm{div}(G)$. Thus, coming back to our original problem, we have:

$$\partial_h Reg(f)(x) = 2 \int_\Omega \langle\, \nabla f(x) \mid \nabla h(x) \rangle \; dx = -2 \int_\Omega h(x)\, \Delta f(x)\, dx.$$

Since this last formula is no more than the scalar product on the space $L_2(\Omega, \mathbb{R})$ of square integrable functions, we end-up with the classical Euler-Lagrange equation: $\nabla Reg(f) = -2\Delta f(x)$. The evolution equation used to filter the data is thus

$$f_{t+1}(x) = f_t(x) - \varepsilon \nabla Reg(f)(x) = f_t(x) + 2\,\varepsilon \Delta f_t(x).$$

4.3.3.2 The vector case

Let us decompose our vector field $F(x)$ into its n scalar components $f_i(x)$. Likewise, we can decompose the $d \times n$ gradient ∇F into the gradient of the n scalar components $\nabla f_i(x)$ (columns). Thus, choosing an orthonormal coordinate system on the space $\mathbb{R}^n$, our regularization criterion is decoupled into n independent scalar regularization problems:

$$Reg(F)(x) = \sum_{i=1}^{n} \int_\Omega \|\nabla f_i(x)\|^2 \; dx = \sum_{i=1}^{n} Reg(f_i).$$

Thus, each component f_i has to be independently regularized with the Euler-Lagrange equation: $\nabla Reg(f_i) = -2\Delta f_i$. With the convention that the Laplacian is applied component-wise (so that we still have $\Delta F = \mathrm{div}(\nabla F) = \nabla^{\mathrm{T}}\, \nabla F = (\Delta f_1, \ldots \Delta f_n)^{\mathrm{T}}$), we end-up with the vectorial equation:

$$\nabla Reg(F) = -2\Delta F \qquad \text{for} \qquad Reg(F) = \int_\Omega \|\nabla F(x)\|\, dx.$$

The associated evolution equation is $F_{t+1}(x) = F_t(x) + 2\,\varepsilon \Delta F_t(x)$.

4.3.3.3 Tensor Fields

Let $(x_1, \ldots x_d)$ be an orthonormal coordinate system of $\mathbb{R}^d$ and $\Sigma(x) \in \mathcal{S}ym_n^+$ a tensor field over $\mathbb{R}^d$. Our regularization criterion is:

$$Reg(\Sigma) = \int_\Omega \|\nabla \Sigma(x)\|^2_{\Sigma(x)} \; dx = \sum_{i=1}^{d} \int_\Omega \|\partial_i \Sigma(x)\|^2_{\Sigma(x)}\,. \tag{4.7}$$

The idea is to write this criterion as the Trace of sums and products of standard Euclidean matrices and to compute its directional derivative $\partial_W Reg$ for a perturbation field W. This expression contains of course derivatives $\partial_i W$ that we need to integrate. However, as everything is expressed in the standard Euclidean chart (matrix coefficients), and assuming the proper Neumann boundary conditions, we shall safely use the previous integration by part formula $\int_\Omega \mathrm{Tr}((\partial_i W)\,\Lambda_i) = -\int_\Omega \mathrm{Tr}(W\,(\partial_i \Lambda_i))$. Notice that we are using the matrix coefficients only as a chart and not as a metric. Eventually, we rewrite the obtained expression in terms of our Riemannian metrics to obtain the formula defining the gradient of the criterion: $\partial_W Reg = \int_\Omega \langle\, W \mid \nabla Reg \rangle_\Sigma$. By identification, we get: $\nabla Reg(\Sigma) = -2\,\Delta\Sigma$, where Δ is the Laplace-Beltrami operator on our manifold that we need to express in the affine-invariant and Log-Euclidean frameworks. Before detailing the Laplace-Beltrami operators, we give the gradient descent on the regularization criterion with the intrinsic geodesic marching scheme of Section 3.3.6:

$$\begin{aligned}
\Sigma_{t+1}(x) &= \exp_{\Sigma_t(x)}\left(-\varepsilon\,\nabla Reg(\Sigma)(x)\right) = \exp_{\Sigma_t(x)}\left(2\,\varepsilon\,\Delta\Sigma(x)\right) \\
&= \Sigma_t(x)^{1/2}\exp\left(2\,\varepsilon\,\Sigma_t^{-1/2}(x)\Delta\Sigma(x)\Sigma_t^{-1/2}(x)\right)\Sigma_t^{1/2}(x) \qquad \text{(AI)} \quad (4.8) \\
&= \exp\left(\log\left(\Sigma_t(x)\right) + 2\,\varepsilon\,\Delta\Sigma(x)\right) \qquad \text{(LE)} \quad (4.9)
\end{aligned}$$

The Affine-Invariant Laplace-Beltrami operator: We show in Appendix A.1 that the Laplace-Beltrami in the affine-invariant case can be expressed as:

$$\Delta\Sigma = \sum_{i=1}^{d} \Delta_i \Sigma \qquad \text{with} \qquad \Delta_i\Sigma = \partial_i^2\Sigma - (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma). \tag{4.10}$$

As we can see, the Euclidean second order directional derivatives $\partial_i^2\Sigma$ are corrected by an additional term due to the curvature of our manifold.

For the numerical computation of the Laplacian, we may approximate the first and second order AI derivatives by Euclidean finite differences. This gives a fourth order approximation of the Laplace-Beltrami operator (details can be found in Appendix A.2). However, this numerical scheme is extrinsic since it is based on (Euclidean) differences of tensors. We propose here an intrinsic scheme based on the exponential chart at the current point. We already know from Eq. 4.4 that $\overrightarrow{\Sigma(x)\Sigma(x+u)}$ is an approximation of the first order directional derivative $\partial_u\Sigma(x)$. We show in Appendix A.2 that $\overrightarrow{\Sigma(x)\Sigma(x+u)} + \overrightarrow{\Sigma(x)\Sigma(x-u)}$ is a fourth order approximation of the Laplace Beltrami operator in the direction u:

$$\Delta_u\Sigma = \partial_u^2\Sigma - (\partial_u\Sigma)\Sigma^{(-1)}(\partial_u\Sigma) = \frac{\overrightarrow{\Sigma(x)\Sigma(x+u)} + \overrightarrow{\Sigma(x)\Sigma(x-u)}}{\|u\|^2} + O(\|u\|^4). \tag{4.11}$$

To compute the complete manifold Laplacian of Eq. 4.10, we just have to compute the above numerical approximations of the tensor field derivatives along d orthonormal basis vectors x_i. However, like for the computation of the gradient, we may improve the stability of the numerical scheme by averaging the derivatives in

all possible directions in the neighborhood $\mathcal{V}$. Assuming a symmetric and isotropic neighborhood, we finally obtain:

$$\begin{aligned}\Delta\Sigma(x) &= \frac{d}{Card(\mathcal{V})}\sum_{u\in\mathcal{V}}\Delta_u\Sigma(x)\\ &\simeq \frac{d}{Card(\mathcal{V})}\sum_{u\in\mathcal{V}}\left(\frac{\Sigma^{1/2}(x)\log\left(\Sigma(x)^{-1/2}\Sigma(x+u)\Sigma(x)^{-1/2}\right)\Sigma^{1/2}(x)}{\|u\|^2}+\right.\\ &\qquad\left.\frac{\Sigma^{1/2}(x)\log\left(\Sigma(x)^{-1/2}\Sigma(x-u)\Sigma(x)^{-1/2}\right)\Sigma^{1/2}(x)}{\|u\|^2}\right) \qquad (4.12)\end{aligned}$$

The Log-Euclidean Laplace-Beltrami operator: The Log-Euclidean formulation of the Laplace-Beltrami operator boils down to the vectorial case on the logarithms of tensors:

$$\Delta\Sigma = \sum_{i=1}^{d}\Delta_i\Sigma \qquad \text{with} \qquad \Delta_i\Sigma = \partial_i^2\log(\Sigma)$$

Using a finite difference scheme over a neighborhood $\mathcal{V}$, one can write:

$$\Delta\Sigma(x) \simeq \frac{d}{Card(\mathcal{V})}\sum_{u\in\mathcal{V}}\frac{\log(\Sigma(x+u)) - 2\log(\Sigma(x)) + \log(\Sigma(x-u))}{\|u\|^2} \qquad (4.13)$$

Remark that contrary to the affine-invariant case, one is not obliged to go back to the tensor space (by taking the matrix exponential) at each iteration. Indeed, the evolution equation acts on the logarithms, which are computed once for all. Taking the exponential is required at the end of the process to obtain the final tensors.

In the following, we illustrate the use of the PDE tensor filtering strategy we presented with the anisotropic filtering.

4.3.4 Anisotropic Filtering

In practice, we would like to filter within homogeneous regions, but not across their boundaries. The basic idea is to penalize the smoothing in the directions where the derivative is important [Perona 1990, Gerig 1992]. If $c(.)$ is a weighting function decreasing from $c(0) = 1$ to $c(+\infty) = 0$, this can be realized directly in the discrete implementation of the Laplacian (Eq. 4.12 and 4.13): the contribution of the spatial direction u to the Laplace-Beltrami operators is weighted by our decreasing function according to the norm of the gradient in that direction. The important point here is that we should evaluate the norm of the directional derivatives of the tensor field with the corresponding metric (affine-invariant or Log-Euclidean). With our finite difference approximations, this leads to the following modified Laplacian:

$$\Delta_{aniso}\Sigma(x) = \frac{d}{Card(\mathcal{V})}\sum_{u\in\mathcal{V}} c\left(\|\partial_u\Sigma(x)\|_{\Sigma(x)}\right)\ \Delta_u\Sigma(x) \qquad (4.14)$$

Figures 4.10 and 4.11 present example results of this simple anisotropic filtering scheme on synthetic and real DTI images. Although the affine-invariant implementation was used, differences between AI and Log-Euclidean implementation are very small and discussed later. We used the function $c(x) = \exp\left(-x^2/\kappa^2\right)$, where the threshold κ controls the amount of local regularization: for a gradient magnitude greater than 2 to 3 times κ, there is virtually no regularization, while the field is almost linearly smoothed for gradient magnitudes below a fraction (say 0.1) of κ. For both synthetic and real data, the histogram of the gradient norm is clearly bimodal so that the threshold κ is easily determined.

In Fig. 4.10, we generated a tensor field with a discontinuity, and add independent Gaussian noises according to Section 4.1.1. The anisotropic smoothing perfectly preserves the discontinuity while completely smoothing each region. In this synthetic experiment, we retrieve tensor values that are very close to the initial tensor field. This could be expected since the two regions are perfectly homogeneous. After enough regularization steps, each region is a constant field equal to the mean of the 48 initially noisy tensors of the region. Thus, similarly to the Euclidean mean of identically and independently distributed measurements, we expect the standard deviation of the regularized tensors to be roughly $7 \simeq \sqrt{48}$ times smaller than the one of the noisy input tensors.

Figure 4.10: **Left:** 3D synthetic tensor field with a clear discontinuity. **Middle:** The field has been corrupted by a Gaussian noise (in the Riemannian sense). **Right:** result of the regularization after 30 iterations (time step $\varepsilon = 0.01$, $\kappa = 0.05$).

In Figure 4.11, we display the evolution of (a slice of) the tensor field, the norm of the gradient and the FA at different steps of the anisotropic filtering of a 3D DTI. One can see that the tensors are regularized in "homogeneous" regions (ventricles, temporal areas), while the main tracts are left unchanged. It is worth noticing that the fractional anisotropy is very well regularized even though this measure has almost nothing in common with our invariant tensor metric.

Figure 4.12 displays closeups around the ventricles to compare the different regularization methods developed so far. One can see that the Riemannian metrics (AI and LE) give much less weight to large tensors, thus providing a regularization which is more robust to outliers. The anisotropic filtering further improves the results by preserving the discontinuities of the tensor field (e.g. at the boundary of the ventricles), but also the discontinuities of the tensor orientation, which is exactly what is needed for fiber tracking in DTI.

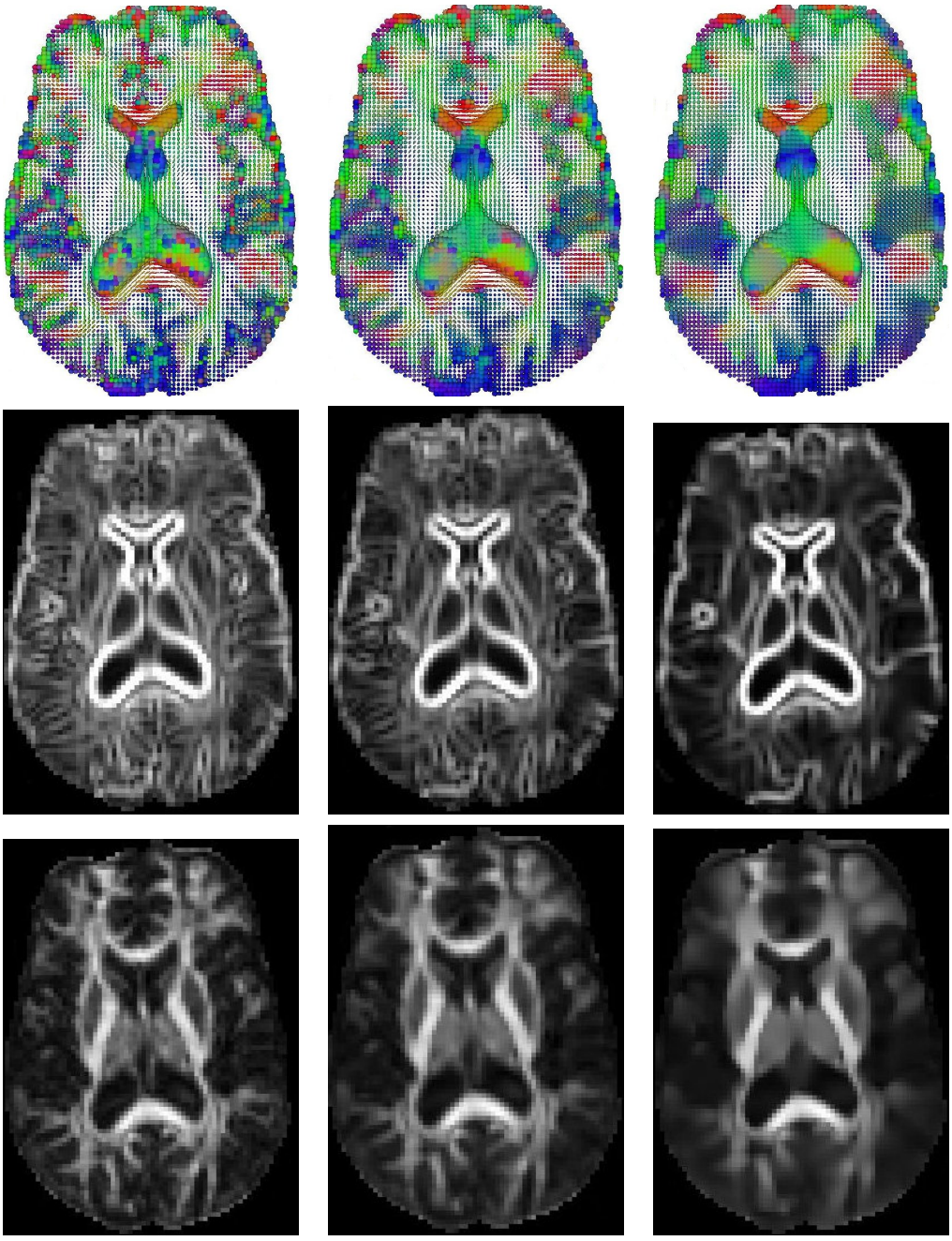


Figure 4.11: Anisotropic filtering of a DTI slice (time step 0.01, $\kappa = 0.05$). From left to right: at the beginning, after 10 and after 50 iterations. **Top:** A 3D view of the tensors as ellipsoids. The results are to be compared with the isotropic Gaussian filtering displayed in Figure 4.8. **Middle:** Riemannian norm of the gradient. **Bottom:** Fractional anisotropy. Note that after 50 iterations we start bluring out some structures like the sulci in the exterior part of the images. The number of iterations should not exceed 30 to 40 for an optimal result.

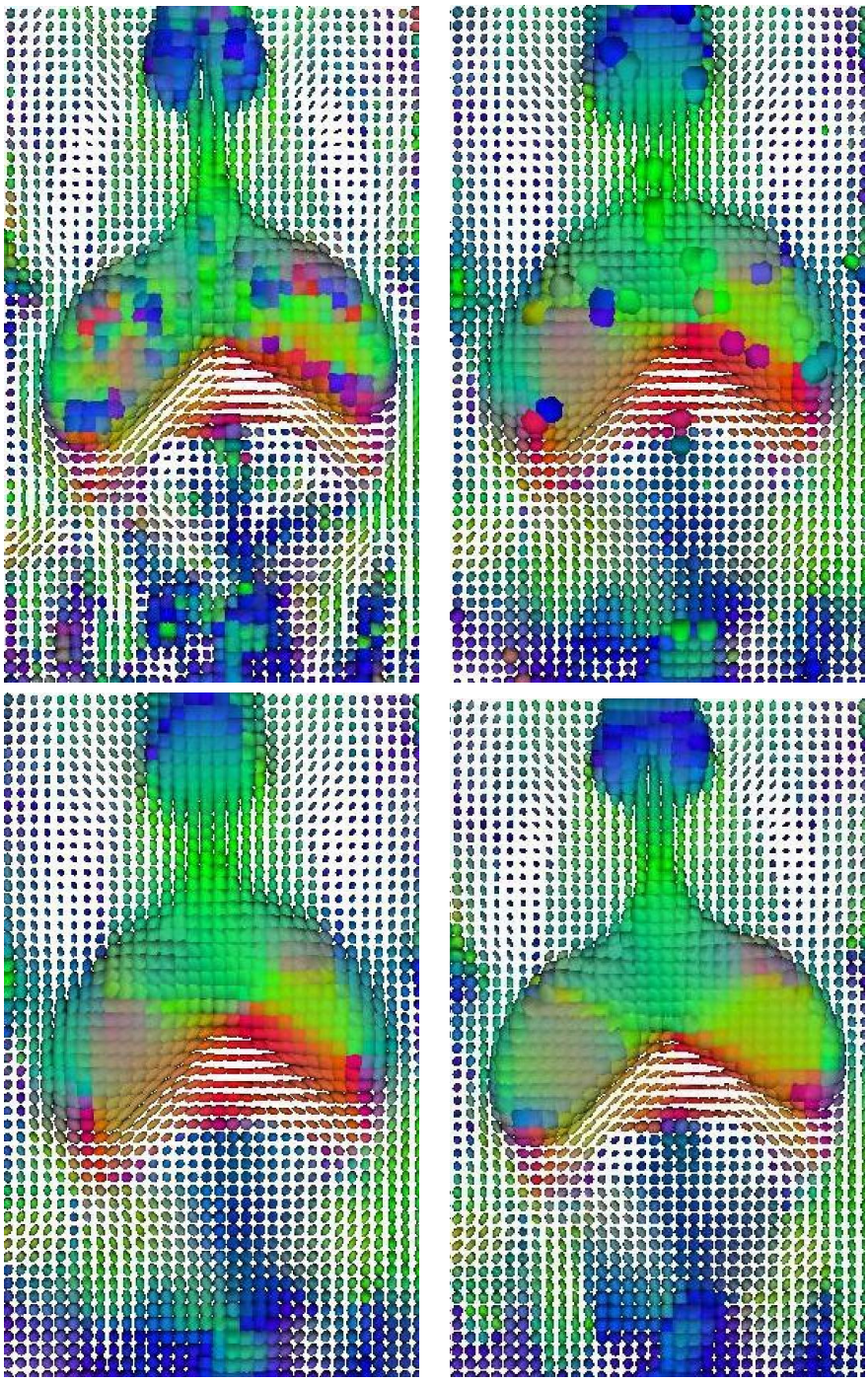


Figure 4.12: Closeup on the results of the different filtering methods around the splenium of the corpus callosum.**Top left:** Original image. **Top right:** Gaussian filtering using the Euclidean metric (5x5 window, $\sigma = 2.0$). This metric gives too much weight to tensors with large eigenvalues, thus leading to clear outliers in the ventricles or in the middle of the splenium tract. **Bottom left:** Gaussian filtering using the Riemannian metrics (5x5 window, $\sigma = 2.0$). Outliers disappeared, but the discontinuities are not well preserved, for instance in the ventricles at the level of the cortico-spinal tracts (upper-middle part of the images). **Bottom right:** Anisotropic filtering in the Riemannian frameworks (time step 0.01, 50 iterations). The ventricles boundary is very well preserved with an anisotropic filter and both isotropic (ventricles) and anisotropic (splenium) regions are regularized. Note that the U-shaped tracts at the boundary of the grey/white matter (lower left and right corners of each image) are preserved with an anisotropic filter and not with a Gaussian filter.

We show on Fig. 4.13 an illustration of the swelling effect with the Euclidean smoothing. First of all, Euclidean anisotropic filtering is difficult to realize because the gradient descent is not a convex operation and negative eigenvalues appear very quickly. To prevent negatives values to come up, one has to take an integration step extremely small (in this case 1.10^{-6}) and increase the number of iterations so that the product: number of iterations $\times$ integration step remains the same in the Euclidean and Riemannian cases. One notices that the Euclidean result seriously suffers from the swelling effect: tensors are restored but are much bigger than the original values. This effect is due to the fact that this type of smoothing can be seen as a weighted mean calculation at each positions and at each iterations. Of course, the weights are chosen so that it gives an anisotropic behavior to the smoothing. We saw previously that in general, Euclidean averaging of tensors produces results that have a larger determinant (i.e., volume) than the samples. If this process is iterated an important number of times, this leads to the swelling effect of Fig. 4.13. As expected, the Riemannian smoothing does not suffer from this drawback and correctly denoised homogeneous regions while preserving the boundary with the ventricles. On Fig. 4.13 bottom right, we displayed the difference, magnified by 100, between the affine-invariant and Log-Euclidean results! Differences are almost inexistent, which gives advantage to the Log-Euclidean framework as its computational time is by far less than the affine-invariant one (see table 4.1 for the computational times). Notice that the only differences appear in regions where tensor anisotropy is high, confirming the observation we made on the linear interpolation of Sec. 4.2.1.

	Euclidean	Affine-Invariant	Log-Euclidean
time (s/iteration)	1.8	20.24	1.8
total time	90002 (50000 ite)	1012 (50 ite)	90 (50 ite)

Table 4.1: Computational times for one iteration of the anisotropic diffusion scheme of Sec. 4.3.4. The dimensions of the tensor field are: $128 \times 128 \times 58$. Implementation was made in C++ using optimized numerical libraries (for the eigen decomposition), on a Pentium M at 2GHz with 1Gb of memory. Euclidean and Log-Euclidean frameworks are obviously much more performant than the affine-invariant one. However, one should retain that the integration step to make the Euclidean smoothing stable must be extremely low (for stability reasons), and consequently we had to increase the number of iterations proportionally: when Log-Euclidean and Affine-Invariant smoothings require only 50 iterations with an integration step of 0.1 to converge, one must perform 50000 iterations with an integration step of 0.0001 in the Euclidean case to obtain the same "amount" of smoothing!

4.4 Tensor Field Restoration

The pure diffusion is efficient to reduce the noise in the data, but it also reduces the amount of information. Moreover, the amount of smoothing is controlled by the time of diffusion (time step ε $\times$ the number of iterations), which is not an easy parameter to tune. At an infinite diffusion time, the tensor field will be com-

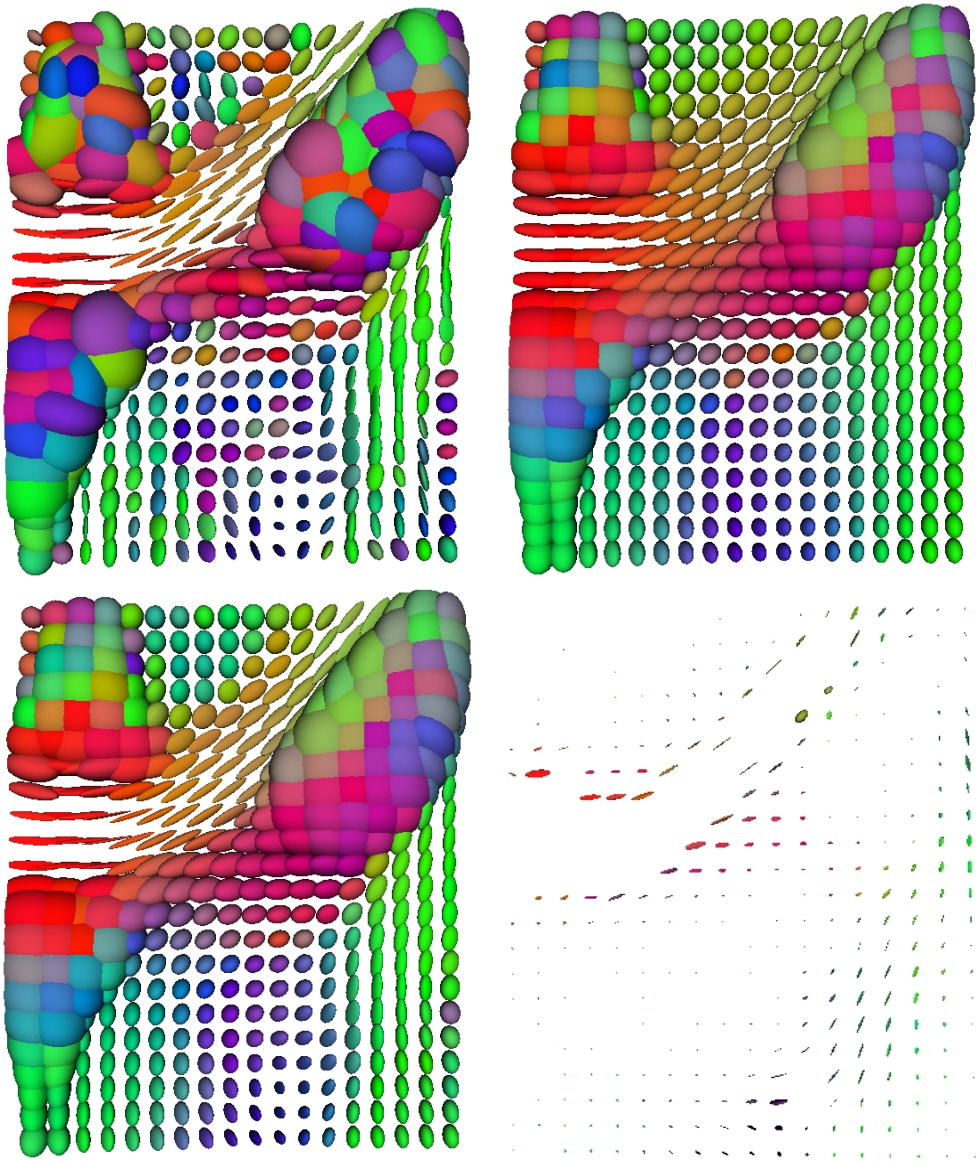


Figure 4.13: **Anisotropic Smoothing: Euclidean vs. Riemannian.** Top Left: Original slice (closeup around the left big ventricle). Top Right: Euclidean anisotropic smoothing. Bottom Left: Log-Euclidean smoothing. Bottom Right: Difference between affine-invariant and Log-Euclidean smoothing magnified by 100.

pletely homogeneous (or homogeneous by part for the anisotropic diffusion scheme), with a value corresponding to the mean of the measurements over the region (with Neumann boundary conditions). Thus, the absolute minimum of our regularization criterion alone is of little interest.

To keep close to the measured tensor field $\Sigma_0(x)$ while still regularizing, a more theoretically grounded approach is to consider an optimization problem with a competition between a data fidelity term (also called data attachment term) and a

possibly non-linear anisotropic regularization term:

$$C(\Sigma) = Sim(\Sigma, \Sigma_0) + \lambda\, Reg(\Sigma)$$

Like before, the intrinsic evolution equation leading to a local minimum is:

$$\Sigma_{t+1}(x) = \exp_{\Sigma_t(x)} \left(-\varepsilon\ \left(\nabla Sim(\Sigma, \Sigma_0)(x) + \lambda\, \nabla Reg(\Sigma)(x)\right)\right),$$

In the following, we will focus on defining the data attachment term. We use the isotropic regularization of Eq. 4.7 as regularization term.

4.4.1 A least-squares data fidelity term

Usually, one considers that the data (e.g., a scalar image or a vector field $F_0(x)$) are corrupted by a uniform (isotropic) Gaussian noise independent at each spatial position. With a maximum likelihood approach, this amounts to considering a least-squares criterion $Sim(F) = \int_\Omega \|F(x) - F_0(x)\|^2\, dx$. Like in the previous section, we compute the first order variation by writing the Taylor expansion:

$$Sim(F + \varepsilon\, H) = Sim(F) + 2\,\varepsilon \int_\Omega \langle\, H(x)\ |\ F(x) - F_0(x)\,\rangle\ dx + O(\varepsilon^2).$$

This time, the directional derivative $\partial_H Sim(F)$ is directly expressed using a scalar product with H in the proper functional space, so that the steepest ascent direction is $\nabla Sim(F) = 2\,(F(x) - F_0(x))$.

On tensors, assuming a uniform (generalized) Gaussian noise independent at each position also leads to a least-squares criterion through a maximum likelihood approach. The only difference is that it uses either one or the other of our Riemannian frameworks:

$$Sim(\Sigma) = \int_\Omega \mathrm{dist}^2\left(\Sigma(x)\,,\ \Sigma_0(x)\right)\ dx = \int_\Omega \left\|\overrightarrow{\Sigma(x)\Sigma_0(x)}\right\|^2_{\Sigma(x)}\ dx,$$

Thanks to the properties of the exponential map, one can show that the gradient of the squared distance is: $\nabla_\Sigma\, \mathrm{dist}^2(\Sigma\,,\ \Sigma_0) = -2\,\overrightarrow{\Sigma\Sigma_0}$ [Pennec 2004a]. Finally, we obtain a steepest ascent direction (gradient) of our criterion which is similar to the vector case (in formula):

$$\nabla Sim(\Sigma)(x) = -2\,\overrightarrow{\Sigma(x)\Sigma_0(x)}, \tag{4.15}$$

which gives for the affine-invariant case:

$$\nabla Sim(\Sigma)(x) = -2\,\Sigma^{1/2}(x) \log\left(\Sigma^{-1/2}(x)\Sigma_0(x)\Sigma^{-1/2}(x)\right)\Sigma^{1/2}(x),$$

and for the Log-Euclidean case:

$$\nabla Sim(L)(x) = -2\ \left(\log\left(\Sigma_0(x)\right) - \log\left(\Sigma(x)\right)\right).$$

4.4.2 A least-squares attachment term for sparsely distributed tensors

Now, let us consider the case where we do not have a dense measure of our tensor field, but only N measures Σ_i at irregularly distributed sample points x_i. Assuming a uniform Gaussian noise (w.r.t. the considered metric) independent at each position still leads to a least-squares criterion:

$$Sim(\Sigma) = \sum_{i=1}^{N} \text{dist}^2\left(\Sigma(x_i)\ ,\ \Sigma_i\right) = \int_\Omega \sum_{i=1}^{N} \text{dist}^2\left(\Sigma(x)\ ,\ \Sigma_i\right)\ \delta(x - x_i)\, dx.$$

In this criterion, the tensor field $\Sigma(x)$ is related to the data only at the measurement points x_i through the Dirac distributions $\delta(x - x_i)$. If the introduction of distributions may be dealt with for the theoretical differentiation of the criterion with respect to the continuous tensor field Σ, it is a real problem for the numerical implementation. In order to regularize the problem, we consider the Dirac distribution as the limit of the Gaussian function G_σ when σ goes to zero. Using that scheme, our criterion becomes the limit case $\sigma = 0$ of:

$$Sim_\sigma(\Sigma) = \int_\Omega \sum_{i=1}^{N} \text{dist}^2\left(\Sigma(x)\ ,\ \Sigma_i\right)\ G_\sigma(x - x_i)\, dx. \tag{4.16}$$

From a practical point of view, we need to use a value of σ which is of the order of the spatial resolution of the grid on which $\Sigma(x)$ is evaluated, so that all measures can at least influence the neighboring nodes.

Now that we came back to a smooth criterion, we may differentiate it exactly as we did for the dense measurement setup. The first order variation is:

$$Sim_\sigma(\Sigma + \varepsilon W) = Sim_\sigma(\Sigma) - 2\,\varepsilon \sum_{i=1}^{N} \int_\Omega \left\langle\ W(x)\ \middle|\ G_\sigma(x - x_i)\ \overrightarrow{\Sigma(x)\Sigma_i}\ \right\rangle\ dx + O(\varepsilon^2),$$

so that we get:

$$\nabla Sim_\sigma(x) = -2 \sum_{i=1}^{N} G_\sigma(x - x_i)\ \overrightarrow{\Sigma(x)\Sigma_i}. \tag{4.17}$$

Of course, this gradient should be expressed with the affine-invariant or Log-Euclidean metric, depending on which framework is preferred.

We will now use the criteria described above for defining a new type of extrapolation by diffusion.

4.4.3 Extrapolation through diffusion

With the sparse data attachment term (4.16) and the isotropic first order regularization term (4.7), we are looking for a tensor field that minimizes its spatial variations

while interpolating (or more precisely approximating at the desired precision) the measurement values:

$$C(\Sigma) = \sum_{i=1}^{N} \int_{\Omega} G_\sigma(x - x_i)\ \mathrm{dist}^2\left(\Sigma(x_i)\ ,\ \Sigma_i\right) + \lambda \int_{\Omega} \|\nabla\Sigma(x)\|^2_{\Sigma(x)}\, dx.$$

According to the previous sections, the gradient of this criterion is

$$\nabla C(\Sigma)(x) = -2 \sum_{i=1}^{N} G_\sigma(x - x_i)\, \overrightarrow{\Sigma(x)\Sigma_i} - 2\,\lambda\,\Delta\Sigma(x).$$

Using our finite difference approximation scheme (Eq. 4.12), the intrinsic geodesic gradient descent scheme (Sec. 3.3.6) is finally:

$$\Sigma_{t+1}(x) = \exp_{\Sigma_t(x)}\left(\varepsilon\left\{\sum_{i=1}^{N} G_\sigma(x - x_i)\, \overrightarrow{\Sigma(x)\Sigma_i} + \lambda' \sum_{u\in\mathcal{V}} \frac{\overrightarrow{\Sigma(x)\Sigma(x+u)}.}{\|u\|^2}\right\}\right) \quad (4.18)$$

Last but not least, we need an initialization of the tensor field $\Sigma_0(x)$ to obtain a fully operational algorithm. This is easily done with any radial basis function approximation like in Section 4.2.3, or a soft closest point interpolation (i.e., nearest neighbor interpolation). Figure 4.14 displays the result of this algorithm on the interpolation between 4 tensors. On can see that the soft closest point approximation is well regularized into a constant field equal to the mean of the four tensors if data attachment term is neglected. On the contrary, a very small value of λ is sufficient for regularizing the field between known tensors (as soon as σ is much smaller than the typical spatial distance between two measurements).

The choice of the initialization is a critical issue from a computational point of view. For instance, starting with a constant (or any harmonic) field is not optimal: there is a null Laplacian everywhere, except at the immediate neighborhood of the sparse tensors, exactly where the data attachment term acts. Thus, we have a potentially destructive competition between the two terms of the criterion in these area: the data attachment term almost zeroes out the diffusion, resulting in a very slow convergence. On the contrary, starting with a nearest neighbor interpolation leads to a Laplacian which is non null on the boundaries of the Voronoi cells of the measurement points, i.e. the farthest possible place from the sparse measures. In that case, the Laplacian regularization will spread from these boundaries with no constraints until it reaches the counterbalancing forces of the data attachment term in the neighborhood of the sparse measurements. Thus, we may expect to reach the maximal efficiency in terms of convergence rate.

This last application puts an end to this chapter. We will now conclude on the Riemannian processing of DT-MRI using the AI or LE metrics.

4.5 Conclusion

In this chapter, we showed that any vector-processing algorithm could be extended to tensors in both affine-invariant and Log-Euclidean frameworks. We detailed

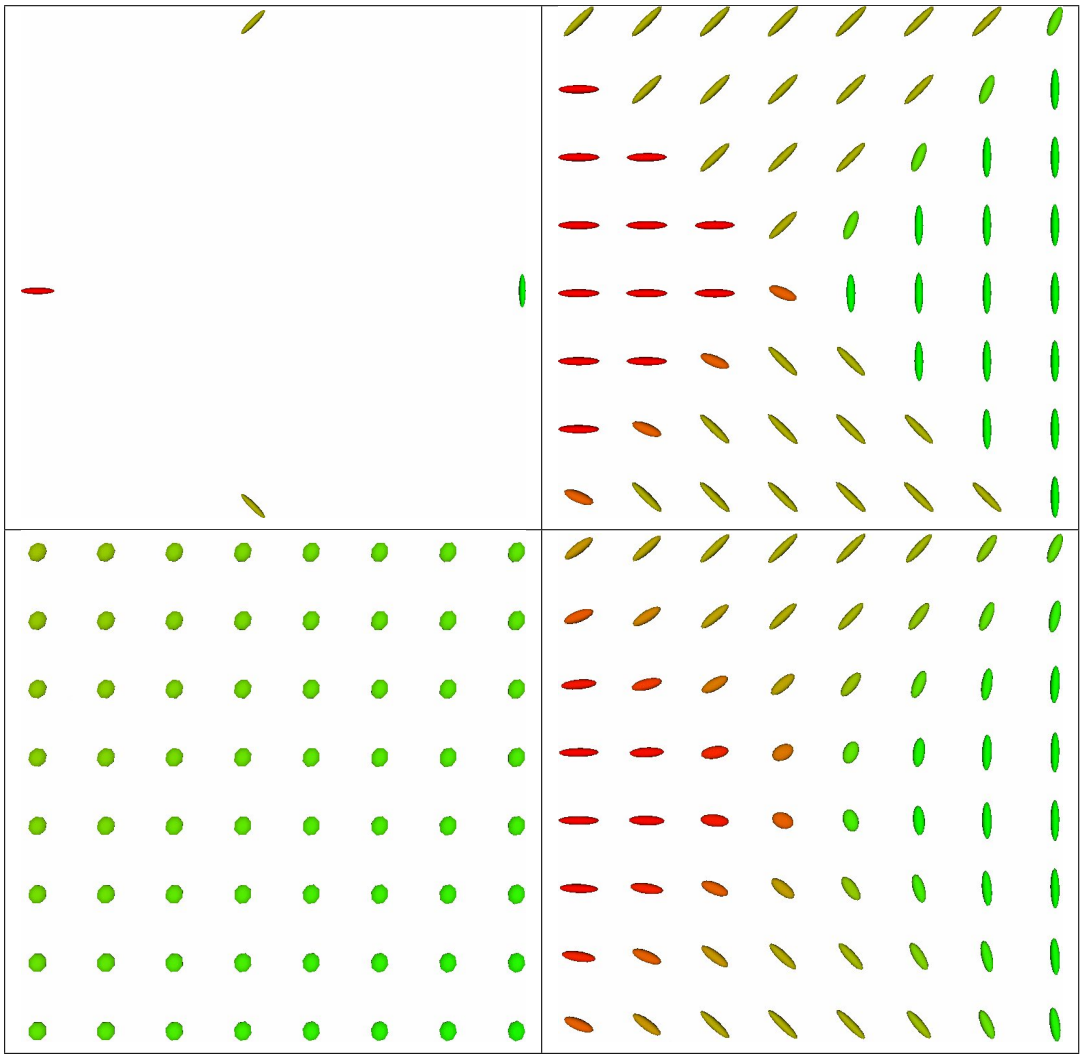


Figure 4.14: Interpolation and extrapolation of tensor values from four measurements using diffusion. **Top left:** The four initial tensor measurements. **Top right:** Initialization of the tensor field using a soft closest point interpolation (mean of the four tensors with a renormalized spatial Gaussian influence). **Bottom left:** result of the diffusion without the data attachment term (1000 iterations, time-step $\varepsilon = 1$, $\lambda = +\infty$). **Bottom right:** result of the diffusion with an attachment term after (1000 iterations, time-step $\varepsilon = 1$, $\lambda = 0.01$, $\sigma = 1$ pixel of the reconstruction grid). The algorithm did in fact converge in about 100 iterations.

the mathematical formulation of various processing like interpolation and Gaussian smoothing, but also PDE solving and statistical analysis. The most striking result is that in the case of DT-MRI, the AI and LE frameworks gave almost similar results! This is something that could be explained by the fact that tensors, in the context of diffusion MRI, are "relatively close" to the identity matrix. And we showed that the LE framework could be seen as the special case of the AI one when vectors were expressed in the tangent space at the identity. The main differences, though, reside in the fact that LE formulation of the processing exposed here is much simpler than with AI metrics, and that computations are much faster. Finally, LE metrics appear as a better candidate to replace the flawed Euclidean calculus. However, other frameworks with other properties are possible: the geodesic loxodromes by Kindlmann et al. [Kindlmann 2007] have the nice property to preserve not only the monotonic evolution of the trace and the determinant during a linear interpolation, but also the FA (which is not the case of our frameworks). However, it is not yet possible to use these loxodromes in a complete and simple framework as we do it here with LE metrics.

The next chapter is the general conclusion of this part, with a summary of the contributions and a discussion on the choice of the metric. This choice is important since the metric is responsible for giving the desired (or un-desired) properties of the result, and we wanted to emphasize this point.

Chapter 5

Conclusions and Discussion: Which Metric for Which Application?

Contents

5.1 Conclusions

We have presented two alternatives to the Euclidean calculus on tensors. Endowed with the affine-invariant metric, the space of positive define symmetric matrices takes a very regular manifold structure. In particular, tensors with null and infinite eigenvalues are both at an infinite distance of any positive definite symmetric matrix: the cone of tensors is replaced by a space which has an infinite development in each of its $n(n+1)/2$ directions. Moreover, there is one and only one geodesic joining any two tensors, and we can even define globally consistent orthonormal coordinate systems of the tangent spaces. Thus, the structure we obtain is very close to a vector space, except that the space is curved.

The second framework, the Log-Euclidean metric, relies on a simple principle: the vector space structure of symmetric matrices is transposed to the tensor space by the matrix exponential. This leads to a striking simple fact: all computations on tensors can be made on their logarithms as we are in a vector space, and mapped back to the tensor space by the matrix exponential. The price to pay for such straightforward principle is rather low: Log-Euclidean metrics are only similitude invariant, while our first family of metric is affine invariant. In fact, one can show that

both are strongly linked: Log-Euclidean metrics correspond to expressing all tensors in the vector space at the identity matrix within the affine-invariant framework.

A second contribution is the application of these frameworks to important geometric data processing problems such as interpolation, filtering, diffusion and restoration of tensor fields. We showed that interpolation and Gaussian filtering can be tackled efficiently through a weighted mean computation. However, if weights are easy to define for regularly sampled tensors (e.g., for linear interpolation), the problem becomes more difficult for irregularly sampled values. The solution we propose is to consider this type of interpolation as a statistical restoration problem where we want to retrieve a regular tensor field between (possibly noisy) measured tensor values at sparse points. This type of problem is usually solved using a PDE evolution equation. We showed that the usual linear regularization (minimizing the magnitude of the gradient) and some anisotropic diffusion schemes can be adapted to our Riemannian frameworks, provided that the metric of the tensor space is taken into account. We also provided intrinsic numerical schemes for the computation of the gradients and Laplace-Beltrami operators. Finally, simple statistical considerations led us to propose least-squares data attachment criteria for dense and sparsely distributed tensor fields. The differentiation of these criteria is particularly efficient thanks to the use of the Riemannian distance inherited from the chosen metric.

We confronted both families of metrics in practice, and showed that results were so close that it is difficult to differentiate them. In fact, differences appear only when the difference in anisotropy is high (e.g., when the largest eigenvalues are different by a factor of at least 10). This situation is hardly encountered when working with tensors as those obtained in DTI for example. However, the computational time can be significantly different. Processing tensors using Log-Euclidean metrics requires taking the matrix logarithm and exponential only once for each tensor (we could talk about constant time operations, i.e., operations that do not depend on the processing itself). Using the affine-invariant metric requires going back and forth between the tensor space and the tangent space at the current processing point, thus making use at each processing step of the matrix exponential, logarithm, inverse and square root. This computational load reduces the performance of the affine-invariant tensor processing when iteratively solving a PDE (like the anisotropic regularization of Sec. 4.3.4, or the tensor field extrapolation by diffusion of Sec. 4.4.3).

Log-Euclidean metrics seems to be more suited for DT-MRI processing: they give results similar to the affine-invariant family for a computational cost close to the Euclidean calculus. One could question, however, if the invariance of the result by a similarity transform is sufficient. For instance, Log-Euclidean smoothing of a tensor field and of the same field warped by a shearing will not give the same result (i.e., smoothing the warped field is not equivalent to warp the smoothed field), while affine-invariant smoothing will. A solution consists in performing a

few iterations of the affine-invariant smoothing after the Log-Euclidean version for example. In that case, the result would be affine-invariant and the time taken to obtain it would be close to the time of the Log-Euclidean processing.

In the next section, we discuss the choice of the metric w.r.t. the application, and show that the Riemannian frameworks developed here may not be the right choice in all situations where tensors are encountered.

5.2 Which Metric for Which Application?

In this discussion, we want to point out that the choice of the metric w.r.t. the application is crucial and should be made carefully. The question is not to figure out whether it is preferable to use the affine-invariant or the Log-Euclidean metric (we showed that both give generally similar results), but rather if these Riemannian metrics are suitable for all applications involving tensors. To illustrate this, let us take the example of another type of tensor: the structure tensor.

5.2.1 Case of the Structure Tensor

The structure tensor has become a useful tool for the analysis of features in images. It is used in edges and corners detection [Förstner 1987], texture analysis [Bigün 1991, Rao 1991], texture and color segmentation [de Luis-Garcìa 2005, de Luis-Garcìa 2008], filtering [Weickert 1999], and even medical image registration [Stefanescu 2005]. We will see in the following section how to calculate this tensor. The estimation of the structure tensor depends on the quality of images. If images are noisy, structure tensor fields will be noisy as well (although a scaling parameter, and consequently a smoothing, appears during estimation, at a fine scale the problem of the noise remains unchanged). One could try to anisotropically smooth the estimated tensor field (using our Riemannian metrics) and assess whether or not this type of smoothing is adapted to structure tensors, and can improve the detection and localization of corners for instance. This is what we investigate in the next sections.

5.2.2 Definition of the Structure Tensor

Let I be an image defined on a domain of $\mathbb{R}^d$. The structure tensor is based on the gradient of I: $\nabla I = (\partial_1 I, \ldots, \partial_d)^T$, where each directional derivative $\partial_i I$ can be computed with a finite difference scheme or by filtering with a first order derivative of a Gaussian. The structure tensor S_σ can be defined as:

$$S_\sigma = G_\sigma * \left(\nabla I \nabla I^T\right)$$

with G_σ being a Gaussian of standard deviation σ. The variance σ controls the smoothness of the resulting tensor field. The noisier the image is, the higher σ must be to obtain a smooth field, but smaller structures may be wiped out. By contrast,

smaller values of σ can help to extract low level features in images, but the resulting structure tensor image may be noisy. Consequently, one would like to perform an anisotropic filtering of the structure tensor field obtained with a low σ, in order to regularize homogeneous regions while preserving the boundaries with low-level features. In the following, we first compute the Riemannian gradient of a structure tensor image and compare it to the classical Euclidean gradient. Then, we perform an anisotropic filtering and discuss the results.

5.2.3 Gradient of a Structure Tensor Image

Following the numerical scheme of Sec. 4.3.2, we computed the gradient norm of a structure tensor image obtained with a σ of 1.0 (Fig. 5.1 left is the original image). Then, we compared it to the Euclidean gradient. We also added noise in the original image to evaluate the robustness of both gradients. Results of comparisons are shown in Fig. 5.1.

First, we can notice that with the affine-invariant metric, outliers appear in the image background (Fig. 5.1 top right). This is intensified when adding noise (Fig. 5.1 bottom right): we see that the background is made with artefacts due to variations of small tensors that result from noise. Indeed, small tensors have as much importance as large ones because of the affine- or similitude-invariance of the metrics. Consequently, the Riemannian gradient of a variation of small tensors or larger ones will be identical. By contrast, the Euclidean gradient remains much less sensitive to tensors with small coefficients, and consequently only the main features are revealed (Fig. 5.1 bottom right and left).

Second, details that are not present in the Euclidean norm appear in the Riemannian gradient: this is the case, for example, of low-contrasted edges in the original image. This also results from the affine- or similitude-invariance of the metrics.

To summarize, the Riemannian frameworks can reveal lower structural information such as low contrasted edges but are highly sensitive to small variations in the tensor image, and thus suffer from a lack of robustness.

Let us now investigate how the anisotropic filtering scheme can restore the noisy structure tensor image.

5.2.4 Anisotropic Filtering of a Structure Tensor Image

We applied the anisotropic filtering scheme of Sec. 4.3.4 on the noisy structure tensor image of Fig. 5.1 bottom left. We used the following parameters: $\kappa = 0.02$, $\varepsilon = 0.1$ and 500 iterations (total diffusion time: 50). Results are presented in Fig. 5.2.

The Riemannian metrics cause the variations of small tensors to be highly contrasted in the norm of the Riemannian gradient and thus to be preserved during the filtering process. Figures 5.2 bottom illustrate this behavior: the top of the original noisy image (middle image) is filled with artefacts that are preserved during smoothing (right image). Taking the Euclidean gradient

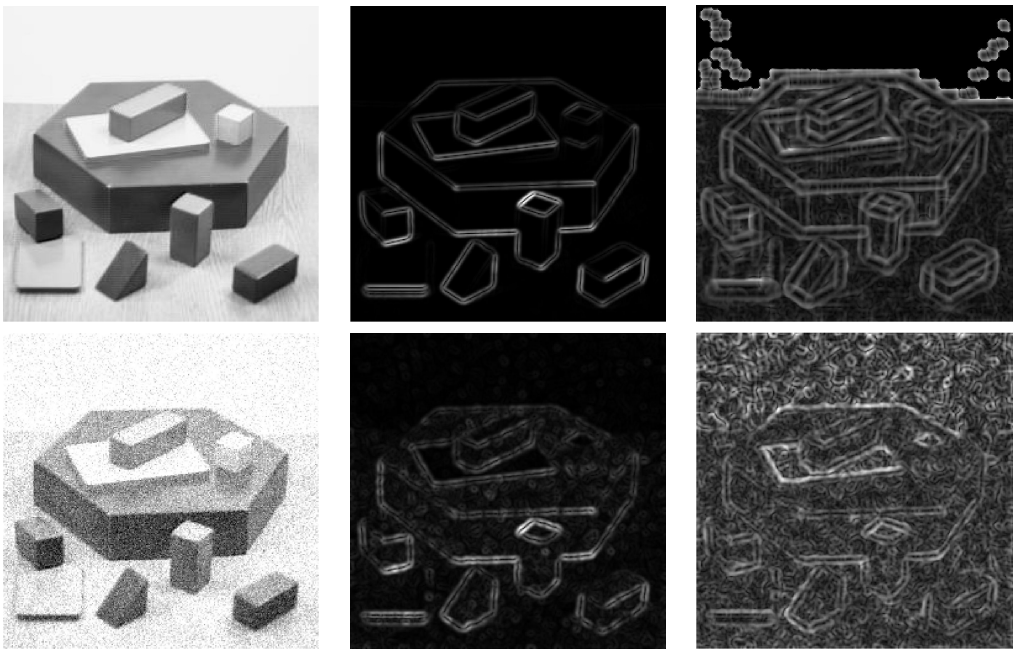


Figure 5.1: **Euclidean gradient versus Riemannian gradient. Top row:** Original image (left), norm of the Euclidean gradient (middle) and Riemannian gradient (right). **Bottom row:** Noisy image (Gaussian noise of variance 0.01 was added) (left), norm of the Euclidean gradient (middle) and Riemannian gradient (right).

(Fig. 5.2 top right) removes the artefacts in the image background. Homogeneous regions are smoother while edges are correctly conserved. However, some artefacts in the background that were expected to disappear are still present and the intensity of most of the relevant edges is lower than in the Riemannian case.

In conclusion, the Riemannian metrics that are well suited for DTI, or more generally for covariance matrices, seems not to be applicable directly to structure tensor images. In fact, the affine and similitude invariance give an identical role to small tensors versus larger ones. Thus, while this allows to extract low-level features, it also suffers from a lack of robustness.

This example illustrates the complexity of choosing the correct metric for a given application. The choice for the Riemannian metrics we made is obviously not the best solution in all situations. We have just shown that they are good candidates to replace the Euclidean calculus when working with diffusion tensors, and later in this manuscript we will show that they can be applied to covariance matrices as well. The most important point is that the choice of the metric is crucial and determines the key functions (exponential and logarithmic maps to name them) of the Riemannian framework. It is important to notice that this framework does not depend on

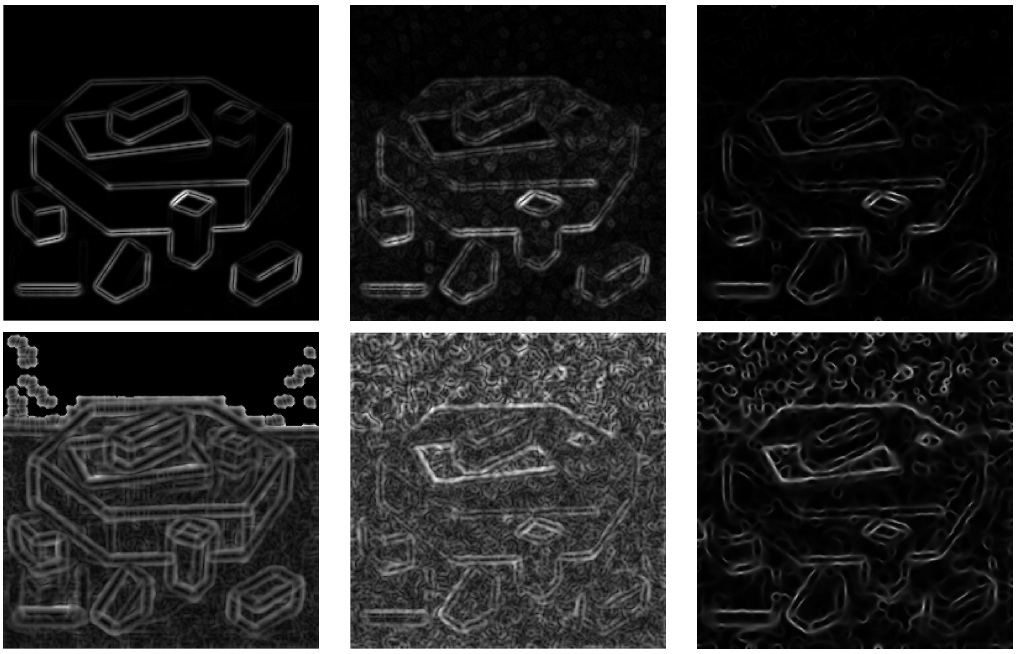


Figure 5.2: **Anisotropic filtering of a noisy structure tensor image.** Diffusion parameters: $\kappa = 0.02$, $\varepsilon = 0.1$ and 500 iterations. **Top row:** From left to right: Original Euclidean gradient, Euclidean gradient with noise, Euclidean gradient after regularization. **Bottom row:** From left to right: Original Riemannian gradient, Riemannian gradient with noise, Riemannian gradient after regularization.

the choice of the metric: basic operations (addition, subtraction), algorithms, etc. remain the same, and only the exponential and logarithmic maps are changing. In other words, the Riemannian framework is a general toolbox for working on manifolds, and only the exp and log functions vary depending on the metric. The metric itself should be carefully chosen depending on the desired properties for the result (properties of invariance for instance). One last remark is that the affine-invariant and Log-Euclidean metrics described here hopefully result in simple expressions for these two key functions, which is generally not the case.

Part II

An Optimal Workflow for Using DT-MRI in Clinical Applications

"Part of the inhumanity of the computer is that, once it is competently programmed and working smoothly, it is completely honest."
Isaac Asimov.

Diffusion tensor MRI has rapidly become popular in clinical research as it is a unique tool to assess in vivo oriented structures within tissues. In particular, brain DTI allows to access the white matter neuronal architecture and is by far the most frequently encountered type of diffusion MRI in clinical applications. The goals of using such brain imaging modality are numerous: one would like for instance to assess whether a tumour is bending or destroying a fiber bundle, or as a preliminary step to surgery, one may want to make sure that no vital fiber tract will be injured during the surgical procedure. The scope of using DT-MRI in clinical applications seems extraordinary promising and brings hope for improving several aspects of brain surgery, like a shorter surgery time, a better comfort for the patient, and an increased survival rate with less side effects. However, data and methods are not fully ready yet to allow such applications. Acquisitions generally have a limited number of encoding gradients and low signal-to-noise ratios (SNR). Indeed, the scanning time is rather short - 15 minutes for most of the French hospitals. The situation may be different in other countries, but the rather limited number of MRI scanners available in France implies that no more than 15 minutes should be spent to scan a patient, and the fact that most of them are 1.5T scanners leads to rather low SNR images. Another reason to shorten the scan time is that scanning itself can be perceived or experienced as innocuous for some patients, but not for others: the patient should not move (or the data will not be valid), the scanner entrance is small (claustrophobia becomes an issue), and scanning itself can be very noisy, which makes the overall experience rather uncomfortable. One solution to improve image quality while keeping the same scanning time would be to switch to higher magnets (3T scanners), which is out of question in general (for the time being) for obvious financial reasons. Consequently, this short scanning time prevents acquiring and averaging the large number of images that is necessary for enhancing SNR or using more elaborate models of the diffusion phenomenon like CHARMED [Assaf 2005], Q-ball and orientation distribution functions (ODF) [Tuch 2004, Tournier 2004b, Descoteaux 2006, Descoteaux 2007a], or higher order tensors (HOT) [Barmpoutis 2007a], which could describe with more accuracy the white matter architecture, especially in regions with crossing fibers. Typical clinical DWI acquisitions are acquired using 6 gradient directions with 4 repeated scans, or 25 gradient directions without repetition.

To be able to work with diffusion tensor images, and by extension, reconstruct white matter fibers in a clinical environment, one needs to adapt methods and software to the clinical constraints and clinician needs. First of all, the estimation of the diffusion tensor from diffusion weighted images should take into account the fact that only a few gradients are used and that images have a rather low SNR. This will be the topic of Chapter 6, where we propose a joint diffusion tensor estimation

and smoothing procedure, which takes into account the real nature of the MRI noise. Then, we switch to software issues. In order to be integrated in a clinical environment, a method should be embedded into a stable, efficient and user-friendly software. Indeed, medical experts cannot spend hours learning a new software, and any tool must be intuitive and reactive to have a chance to be used in clinics. For this purpose, we introduce in Chapter 7 MedINRIA, a software package for DT-MRI processing targetting explicitly the medical experts. Finally, we demonstrate in Chapter 8 the feasibility of DTI in a clinical environment with a study on the possible applications of diffusion tensor imaging in spinal injuries.

Chapter 6

Joint Estimation and Smoothing of Clinical DT-MRI with Log-Euclidean Metrics

Contents

Diffusion tensor estimation is a necessary step prior to working with a diffusion tensor field. Indeed, the MRI scanner does not produce tensor images directly, but provides us with a diffusion weighted MRI data: It is generally composed with one baseline image, called the S_0 (or $b = 0$ image) (which is nothing else than a T_2 weighted image), and at least six diffusion weighted images. Each of these diffusion weighed images corresponds to a diffusion gradient, i.e., a measure of the water diffusion phenomenon in a spatial direction. The diffusion tensor relates the signal loss observed in the DWI compared to the baseline image via the diffusion equation of Stejskal-Tanner [Basser 1994b]:

$$S_i = S_0 \exp(-b\mathbf{g_i}^\top \mathbf{D}\mathbf{g_i}), \tag{6.1}$$

where S_i is the DWI corresponding to the spatial gradient g_i, S_0 is the baseline image (b_0 image), D the diffusion tensor and b the $b - value$, a factor related to

the acquisition parameters and diffusion time (generally $b = 1000\ s/mm^2$). The tensor estimation consists in retrieving D from the measures (S_i). As MR images are subject to noise, the estimation of the diffusion tensor field from the MR measurements is noise-sensitive. We give below a quick overview of the literature on diffusion tensor estimation.

6.1 Introduction

6.1.1 How Does DTI Estimation Work?

What is generally done for tensor estimation is a linearization of Eq. 6.1 to obtain a linear system [Basser 1994b, Westin 2002, Tschumperlé 2003, Deriche 2004]:

$$\log(S_i) = \log(S_0) - b g_i^\top D g_i, \forall i \in [1, N]$$

N being the total number of acquisition (i.e., encoding gradients). Solving this system in a least square sense leads to the minimization of a quadratic criterion with algebraic methods and has an analytical solution:

$$\tilde{D} = \arg\min_D \sum_{i=1}^{N} \left[\log\left(\frac{S_i}{S_0}\right) + b g_i^\top D g_i \right]^2$$

From a pure signal processing point of view, this estimation corresponds to a maximum likelihood (ML) estimator with a model of a Gaussian noise on the image logarithms:

$$\log(\hat{S}_i) = \log(S_0) - b \mathbf{g_i}^\top \mathbf{D} \mathbf{g_i} + N_i(0, \sigma_{lg}), \tag{6.2}$$

where $\hat{S}_i$ is the measured DWI intensity, and $N_i(0, \sigma_{lg})$ is a centered Gaussian noise of variance σ_{lg}. We call this noise model log-Gaussian in the following. Naturally, one may wonder if this assumption correctly reflects the noise appearing in real DW images. In fact, when the SNR is high, one can show that the noise on the image logarithms is indeed well approximated by a Gaussian distribution [Salavador 2005], which justifies the linearization of Eq. 6.1. In the same conditions, the noise can also be well approximated by a Gaussian distribution within the brain on the image directly (not their logarithm), as in [Chang 2005]:

$$\hat{S}_i = S_i + N_i(0, \sigma_g). \tag{6.3}$$

Note that in [Chang 2005], the authors also use a strategy to automatically detect and reject image outliers (i.e., bad diffusion weighted images) to make the estimation more robust to a bad acquisition, which is not applicable, for instance, when working with the minimum number or images (6).

For low SNR images typical of clinical acquisitions, one needs to go back to the true nature of the MRI noise, which is Rician: it corresponds to a Gaussian noise on both real and imaginary parts of the complex MR signal. The measured intensity

is the magnitude of this signal. In other words, one can model the measured signal as [Sijbers 1998]:

$$\hat{S}_i = \sqrt{[S_i + N_{re}(0, \sigma_r)]^2 + N_{im}(0, \sigma_r)^2}, \tag{6.4}$$

where N_{re} and N_{im} are two independent Gaussian noises of the same variance. This is equivalent to adding Gaussian noise to the k-space data before computing the signal magnitude. Wang et al. [Wang 2004a] proposed an estimation criterion on the complex DWI signal that is adapted to this type of noise. However, we generally cannot access the complex signal but only its magnitude (although it could be accessed if required), and we will show in the following that we actually don't need it.

There is also a major drawback to using such estimation procedures: nothing ensures the resulting tensor to be a positive definite symmetric matrix! By construction, the result will be a symmetric matrix, but because of noise and possibly uncorrected distorsions, it can eventually lead to tensors with null or negative eigenvalues, which are physically meaningless and should be avoided at any cost.

In addition to tensor estimation, and in view of fiber reconstruction, one could think of improving the spatial regularity of the diffusion tensor field in order to remove the maximum amount of noise we can. A naive tractography algorithm, which consists in integrating the principal direction of diffusion (PDD) field (major tensor eigenvector), is highly noise-sensitive and can easily produce erroneous fiber trajectories. Increasing the spatial regularity might help recover longer and smoother fibers. One needs to be careful, however, because the diffusion tensor field should be regularized without blurring the transitions between distinct fiber tracts, which delimit anatomical and functional brain regions. We quickly summarize below the state-of-the-art in diffusion tensor field smoothing.

6.1.2 Diffusion Tensor Fields Regularization

A first idea consists in smoothing independently each DWI, as done for instance very recently in [Basu 2006] with a Rician noise model. This results in a smoother tensor field that preserves some of the transitions. However, it also blurs the transitions between some regions as the boundaries of the tensor field are not visible in each DWI taken separately. For instance, in brain DTI, only the combination of all the DW images reveals the complex neural structure of the white matter. Consequently, we believe that it is better to detect the transitions on the tensor field itself. Moreover, working on the DWI directly can be tricky as we don't know what would be the consequences of our processing on the tensor field after estimation. Working on tensors, and within a framework that allows us to control the result (i.e., be sure that whatever we do we still get a tensor at the end, and that we preserve some properties, like invariance) is a much more satisfactory approach.

Some earlier regularization methods have also been proposed as a post-processing step after the estimation of the tensor field. For instance, [Coulon 2004] regularizes the principal eigenvector (associated to the largest eigenvalue), while [Chefd'hotel 2004] uses the spectral decomposition of tensors to independently regularize their eigenvectors and eigenvalues. These two methods rely on the spectral decomposition of tensors: considering only the principal eigenvector for smoothing induces a loss of information and an uncertainty of the spectral decomposition in regions with flat tensors (in this case, what is the major eigenvector?). Moreover, discontinuity problems arise when smoothing the field of orthogonal matrices as the spectral decomposition is not unique. Using the Riemannian metrics described in Chapter 3 (affine-invariant or Log-Euclidean) allows us to generalize anisotropic regularization scheme like those of Perona-Malik [Perona 1990] or Gerig et al. [Gerig 1992] directly to tensors (Sec. 4.3.4). This last approach is more satisfactory because it acts on the tensors directly, thus preserving the discontinuities of the tensor field which is a desirable feature.

Furthermore, it would be more interesting to consider the regularization as a spatial prior on the tensor field during the estimation step itself. This would allow to optimally weight the information brought by the observed images and the expected spatial regularity. Such a Maximum A Posteriori (MAP) estimation should extract the maximal amount of meaningful information from very noisy clinical DWI data. In that spirit, Wang et al. [Wang 2004a] proposed to parameterize the space of tensors by the vector space of lower triangular matrices thanks to the Cholesky decomposition. This provided them with a computational framework that could handle a joint estimation and regularization of the tensor field from the complex DWI signal. It raises the question of the framework used for solving such non-linear estimation criteria. Indeed, one cannot generally find an analytical solution when using a noise model different from the log-Gaussian (i.e., solving a linear criterion). Moreover, one needs a unified framework that would allow us to perform both tensor estimation and smoothing at the same time. Add-hoc methods like smoothing only the PDD field, or processing independently eigenvectors and eigenvalues do not allow this. By contrast, the Cholesky decomposition and the Log-Euclidean metrics are two possible frameworks for solving this problem of joint estimation and smoothing of DTI.

To illustrate that point, we investigate and compare in Section 6.1.3 different tools for computing with tensors: the Euclidean metric, the Cholesky factorization and the Log-Euclidean metrics. It turns out that LE metrics are computationally as efficient as the Cholesky decomposition while ensuring the positive definiteness of tensors and canceling the swelling effect in regularization.

Based on the LE computational framework, we detail in Section 6.2 an original variational method for a joint estimation and smoothing of DTI. We first derive the ML estimation of tensors with the log-Gaussian (Eq. 6.2), Gaussian (6.3) and Rician (Eq. 6.4) noise models. We show that Rician noise induces a *shrinking effect* when

other noise models are used for tensor estimation. By adding an anisotropic spatial prior in a second step, we turn these three ML into three MAP methods where the estimation and the regularization of the tensor field are jointly performed.

A quantitative analysis of our six new methods on synthetic data shows in Sec. 6.3 that the Rician ML method correctly handles the shrinking effect even with a very low SNR, while other methods under-estimate the tensor volume by as much as 40%. Then, we switch to experiments with real clinical data on a medium quality brain DTI dataset (tumor case with only 6 gradient directions) and a low quality experimental acquisition of the spinal cord (same sequence). The visual inspection of the reconstructed tensor field shows that the MAP method nicely preserves the separation between different regions. A careful analysis in two specific regions shows that ML and MAP Rician methods exhibit a larger tensor volume and ADC than Gaussian and log-Gaussian methods. Lastly, we illustrate that the MAP Rician method qualitatively improves the fiber tracking.

6.1.3 Tools for Tensor Computing

This section is a brief summary of Chapter 3. Tensor computing is difficult due to the severe limitations of the standard Euclidean calculus: while convex operations are stable (e.g. the mean of a set of tensors is a tensor), one can quickly reach the boundaries of the space with complex operations (e.g., gradient descent or partial differential equations) and null or negative eigenvalues may appear. To overcome this limitation, [Wang 2004a] proposed to parameterize a tensor $\mathbf{D}$ by its Cholesky factors. A Cholesky decomposition of $\mathbf{D}$ is given by: $\mathbf{D} = \mathbf{L}\mathbf{L}^\top$, where $\mathbf{L}$ is a lower triangular matrix. A lower triangular matrix $\mathbf{M}$ is such that $\forall(i,j)$, if $i > j$ then $M(i,j) = 0$ (all terms above the diagonal are null). For any lower triangular matrix $\mathbf{L}$, the matrix $\mathbf{L}\mathbf{L}^\top$ is positive (this is easily shown using SVD). However, the definiteness property is not ensured (null eigenvalues are possible). The authors argue that forbidding negative eigenvalues is sufficient because one cannot numerically distinguish very small eigenvalues from null ones. Thus, forbidding explicitly null eigenvalues has no practical justification.

While we agree that very small eigenvalues are not distinguishable from null ones, we believe that both are very unlikely to exist from a physical point of view (this would mean an almost null diffusion in one direction of space, which is physically not possible in the conditions where the measures are made), and should be as far as possible from any reference tensor (even if they are numerically close). In other words, a tensor with very small eigenvalues has a very low probability to appear, as well as a tensor with very large eigenvalues. Both of them must be numerically nearly impossible to reach.

As described in Chap. 3, the Riemannian metrics offer a solution to this constraint: endowed with such metric, the tensor space is replaced by a regular manifold where matrices with null and negative eigenvalues are at an infinite distance from any tensor. Among these, Log-Euclidean metrics turn out to have a computational

cost similar to the Euclidean case (Sec. 3.4). The key idea is to take the matrix logarithm of a tensor $\mathbf{D}$: $\mathbf{L} = \log(\mathbf{D})$, and to run computations on $\mathbf{L}$. The new processed value $\tilde{\mathbf{L}}$ obtained is turned back into a tensor by taking the matrix exponential: $\tilde{\mathbf{D}} = \exp(\tilde{\mathbf{L}})$. We showed in 3.4.2 that it yields excellent theoretical properties, such as the monotone interpolation of the determinants, and the prevention of the swelling effect.

Advantages of the Log-Euclidean Framework:

To compare the LE metric, the standard Euclidean framework and the more elaborate Cholesky decomposition, we believe that understanding the structure of the tensor space endowed with each metric is important. Let us illustrate this by realizing a geodesic shooting in the tensor space endowed alternately with an Euclidean, Cholesky and Log-Euclidean structure. It consists in following the geodesic starting at one given tensor in a given direction during a certain time. It is exactly the same as a car race: we give a starting position, and an initial speed (initial speed vector), and we look at the trajectory that the car takes. As the tensor space is embedded in the space of symmetric matrices, a tangent vector is simply a symmetric matrix. In the Euclidean case, one computes: $\mathbf{D_e}(t) = \mathbf{D} + t\dot{\mathbf{D}}_\mathbf{e}$. In Figure 6.1, we chose a diagonal tensor $\mathbf{D} = ((4,0,0),(0,1,0),(0,0,1))$, a diagonal tangent vector (a symmetric matrix belonging to the tangent plane at the considered tensor) $\dot{\mathbf{D}}_\mathbf{e} = ((-8,0,0),(0,-2,0),(0,0,-2))$ and $t \in [0,2]$. A "geodesic" with the Cholesky factors is given by a straight line in the space of lower triangular matrices: $\mathbf{L(t)} = \mathbf{L} + t\dot{\mathbf{L}}$. The resulting tensor curve is simply $\mathbf{D_c}(t) = \mathbf{L(t)L(t)}^\top$. To start at the same point with the same tangent vector as in the Euclidean case, $\mathbf{L}$ is the Cholesky factor of $\mathbf{D}$ and $\dot{\mathbf{L}}$ is solution of $\dot{\mathbf{L}}\,\mathbf{L}^\top + \mathbf{L}\,\dot{\mathbf{L}}^\top = \dot{\mathbf{D}}_\mathbf{e}$. In our case, we obtain $\dot{\mathbf{L}} = ((-2,0,0),(0,-1,0),(0,0,-1))$. With the LE metric, the geodesic is finally: $\mathbf{D_{log}} = \exp(\log(\mathbf{D}) + t\dot{\mathbf{D}}_\mathbf{log})$, where tangent vectors are given in the diagonal case by $\dot{\mathbf{D}}_\mathbf{log} = \exp(\mathbf{D})^{-1}\dot{\mathbf{D}}_\mathbf{e}$. Results of the geodesic shootings are displayed in Fig. 6.1.

In the Euclidean case, as expected, one quickly reaches the boundaries and non-displayed values are actually non-positive tensors. The Cholesky case is algorithmically well posed as non-positive matrices do not appear. However, we still reach zero eigenvalues, i.e. the null matrix appears, and values beyond it are the mirrored versions of the first ones. This means that the null matrix is reached on the trajectory during a gradient descent. Moreover, one may question the physical meaning of the mirrored values obtained beyond the null matrix. In the LE case, the null matrix is never reached and all tensors are by nature positive definite.

As a second advantage, the Log-Euclidean framework completely overcomes the *swelling effect* which can be observed in both Euclidean and Cholesky cases, and is illustrated in Fig. 6.2 with the example of the regularization. This effect causes tensors to grow after a processing. We generated a synthetic noisy tensor field and applied the anisotropic regularization of Sec. 6.2.4 with 3 different ways to process tensors. First, we used the Euclidean calculus, second we computed on the Cholesky factors and third we used a LE metric. Both Euclidean and Cholesky

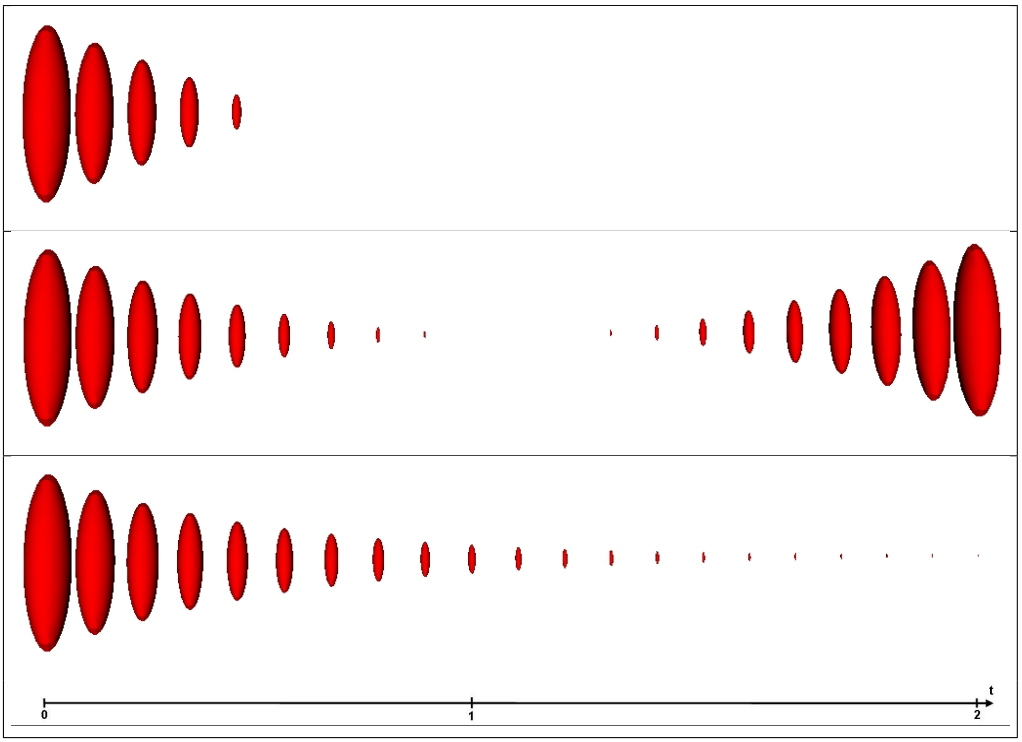


Figure 6.1: **Geodesic shootings simulating a gradient descent. The x axis is the time t travelled along the geodesic ($\mathbf{t} \in [\mathbf{0}, \mathbf{2}]$). Top:** The Euclidean case. 2/3 of tensors are not positive definite and thus are not displayed. **Middle:** The Cholesky case: the null matrix is reached (exact middle value) and values beyond are the mirrored versions of the previous ones. **Bottom:** The Log-Euclidean case: all tensors are positive definite, and the null tensor is never reached.

regularization suffer from the swelling effect, i.e. the denoised tensors are larger than the original values, the effect being less pronounced in the Cholesky case. With the LE framework, the swelling effect vanishes and the tensors are correctly denoised.

To quantify the benefit of the Log-Euclidean framework, we computed the root mean square error (RMSE) between the restored and original fields. Not to influence one particular metric, we computed the RMSE for the three metrics: Euclidean, Cholesky and LE. Results are summarized in Table 6.1. Whatever the metric, the LE framework gives results quantitatively better than the 2 other frameworks.

Table 6.1: **RMSE between the restored and original tensor fields.** For the three given metrics, the Log-Euclidean regularization produces results the closest to the original data.

	Euc. Reg.	Chol. Reg.	LE Reg.
Euclidean RMSE	0.228	0.172	0.051
Cholesky RMSE	0.152	0.092	0.015
Log-Euclidean RMSE	0.532	0.313	0.111

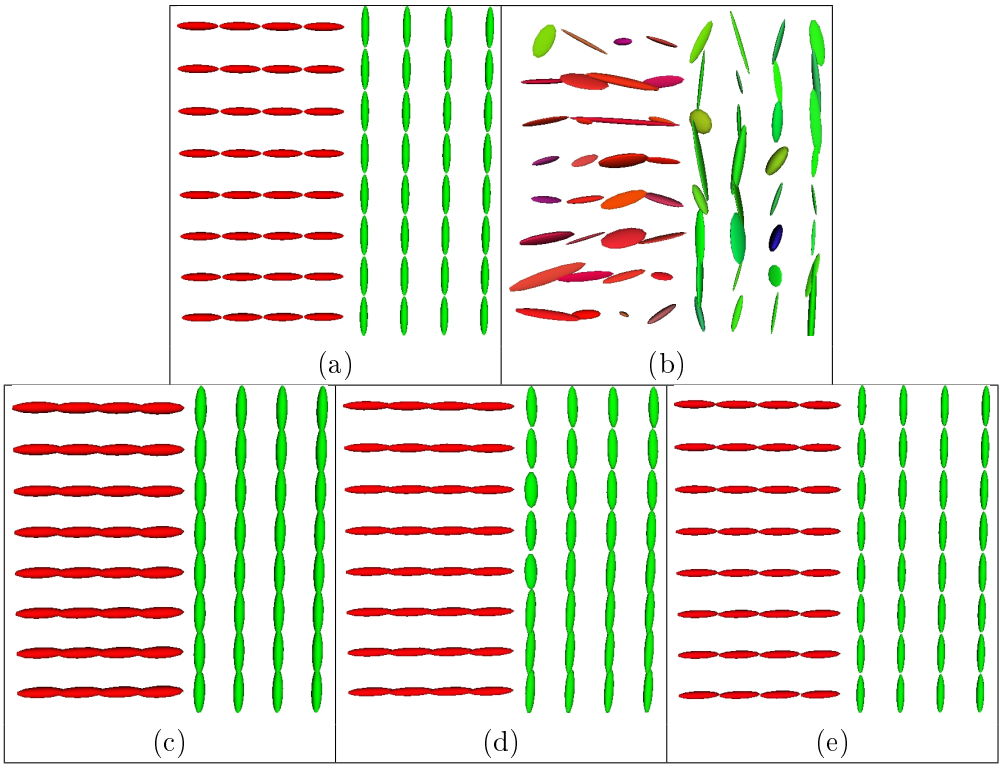


Figure 6.2: **Anisotropic regularization of a noisy tensor field. (a):** Original synthetic field. **(b):** Noisy field. **(c):** Euclidean regularization. **(d):** Cholesky regularization. **(e):** Log-Euclidean regularization. Note the swelling effect in the Euclidean and Cholesky cases.

To conclude, the Log-Euclidean framework is well adapted to the processing of diffusion tensors: it not only overcomes the limitations of the Euclidean calculus (negative eigenvalues), but it also removes the swelling effect which can be observed in both Euclidean and Cholesky frameworks. Consequently, we choose this family of metrics to solve our problem of joint estimation and smoothing of clinical DTI.

6.2 A Variational Formulation with Three Noise Models

The joint estimation and regularization of DTI can be tackled by a variational formulation, i.e. one has to minimize the energy functional:

$$E(\mathbf{L}) = \frac{1}{2}\text{Sim}(\mathbf{L}) + \frac{\lambda}{2}\text{Reg}(\mathbf{L}), \tag{6.5}$$

with Sim(.) being the data attachment term (estimation) and Reg(.) being the regularization term. In a statistical setting, Sim(.) is usually the log-likelihood of the measurements knowing the parameters, while Reg(.) is the prior knowledge on the parameters. λ is a normalization factor between the two terms. To fully make

use of the advantages of the LE framework, we directly parameterize the diffusion tensor by its logarithm $\mathbf{L} = \log(\mathbf{D})$.

6.2.1 Log-Gaussian Noise

The linearized version of the Stejskal and Tanner diffusion equation system gives us the following energy to minimize, in the least-squares sense:

$$\mathrm{Sim}_{log}(\mathbf{L}) = \int_\Omega \sum_{i=1}^{N} \left(\log\left(\frac{S_0}{\hat{S}_i}\right) - b\mathbf{g_i}^\top \exp(\mathbf{L})\mathbf{g_i} \right)^2, \tag{6.6}$$

where N is the number of encoding gradients. Of course, if we had parametrized this objective function with the unknown tensor $\mathbf{D}$ directly, we could have derived an analytical solution directly as in [Westin 2002]. However, by using the Log-Euclidean framework, we are adding a constraint to the estimation: the result must be positive definite! Indeed, we will optimize the criterion w.r.t. the tensor logarithm, and exponentiate the result to obtain a tensor. In order to minimize the criterion of Eq. 6.6, we need to differentiate it. The differentiation is easy to perform in the LE framework and gives:

$$\nabla \mathrm{Sim}_{log}(\mathbf{L}) = -2 \left(\log\left(\frac{S_0}{S_i}\right) - b\mathbf{g_i}^\top \exp(\mathbf{L})\mathbf{g_i} \right) \partial_{\mathbf{G_i}} \exp(\mathbf{L}), \tag{6.7}$$

with $\mathbf{G_i} = \mathbf{g_i}\mathbf{g_i}^\top$, and $\partial_{G_i} \exp(L)$ is the directional derivative of the matrix exponential in direction G_i (see Appendix C for a practical implementation of it). Finally, the minimization is achieved through a simple first order gradient descent: $\mathbf{L_{t+1}} = \mathbf{L_t} - dt \nabla \mathrm{Sim}_{log}(\mathbf{L_t})$. $\mathbf{L}$ belongs to a vector space and this evolution equation is actually a geodesic marching. After convergence, one simply needs to exponentiate the vector $\mathbf{L}$ to obtain a tensor: $\mathbf{D} = \exp(\mathbf{L})$.

6.2.2 Gaussian Noise

Assuming a Gaussian noise on the original DWI intensities, the ML estimation boils down to a least-squares estimation of the tensor field from the images themselves rather than from their logarithm versions. For more clarity, we denote the predicted DWI intensity by $S_i(\mathbf{L}) = S_0 \exp(-b\mathbf{g_i} \exp(\mathbf{L})\mathbf{g_i}^\top)$. This gives the following criterion [Chang 2005]:

$$\mathrm{Sim}_{gaussian}(\mathbf{L}) = \sum_{i=1}^{N} \int_\Omega (S_i(\mathbf{L}) - \hat{S}_i)^2. \tag{6.8}$$

The differentiation of Eq. [6.8] gives:

$$\nabla \mathrm{Sim}_{gaussian}(\mathbf{L}) = -2b \sum_{i=1}^{N} \left(S_i(\mathbf{L}) - \hat{S}_i \right) S_i(\mathbf{L}) \, \partial_{\mathbf{G_i}} \exp(\mathbf{L}). \tag{6.9}$$

Similarly to Eq. [6.6], the minimization is achieved through a first order gradient descent.

6.2.3 Rician Noise

In this section, we consider that the noise in MR images is Rician, i.e. the measured magnitude of the DWI can be modeled as:

$$\hat{\mathbf{S}}_\mathbf{i} = \sqrt{[\mathbf{S_i} + \mathbf{N_{re}}(\mathbf{0}, \sigma)]^\mathbf{2} + \mathbf{N_{im}}(\mathbf{0}, \sigma)^\mathbf{2}}$$

where $N_{re}(0, \sigma)$ and $N_{im}(0, \sigma)$ are independent centered Gaussian noises of variance σ acting respectively on the real and imaginary part of the signal. The square magnitude of the observation $\hat{S}_i$ is:

$$\hat{S}_i^2 = (S_i + N_{re}(0, \sigma))^2 + N_{im}(0, \sigma)^2.$$

Taking the mean of the last expression gives:

$$\begin{aligned} E\left[\hat{S}_i^2\right] &= E\left[(S_i + N_{re}(0, \sigma))^2\right] + E\left[N_{im}(0, \sigma)^2\right] \\ &= E\left[S_i^2\right] + 2\sigma^2. \end{aligned}$$

Therefore, the observed square magnitude $\hat{S}_i^2$ is a non-central chi-squared random variable. In this case, the DWI signal (not the squared version) is shifted by approximatively $\sigma^2/(2S_i)$ [Sijbers 1998]. This means that the Rician noise induces a shrinking effect of tensors: The DWI signal tends to be greater than it should be, and the resulting tensors tend to be smaller than they actually are (a higher signal means a lower diffusion). This effect is even more obvious when the SNR is low. To correct for this shrinking effect, we propose the ML estimator for the Rician noise.

For a Rician noise of variance σ^2 on the data, the pdf of the measured signal $\hat{S}$ knowing the expected signal S is [Sijbers 1998]:

$$p(\hat{S}|S) = \frac{\hat{S}}{\sigma^2} \exp\left(-\frac{\hat{S}^2 + S^2}{2\sigma^2}\right) I_0\left(\frac{S\hat{S}}{\sigma^2}\right), \tag{6.10}$$

where I_0 is the modified 0^{th} order Bessel function of the first kind. The ML estimator for the pdf of Eq. [6.10] is:

$$\mathrm{Sim}_{Rician}(\mathbf{L}) = -\sum_{i=1}^{N} \log\left(p\left(\hat{S}_i|S_i\left(\mathbf{L}\right)\right)\right). \tag{6.11}$$

The differentiation of Eq. [6.11] gives:

$$\nabla \mathrm{Sim}_{Rician}(\mathbf{L}) = -1/\sigma^2 \sum_{i=1}^{N} (S_i(\mathbf{L}) - \alpha\hat{S}_i) S_i\left(\mathbf{L}\right) \partial_{\mathbf{G_i}} \exp(\mathbf{L}), \tag{6.12}$$

with $\alpha = I_0'/I_0(\hat{S}_i S_i/\sigma^2)$ (see Appendix D for a practical evaluation of α). The formula is similar to the gradient of Eq. [6.9], except that a correcting factor α depending on the signal and the noise variance appears. A simple estimator of the noise variance is based on the following. Typical MRIs include regions outside of the patient. Considering the fact that the square magnitude of such regions is null, taking its mean gives us an estimation of $2\sigma^2$.

6.2.4 An Anisotropic Regularization Term

Let us now look at the spatial prior on tensors: We expect the tensor field to vary spatially slowly within homogeneous regions (where the spatial gradient is small), while it can drastically change at the boundaries of these regions. The log-probability of such a prior can be efficiently represented by the ϕ-functionals usually used for anisotropic regularization: $\mathrm{Reg}(\mathbf{L}) = \int_\Omega \phi\left(\|\nabla \mathbf{L}\|\right)$. The ϕ-function gives an anisotropic behavior to the regularization, i.e., it will preserve the edges of the tensor field while smoothing homogeneous regions. We recall here the implementation details of such a regularization procedure using a finite differences strategy (as shown in Sec. 4.3.4).

As we are working on a vector space, the gradient of the regularization criterion can be expressed as follows, using $\psi(s) = \phi\prime(s)/s$:

$$\begin{aligned} \nabla \mathrm{Reg}(\mathbf{L}) &= -2\mathrm{div}\left(\psi\left(\|\nabla \mathbf{L}\|\right)\nabla \mathbf{L}\right) \\ &= -2\psi\left(\|\nabla \mathbf{L}\|\right)\Delta \mathbf{L} - 2\sum_{i=1}^{3}\partial_i\left(\psi\left(\|\nabla \mathbf{L}\|\right)\right)\partial_i \mathbf{L}. \end{aligned} \tag{6.13}$$

For the experiments, we used $\phi(s) = 2(1 + s^2/\kappa^2)^{1/2} - 2$ and $\psi(s) = (1 + s^2/\kappa^2)^{-1/2}$ as in [Chefd'hotel 2002]. κ can be seen as a normalization factor for the gradient. The key for the numerical implementation is the computation of the matrix and scalar fields $\partial_i \mathbf{L}$, $\Delta \mathbf{L}$ and $\|\nabla \mathbf{L}\|$ of $\mathbb{R}^3$. Using a finite difference scheme, these are simply:

$$\begin{aligned} \partial_i \mathbf{L}(x) &= \frac{\mathbf{L}(\mathbf{x}+\mathbf{x_i}) - \mathbf{L}(\mathbf{x}-\mathbf{x_i})}{2\|x_i\|}, \\ \Delta \mathbf{L}(\mathbf{x}) &= \sum_{i=1}^{3}\frac{\mathbf{L}(\mathbf{x}+\mathbf{x_i}) - 2\mathbf{L}(\mathbf{x}) + \mathbf{L}(\mathbf{x}-\mathbf{x_i})}{\|x_i\|^2}, \\ \|\nabla \mathbf{L}(\mathbf{x})\|^2 &= \sum_{i=1}^{3}\|\partial_i \mathbf{L}(\mathbf{x})\|_{LE}^2, \end{aligned}$$

$\|.\|_{LE}$ being the Log-Euclidean metric of Sec. 3.4.1. Finally, by combining the gradient of one of the criteria (Eq. [6.6], [6.8] or [6.11]) with the gradient of Eq. [6.13], we obtain the gradient of the full criterion of Eq. [6.5]. Consequently the evolution equation of the joint estimation and smoothing of DTI is:

$$\begin{aligned} \mathbf{L_{t+1}} &= \mathbf{L_t} - dt\nabla E(\mathbf{L_t}) \\ &= \mathbf{L_t} - dt/2\left(\nabla \mathrm{Sim}(\mathbf{L_t}) + \lambda \nabla \mathrm{Reg}(\mathbf{L_t})\right). \end{aligned}$$

Of course, one has to take the exponential of the solution to obtain a tensor.

We call the full criterion a Maximum A Posteriori estimator, because a spatial prior (the regularization term) is taken into account. We derive now three new potential estimators: the MAP log-Gaussian (ML log-Gaussian estimator + regularization), the MAP Gaussian (ML Gaussian + regularization) and finally the MAP Rician (ML Rician + regularization). We now investigate the effects of these estimators on synthetic and real datasets.

6.3 Quantitative and Qualitative Evaluation

To quantify the benefits of this methodology, we first perform 7 types of estimation on synthetic data: a classic estimation with an algebraic resolution (Classic), the ML log-Gaussian (Eq. [6.6]), the ML Gaussian (Eq. [6.8]) and the ML Rician (Eq. [6.11]). Then, the regularization term is added to turn each ML estimator into MAP estimations (MAP log-Gaussian, MAP Gaussian, MAP Rician), which gives a total of 7 different estimations. Second, we apply the same methodology on 2 clinical datasets, of medium and low quality. Results are presented and discussed in the sequel.

6.3.1 Synthetic Data

We generated a synthetic $16 \times 16 \times 16$ tensor field containing two homogeneous regions with anisotropic tensors (Fig. 6.3 left) as in [Wang 2004a]: the first region (R1) contains tensors whose coefficients are: $(0.970, 0.0, 1.751, 0.0, 0.0, 0.842)$ stored this way: $(d_{xx}, d_{xy}, d_{yy}, d_{xz}, d_{yz}, d_{zz})$. The second region (R2) contains tensors defined as: $(1.556, 0.338, 1.165, 0.0, 0.0, 0.842)$. A corresponding S_0 image is created with a constant value of 10. The DWI are synthetically produced using the Stejskal & Tanner equation with 6 diffusion gradients simulating the real data of Sec. 6.4. Finally, a Rician noise is added to each simulated DWI, including the S_0, with three standard deviations: 0.5 (SNR$\simeq 10dB$), 1.0 (SNR$\simeq 8dB$) and 1.5 (SNR$\simeq 6dB$). Fig. 6.3 b, c and d show a slice of a DWI for the three levels of noise. Parameter κ for smoothing was set to 0.05. We found that this value gives good smoothing results while keeping most of the transitions. Going up to 0.1$\sim$0.5 will smooth a little more the data, but will not preserve much the interfaces in the tensor field, while going down to 0.01 will preserve almost everything and thus result in less smoothing. Moreover, as LE metric are similitude-invariant, κ does not depend on the scale of the tensors, and the given values can be used with any type of dataset. We set $\lambda = 1.0$ and $dt = 100$ iterations, which gives correct results in terms of speed of convergence, stability of the gradient descent, and influence of the regularization for the MAP estimators. A good range for λ is $[0.25, 1]$. Results of the Classic, ML log-Gaussian, ML Gaussian and ML Rician estimations are shown respectively in Fig. 6.4 b, c, d and e. Results of the MAP log-Gaussian, MAP Gaussian and MAP Rician estimations are presented respectively in Fig. 6.4 f, g, h.

To quantitatively compare the methods, we compute the mean error, the variance, the minimum and maximum errors between each estimation and the original data with the LE metric. Results are summarized in Table 6.2. Since non-positive tensors appear using a classical estimation, the LE metric gives an infinite error. With the ML log-Gaussian estimation, all tensors are by nature positive definite. However, negative tensors in a classical estimation are turned into tensors whose eigenvalues are as close as possible to zero. Indeed, the energy is minimum when these tensor eigenvalues are negative. This results in a slow convergence and high errors. On the contrary, the ML Gaussian estimator prevents the tensors to degener-

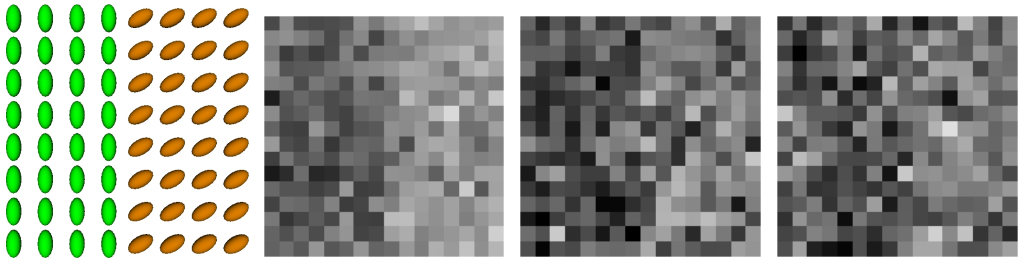


Figure 6.3: **A slice of a generated DWI with $\sigma_n = 0.5$ (b), $\sigma_n = 1.0$ (c), $\sigma_n = 1.5$ (d).** The encoding gradient is $\mathbf{g} = (-1, 1, 0)^\top$.

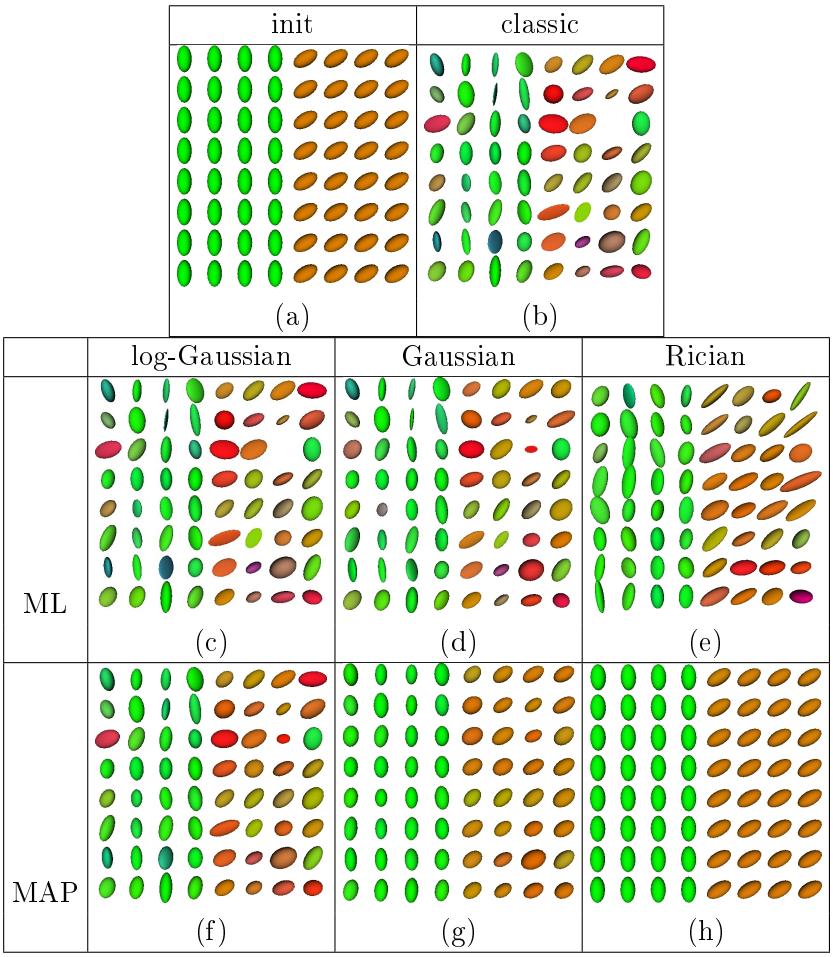


Figure 6.4: **2D slices of 7 different estimations of a noisy synthetic DTI dataset ($\sigma_n = 1.0$). (a):** The 3D synthetic field. **(b):** The classic estimation (non-displayed ellipsoids correspond to non-positive tensors). **(c):** The ML log-Gaussian, **(d):** The ML Gaussian, **(e):** The ML Rician, **(f):** The MAP log-Gaussian, **(g):** The MAP Gaussian, **(h):** The MAP Rician.

ate. Finally, the ML Rician estimator allows us to correct for the bias induced by the noise in each DWI and the shrinking effect vanishes, resulting in better quantitative results. The MAP estimators present the nice feature to preserve discontinuities between tensors, even when the tensors shapes are similar. This is an interesting feature when one would like, for instance, to smooth brain DTI: regions delimiting different fiber tracts have tensors about the same size but with different orientations. With our MAP estimators, these limits will be preserved.

Table 6.2: **Estimation of a synthetic dataset with 7 methods.** RMSE are calculated with the Log-Euclidean metric. FA and ADC columns indicate mean values of the estimations. The percentage is the relative difference to the true mean values (FA= 0.39, ADC= 3.56 $10^{-3}mm^2s^{-1}$). FA increases with noise. Using MAP estimators permits to cancel this effect. Conversely, ADC is underestimated when not using the ML/MAP Rician estimators.

$\sigma_n = 0.5$

	NPT	Mean Error	Variance	Min Error	Max Error	FA	ADC ($10^{-3}mm^2s^{-1}$)
Classic	7	∞	∞	0.077	∞	0.403 (2.7%)	3.567 (0.1%)
ML log-Gaussian	0	0.459	0.342	0.077	13.670	0.403 (2.7%)	3.567 (0.1%)
ML Gaussian	0	0.439	0.052	0.077	3.389	0.397 (1.2%)	3.534 (-0.8%)
ML Rician	0	0.250	0.007	0.051	0.666	0.402 (2.4%)	3.575 (0.3%)
MAP log-Gaussian	0	0.163	0.005	0.035	0.937	0.393 (0.1%)	3.573 (0.3%)
MAP Gaussian	0	0.141	0.003	0.032	0.374	0.385 (-1.9%)	3.517 (-1.3%)
MAP Rician	0	0.075	0.002	0.011	0.262	0.390 (-0.6%)	3.558 (-0.1%)

$\sigma_n = 1.0$

	NPT	Mean Error	Variance	Min Error	Max Error	FA	ADC($10^{-3}mm^2s^{-1}$)
Classic	224	∞	∞	0.128	∞	0.428 (9%)	3.552 (-0.3%)
ML log-Gaussian	0	1.641	8.777	0.128	14.255	0.428 (9%)	3.552 (-0.3%)
ML Gaussian	0	1.086	0.718	0.128	5.413	0.410 (4.5%)	3.427 (-3.8%)
ML Rician	0	0.718	0.064	0.095	2.397	0.412 (5%)	3.58 (0.4%)
MAP log-Gaussian	0	0.584	0.051	0.011	2.032	0.409 (4%)	3.564 (0.1%)
MAP Gaussian	0	0.543	0.034	0.037	1.648	0.369 (-6%)	3.371 (-5.4%)
MAP Rician	0	0.120	0.004	0.017	0.545	0.390 (-0.6%)	3.534 (-0.8%)

$\sigma_n = 1.5$

	NPT	Mean Error	Variance	Min Error	Max Error	FA	ADC($10^{-3}mm^2s^{-1}$)
Classic	717	∞	∞	0.205	∞	0.446 (14%)	3.423 (-4%)
ML log-Gaussian	0	3.518	22.921	0.205	14.422	0.446 (14%)	3.423 (-4%)
ML Gaussian	0	1.889	2.117	0.205	7.212	0.422 (7.5%)	3.207 (-10%)
ML Rician	0	1.525	0.542	0.261	6.332	0.441 (12%)	3.61 (1.3%)
MAP log-Gaussian	0	0.989	0.140	0.188	3.039	0.425 (8%)	3.434 (-3.6%)
MAP Gaussian	0	0.984	0.134	0.208	2.932	0.366 (-6.7%)	3.107 (-13%)
MAP Rician	0	0.394	0.046	0.033	1.889	0.410 (4.4%)	3.474 (-2.5%)

Effects of these estimators on FA and ADC are detailed in Tab. 6.2 (last two columns). One notices that FA increases with noise, as reported in [Anderson 2001, Skare 2000]. ML Gaussian and Rician estimators are slightly less sensitive to noise w.r.t. FA than a Classic or ML Log-Gaussian estimator (the growth in FA is lower). The main advantage comes with the regularization (MAP estimators) which lower the FA compared to ML estimators. Conversely, ADC is underestimated when noise increases, which is a consequence of the shrinking effect. The ML and MAP Rician estimators suffer less from this shrinking, managing to recover almost the loss in tensor size.

To further illustrate the shrinking effect, we evaluated the percentage of tensor volume lost during the estimation (the tensor volume is its determinant) for the three levels of noise. Results are presented in Table 6.3. The shrinking effect is obvious when the SNR is low. As depicted in Table 6.3, the ML Rician estimator corrects for this effect, even when the noise variance is high.

These experiments show that using the ML estimator for the correct noise model helps to correct for the shrinking effect one can observe with a Rician noise and a low SNR. Moreover, the anisotropic regularization enforces the spatial correlation while preserving discontinuities of the diffusion tensor field, making our estimators suitable for clinical datasets with low SNRs.

Table 6.3: **Illustration of the shrinking effect.** Mean volumes are mean tensor determinants of each estimation. The percentage of volume loss increases with the noise variance. Note that the ML Rician estimator is correcting for this effect.

$\sigma_n = 0.5$

	Original Data	ML log-Gaussian	ML Gaussian	ML Rician
Mean volume	1.43	1.41	1.39	1.43
Volume loss	NA	1.4%	2.9%	0.0%

$\sigma_n = 1.0$

	Original Data	ML log-Gaussian	ML Gaussian	ML Rician
Mean volume	1.43	1.25	1.15	1.42
Volume loss	NA	12%	19.2%	0.7%

$\sigma_n = 1.5$

	Original Data	ML log-Gaussian	ML Gaussian	ML Rician
Mean volume	1.43	1.11	0.81	1.40
Volume loss	NA	22%	43%	2%

6.4 Clinical Data

We tested the methods on 2 clinical datasets of medium and low quality. First, we used a brain dataset (Fig. 6.5) acquired on a 1.5T scanner with 7 encoding gradients (Basser sequence [Basser 1994a], b-value of 1000 s.mm^{-2}). The image dimensions are $128 \times 128 \times 30$ and the spatial resolution is $1.875 \times 1.875 \times 4\,\text{mm}^3$. Second, we used an experimental acquisition of the spinal cord on a 1.5T scanner[1] (Fig. 6.6 and 6.10) obtained with the same 7 encoding gradients and b-value as previously. The dimensions are $128 \times 128 \times 24$ (acquisition is coronal) with a spatial resolution of $1.4 \times 1.4 \times 1.4\,\text{mm}^3$. This new type of acquisition is currently actively investigated in clinical research (e.g., see [Facon 2005]) and is difficult to perform. Indeed, the patient often cannot stay long enough in the scanner due to pathology. Moreover, the small entrance of the scanner forces the patient to have an uncomfortable position, and the scanning time must be shortened. Finally, the coil cannot be perfectly adapted to the body as it is for the head. The images are consequently much noisier than for the brain MRI. Note that these datasets were actually collected in a clinical environment: the brain dataset corresponds to a patient with a tumor, and the spinal cord one was acquired to check for possible compressions [Facon 2005]. Estimation of the SNR in the brain dataset gives $\sigma = 12$ (SNR$\simeq 8dB$). Estimation of the SNR of the spinal cord dataset gives $\sigma = 14$ (SNR$\simeq 6dB$). Parameters used for the estimation are: $\kappa = 0.05$, $\lambda = 0.25$, $dt = 1.0$ and 100 iterations. Each of the ML and MAP log-Gaussian, Gaussian and Rician estimations took about 12 minutes to run on a PC with a Pentium M at 2GHz with 1 Gb of memory.

Figures 6.5 show a closeup of the splenium region and nearby. We clearly see that the missing tensors in a Classic estimation (fig. 6.5 c) that are not positive-definite are not completely replaced with a ML log-Gaussian estimation (fig. 6.5 d) due to degenerate tensors. Using a ML Gaussian estimation (fig. 6.5 e) and the ML Rician (fig. 6.5 f) results in a field where all tensors are positive-definite. Below, we investigate quantitatively the effects of these estimators on the tensor field.

[1] The authors would like to thank Denis Ducreux, MD, for providing the brain and spinal cord DTI dataset.

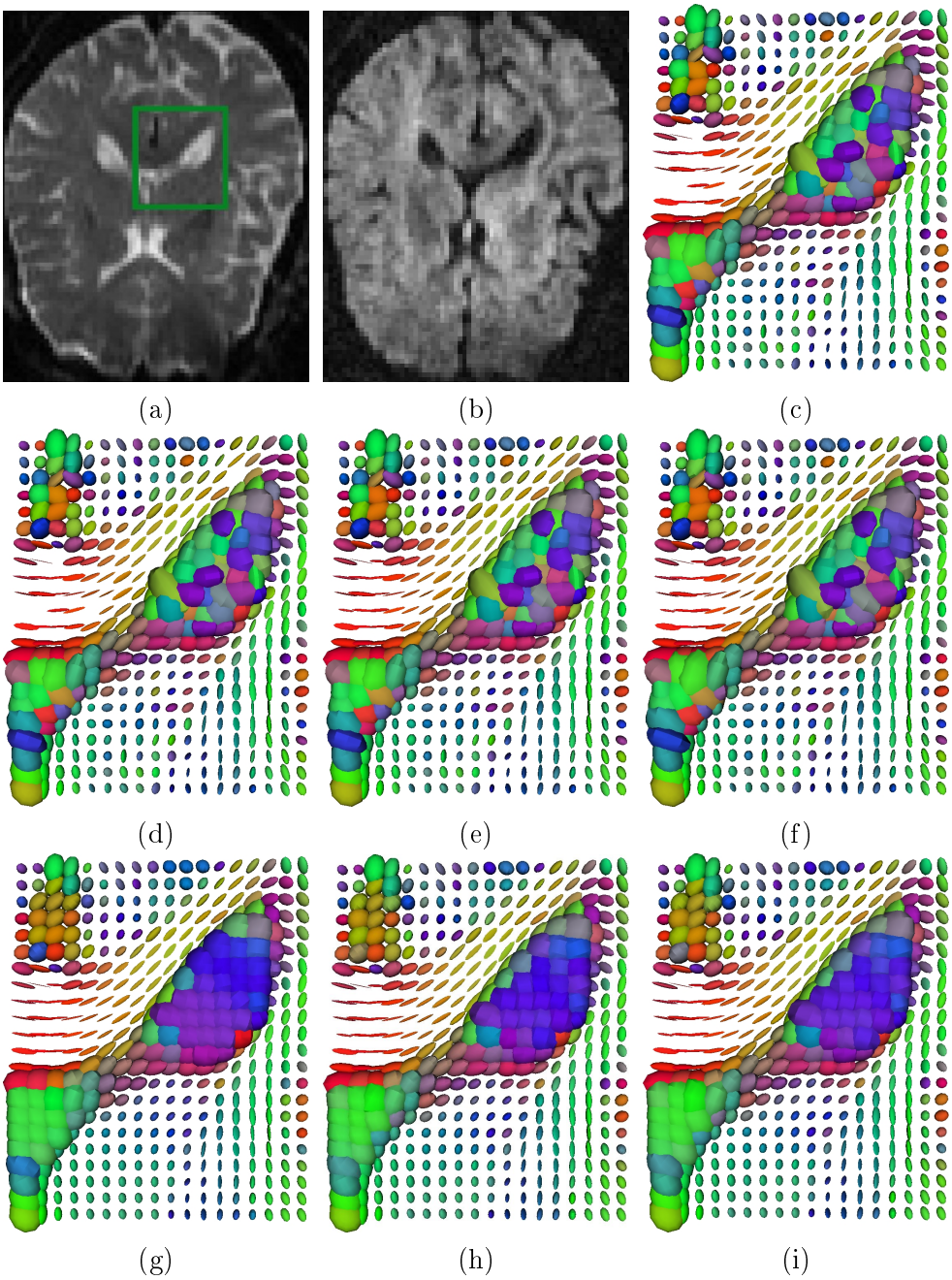


Figure 6.5: **Tensor field estimation of a brain DTI dataset. (a):** The b_0 image. The other figures are a closeup on the region delimited by the white square. **(b):** The DWI corresponding to the encoding gradient $\mathbf{g} = (1, 0, 1)$. **(c):** Classic estimation. **(d):** ML log-Gaussian. Some tensors are still missing because their eigenvalues are very close to zero. **(e):** ML Gaussian. All tensors remain positive definite. **(f):** ML Rician. All tensors are positive definite and are slightly bigger. **(g):** MAP log-Gaussian. The regularization term prevents the appearance of non-positive tensors. Note that the boundary between the ventricles and the splenium was preserved. **(h):** MAP Gaussian and **(i):** MAP Rician. The three MAP estimators give very close results.

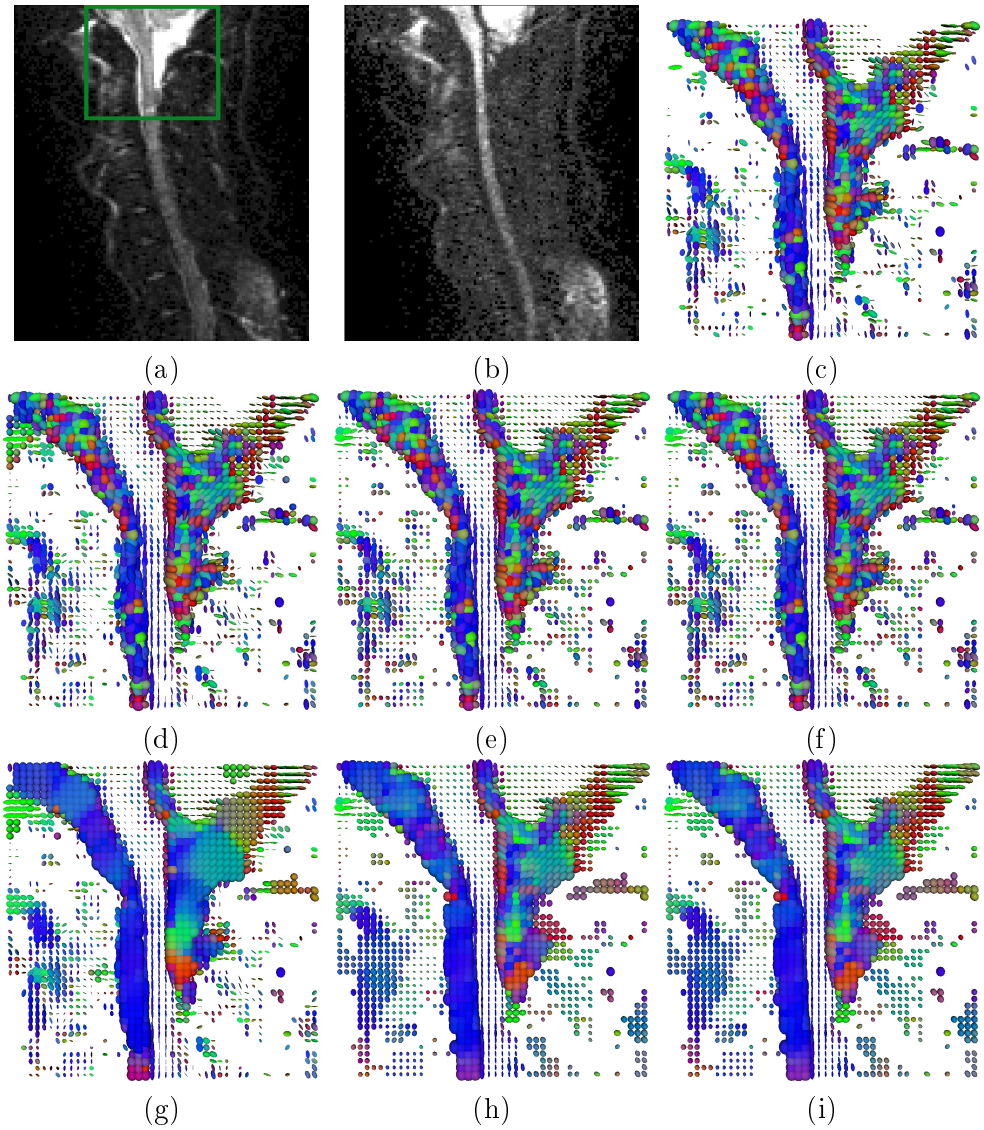


Figure 6.6: **Tensor field estimation of a spinal cord dataset. (a):** The b_0 image. The other figures are a closeup of the region delimited by the white square. **(b):** The DWI corresponding to the encoding gradient $\mathbf{g} = (1, 0, 1)$. **(c):** Classic estimation. Many tensors are missing in and around the spinal cord. **(d):** ML log-Gaussian. Same tensors than with the Classic estimation are missing. **(e):** ML Gaussian. All tensors are positive definite and regions outside the spinal cord are coherent and show no artificial anisotropy. **(f):** ML Rician. **(g):** MAP log-Gaussian, **(h):** MAP Gaussian and **(i):** MAP Rician. The spinal tract is smoothed and boundaries with nearby isotropic regions are preserved.

To evaluate the quality of the tensor fields estimated, we computed the mean apparent diffusion coefficient (ADC), the mean fractional anisotropy (FA) and the mean volume (VOL) in 2 distinct regions with different diffusion properties. First,

we chose the ventricles, where the diffusion is high but isotropic, and second, we chose the corpus callosum (including the splenium and genu), where the diffusion is known to be restricted by neural fibers, thus exhibiting high FA values. Regions were segmented manually on axial slices. Results for the 7 estimations are presented in Table 6.4. As expected, the ML/Map Rician estimations produce the greatest volumes (we found an increase of about 14%), and the highest ADC values. A good property is that in anisotropic regions like the corpus callosum, the gain of volume does not induce a loss of anisotropy (see Table 6.4 right, last column).

Finally, adding the regularization term (fig. 6.5 g, h and i) not only smooths isotropic regions like the ventricles without blurring the nearby splenium tract, but also replaces non-positive tensors of the ML log-Gaussian estimation by positive ones. Other experiments showed that the same comments apply for the interface between grey and white matter. FA maps are shown in Fig. 6.7 first row. However, the effect on the tensor volume, ADC and FA is more difficult to analyze. In isotropic regions like the ventricles, the nearby anisotropic regions have a small influence

Table 6.4: **Quantitative comparison of 7 diffusion tensor estimations of a brain DTI dataset.** For each estimation, the mean volume, ADC and FA were evaluated in the ventricles, and the corpus callosum. In isotropic regions like the ventricles, one notices that the tensor volume is on average 14% (ML Rician) or 10% (MAP Rician) larger than with a Classic estimation. The ADC shows slightly higher values with the ML Rician estimation (4% growth) and MAP Rician (2% growth) than with the Classic estimation. The same remarks apply in the corpus callosum but is less marked.

Ventricles

	Volume	ADC ($10^{-3}mm^2s^{-1}$)	FA
Classic	14.43	6.63	0.26
ML log-Gaussian	14.31	6.62	0.25
MAP log-Gaussian	13.82	6.49	0.18
ML Gaussian	14.39	6.63	0.25
MAP Gaussian	14.18	6.58	0.22
ML Rician	16.47	6.87	0.27
MAP Rician	15.94	6.78	0.23

Corpus Callosum

	Volume	ADC ($10^{-3}mm^2s^{-1}$)	FA
Classic	0.63	2.66	0.64
ML log-Gaussian	0.63	2.66	0.64
MAP log-Gaussian	0.61	2.61	0.57
ML Gaussian	0.63	2.66	0.64
MAP Gaussian	0.65	2.65	0.61
ML Rician	0.65	2.68	0.64
MAP Rician	0.67	2.68	0.62

during the regularization process. Thus, tensors on the boundaries are corrupted by small anisotropic ones. Consequently, mean tensor volumes in the ventricles with MAP estimators are lower than those without regularization. The same remark applies to the ADC (see Table 6.4 left, second column). Conversely, in regions with anisotropic tensors, large isotropic neighbors may influence the results, leading to higher volume and ADC (Tab. 6.4 right, first and second columns). Effects on FA maps are shown in Fig. 6.7 last row.

Figure 6.6 shows the results of the estimations of the spinal cord dataset. A closeup is made on the top of the spinal cord. The same remarks as for the brain dataset apply: the Classic estimation (Fig. 6.6 c) and the ML log-Gaussian (Fig. 6.6 d) lead to approximately the same results. Working with ML Gaussian and Rician estimators (Fig. 6.6 e and f) ensures that all tensors remain positive, with the advantage that the ML Rician estimator corrects for the shrinking effect: tensor volumes inside the spinal cord have grown by about 30% compared to the Classic estimation (the spinal cord was manually segmented from the baseline image). We found a greater growth of tensor volumes in the spinal cord than in the brain (within the brain, tensors have grown on average by 10%). This difference can be explained by the noise level: the SNR of the spinal cord dataset is the lowest. Effects on FA maps are presented in Fig. 6.8 top row. The regularization term (Fig. 6.6 g, h and i) smooths the field while preserving the boundaries with the spinal tract.

With the MAP log-Gaussian estimation, some artificial anisotropic tensors appear: The FA map (Fig. 6.8 d) presents high values outside the spinal cord, and is thus noisier than those obtained with MAP Gaussian and Rician estimators (Fig. 6.8 e and f).

This exemplifies the importance of the choice of the noise model: For low quality data, considering a log-Gaussian noise may not be the right choice, even with a MAP estimation. Switching to Gaussian or Rician noise models can largely improve the quality of the tensor estimation.
We now study how the MAP Rician estimator impacts the quality of fiber tracking on these two datasets.

6.4.1 Improvement of Tractography

Tractography, or fiber tracking, is a process which runs at the end of the DTI processing pipeline. Among the numerous available methods for tracking fibers, we chose a relatively fast and easy to implement one [Fillard 2003] and show how the tracking can be improved by our variational estimation combined with regularization. Criteria for stopping the tracking are: a threshold on FA (if FA is too low, the tracking is stopped) and on the curvature (to forbid unlikely fibers having a high curvature). Prior to the tracking, tensor fields are resampled to obtain isotropic voxels: in general, the out-plane resolution is very low (e.g., the brain dataset here) and interpolating the tensors improves the regularity of the fibers. Resampling is interpreted as a weighted mean with trilinear coefficients. Such a mean is computed in the logarithmic domain and then mapped back to the tensor space with

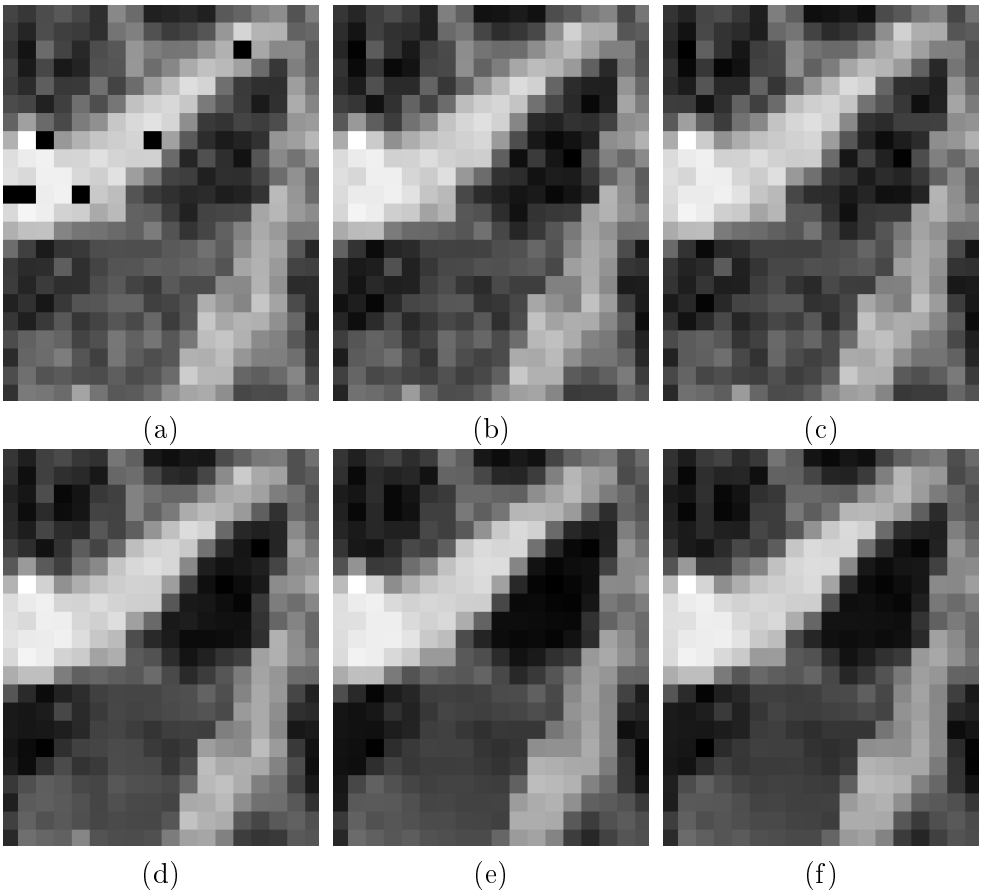


Figure 6.7: **Fractional Anisotropy (FA) of the tensor fields obtained in Fig. 6.5. (a):** ML log-Gaussian estimation. Outliers (black dots) in the splenium are caused by degenerate tensors. **(b):** ML Gaussian. One notices no outliers. **(c):** ML Rician. Same as previously. **(d):** Map log-Gaussian, **(e):** Map Gaussian and **(f):** Map Rician. With the three MAP estimators, the FA contrast between the ventricles and the splenium is well enhanced.

the matrix exponential: $\mathbf{D} = \exp(\sum_{i=1}^{N} \omega_i \log(\mathbf{D_i}))$, where ω_i are classical trilinear weights. We showed in [Arsigny 2006c] that such an interpolation has good practical properties in the context of DT-MRI, compared to a Euclidean interpolation. We tracked the fibers from the tensor fields obtained after the Classic estimation plus resampling and the MAP Rician estimator plus resampling. The parameters used for the tracking are: FA threshold: 0.3 (which is high but allows to extract only a subset of fibers in very anisotropic regions to facilitate the comparison), maximum angle of deviation: 90°. Results of tracking in the brain and the spinal cord are shown in Fig. 6.9, and 6.10. With the MAP Rician estimator, the tracking is qualitatively much smoother in both cases and shows less dispersion. The smoothness of the tensor field leads to more regular and longer fibers: tracts that were stopped due to the noise are now fully reconstructed. The FA threshold used ensures that

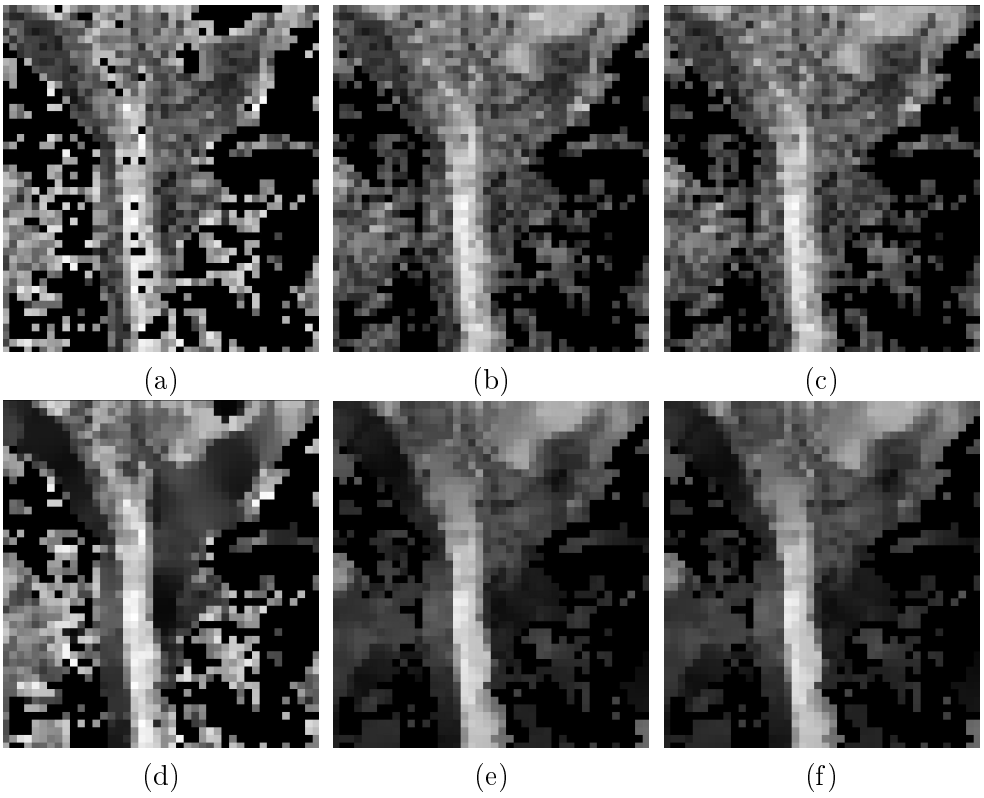


Figure 6.8: **Fractional Anisotropy (FA) of the tensor fields obtained in Fig. 6.6. (a):** ML log-Gaussian estimation. Outliers (black dots) in the spinal cord are caused by degenerate tensors. **(b):** ML Gaussian. One notices no outliers. **(c):** ML Rician. Same as previously. **(d):** Map log-Gaussian, **(e):** Map Gaussian and **(f):** Map Rician. With the three MAP estimators, the FA contrast between the spinal cord and around is well enhanced.

all fibers belong to white matter, and do not result from a tracking in CSF or grey matter.

6.5 Discussion

We introduced a new methodology to process DTIs of medium and low quality (typical of clinical applications) through a joint estimation and regularization of the diffusion tensor field. In particular, the estimation, which assumes that the data are corrupted by a Rician noise, is achieved through a maximum likelihood strategy adapted to the nature of this noise. This approach has the advantage to correct for the bias induced by the Rician noise in the DWI, and consequently not to underestimate the true volume of tensors (shrinking effect). Other estimation criteria which make the assumption of a Gaussian noise on the logarithm of the signal and on the signal itself are compared. These estimators are combined with an anisotropic regularization of the tensor field, so that transitions between homogeneous fiber tracts

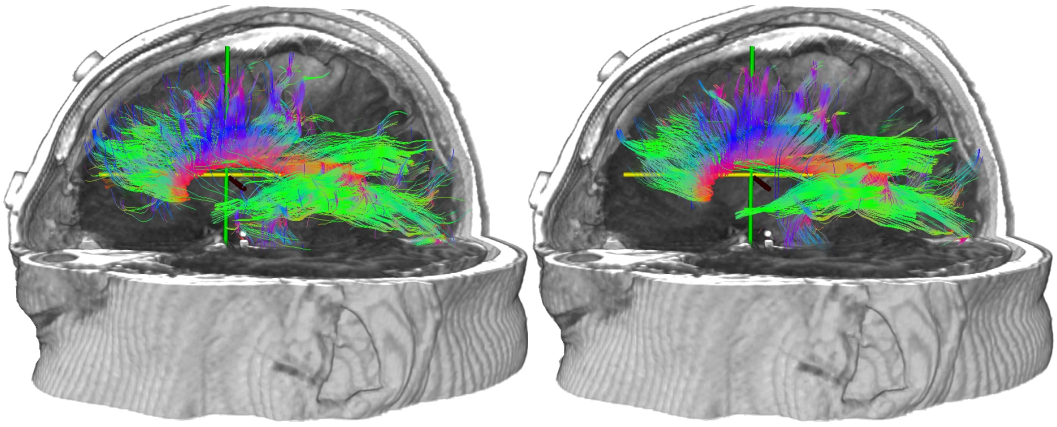


Figure 6.9: **Improvement of fiber reconstruction.** A seed region was placed inside the corpus callosum. Results of the fiber reconstruction after a classical estimation (**Left**), and after the MAP Rician estimation (**Right**). Fibers are overlapped with a volume rendering of the T1 image.

are preserved. To optimize these criteria, we use a Log-Euclidean metric that provides a fast and easy to use framework to process tensors. This tensor computing framework completely overcomes the limitations of the standard Euclidean calculus, and is well adapted to the processing of diffusion tensors.

Results on synthetic data show that considering a ML estimation adapted to a Rician noise model corrects for the shrinking effect, while assuming other noise models results in a loss of tensor volume after estimation. Results on real clinical

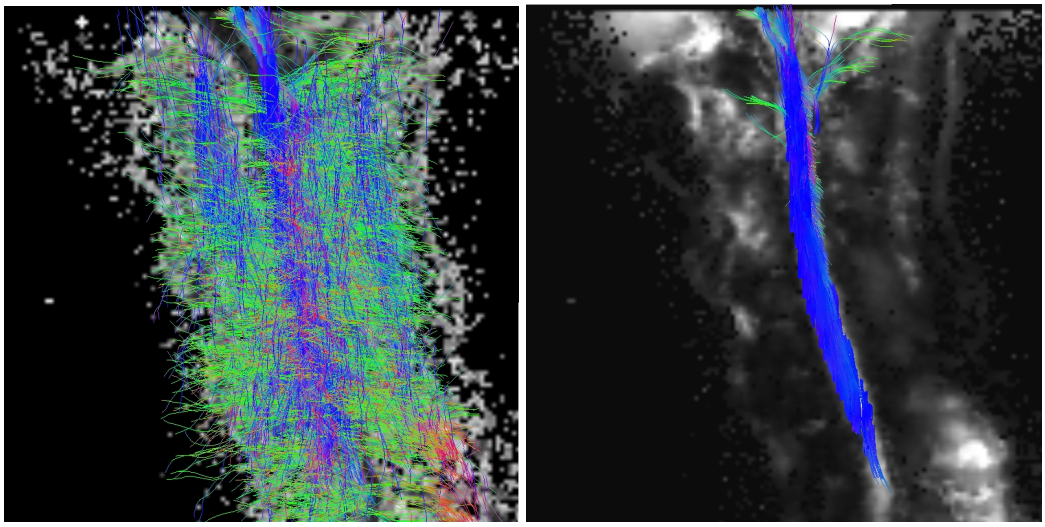


Figure 6.10: **Spinal cord fiber tract reconstruction.** A region containing the spinal cord was used for the tracking. **Left:** The spinal cord reconstructed after the Classic estimation. **Right:** The same tract after our proposed variational framework. Fibers are overlapped with a slice of the FA map.

datasets show that the use of ML estimators can be valuable in clinical studies: The MAP estimations nicely smooths homogeneous regions without blurring transitions between different tract in the case of brain DTI. Moreover, with a dataset of low quality, we showed that the choice of the noise model is important, as outliers may persist even with the regularization. The ML Rician estimator turns out to be the best choice in that case. Finally, the promising improvement of the fiber reconstruction of these data shows that even clinical DTIs can be used for tractography.

In the next chapter, we focus on the second aspect of our optimal workflow for DTI: the integration of DTI processing and visualization methods into a software usable by anyone.

Chapter 7

Towards a Clinical Integration of DTI with the Software MedINRIA

Contents

The increase and diversity of sources of information in medical imaging has raised the need for adapted tools for visualization and analysis of these data. For instance, the progress made in magnetic resonance imaging (MRI) for the past two decades has prodigiously improved the quality and resolution of 3D images. Moreover, these images can take various forms: anatomical MRI, functional MRI (fMRI) and diffusion tensor MRI are only a few examples. The analysis of these data is challenging too. For instance, computing neural fibers in DTI or segmenting lesions in multiple sclerosis (MS) require adapted and sophisticated methods. Moreover, at the Asclepios team, we are very often collaborating with medical experts and we are frequently facing the problem of how to efficiently transfer new techniques from the "methodological world" towards clinical research.

Most of state-of-the-art algorithms are not easily accessible to clinicians. These methods are in general available in relatively complex software, like Slicer [Pieper 2004], BrainVisa [Cointepas 2001] or Camino [Cook 2006]. Although these software suites are extremely powerful, they can be quite hard to fully master. Clinician's needs are simple yet very difficult to satisfy in a computer science point of view: they want ergonomic, reactive and intuitive softwares that offer real-time interactions with data. We started the MedINRIA project to offer medical experts a user-friendly interface to algorithms developed by the Asclepios project and all the potential collaborators. We thought that too many nice piece of methodologies would never be used because they do not have a proper user interface so that end-users can really test and use them. It was time to propose a solution to this problem.

MedINRIA can be seen as a collection of softwares targeting the clinicians. From a programmer point of view, MedINRIA offers a set of libraries for data manipulation and visualization so that programmers only need to focus on the integration of a new method, and not on the data visualization part nor on the user interface creation. We released these libraries as open-source and call them vtkINRIA3D [Toussaint 2007a]. From the user point of view, MedINRIA has the native look-and-feel of any system (Windows, MacOSX and Linux), and is a user-friendly graphical interface with a set of comprehensive buttons and online help.

Up to now, MedINRIA contains software for image visualization, DT-MRI analysis, tensor visualization, image registration and semi-automatic segmentation of MS lesions. In the following of this chapter, we will only focus on the applications useful for DT-MRI analysis (everything except the multiple sclerosis related applications). This chapter is organized as follows: In Sec. 7.1, some basics and key concepts of the MedINRIA's architecture are introduced. Then, a description of each application is given, from their general purpose to the analysis and interactions with medical data they introduce. We conclude in Sec. 7.6 with a discussion on the future work.

7.1 Software Description

Schematically, MedINRIA is composed of three main blocks, as shown in Fig. 7.1. First, the processing block is build upon the Insight ToolKit (ITK [Ibanez 2003]). It has all ITK features: flexibility (it can process any type of data), a multi-threaded implementation and a pipelined architecture. Second, the visualization block is made with the Visualization ToolKit (VTK [VTK 2007]), itself based on OpenGL, and offers the same pipelined structure as ITK. Finally comes the user interface block written in wxWidgets [wxWidgets 2007]. This block is responsible for transmitting user-defined parameters and user interactions to the two first blocks. Unlike the processing and visualization blocks, this one can only transmit data in one direction, and consequently cannot receive any information from the other blocks. The main reason is to keep the user-interface detached from the processing and visualization part. Thus, one could easily re-write the user-interface with another library, or even write command line tools instead of graphical applications.

When started, all MedINRIA's tools can be launched from a single menu. Several applications are available, each of them being dedicated to a specific task or type of data. We give below a description of four of them:

- *Image Viewer*, a simple yet powerful 3D medical image viewer;
- *DTI Track*, DT-MRI processing and visualization with Log-Euclidean metrics;
- *Tensor Viewer*, a diffusion tensor field visualization software;
- and *Image Fusion*, a software for manual rigid to fully automatic non-linear alignment of any type of 3D images.

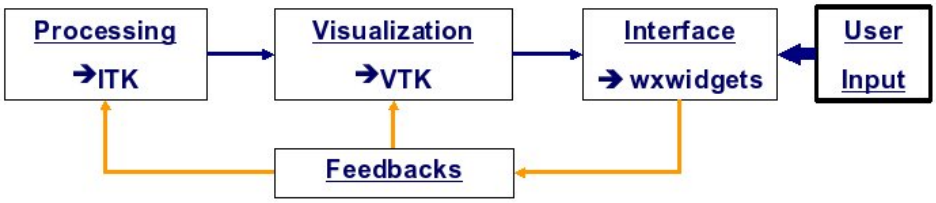


Figure 7.1: **The MedINRIA architecture.** MedINRIA is composed of three main blocks: 1. The processing block based upon ITK, 2. The visualization block using VTK and 3. The user interface block made in wxWidgets. The blue arrows indicate the data flow, from the processing block to the user interface. The user-interface can only send controls to the processing and visualization blocks (the opposite is forbidden). Thus, the interface can be re-written using another library with minimum effort.

We start by briefly introducing the general purpose of each application as well as their utility in DT-MRI processing, and continue by describing the processing and interactions technique they propose, to give to medical experts a fast and efficient insight into their data.

7.2 ImageViewer: A Simple yet Powerful Image Viewer

This first application aims at displaying 3D images and interacting with them. It is also used to convert clinical data (DICOM) into volumetric images. *ImageViewer* introduces interesting original features like: image tabbed browsing, a preview screen (Fig. 7.3 left) where a set of images can be compared (interactions are synchronized: slice or contrast change of one window affects all the other windows as well), which gives an ergonomic navigation through images.

The most important part of this first application w.r.t. DT-MRI analysis in clinics is the possibility to import diffusion weighted sequences from raw DICOMs. We describe this feature in the next section. Then, we focus on the user interactions with 3D images as this is the most innovative part for an intuitive manipulation of data.

DICOM Conversion to Volumes: In medical imaging, DICOM is the file format outputted by any scanning device. Consequently, a DICOM importation tool is mandatory for integrating software in a medical environment. In *Image Viewer*, a wizard simply asks the user to specify the DICOM folder (where all individual DICOM images are stored by the scanner), and files are automatically arranged to construct the 3D volumes of the patient. Manual interaction is possible for difficult cases (interlaced files for instance). Special care should be made when importing a DWI sequence. Indeed, DWI are composed of several images, often with interlaced slices due to the sequence acquisition (every first slice of all DWI are acquired first, then every second slices, etc.). In this case, volumes must be correctly reconstructed.

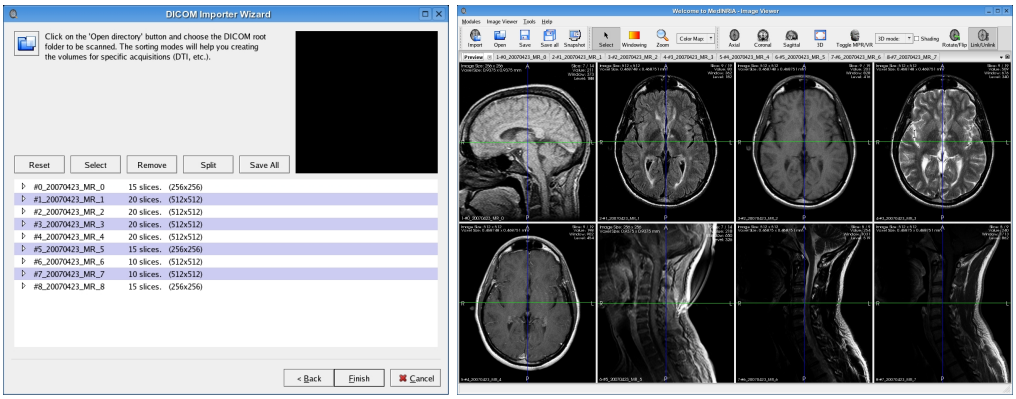


Figure 7.2: **Importing DICOM files with MedINRIA. Left:** The importer wizard simply asks the user for the DICOM folder and automatically reconstructs the 3D images. **Right:** After reconstruction, the preview screen of *Image Viewer* is loaded.

MedINRIA uses specific tags present in the DICOM files to differentiate between distinct DW images and successfully reconstructs interlaced volumes. As shown in Fig. 7.2 left, once the importation process has ended, the user can browse into the list of DICOM files composing a volume, and consequently visually inspect the result. After reconstruction, all volumes are imported into *Image Viewer* and the preview screen is loaded: all images can be compared and interaction are synchronized (Fig. 7.2 right).

2D Image Visualization: Classical 2D visualization of 3D volumes is performed by extracting slices in three orthogonal directions (axial, sagittal and coronal). Slice navigation is simply done by clicking on a view, and then moving the mouse upward or downward. The positioning is synchronized, i.e., each of the three orthogonal views show the same position at the same time. Contrast adjustment is also synchronized, and can be set with the mouse by changing the interaction mode via the menu bar (icons are comprehensive). Zoom and pan functions are also provided.

3D Image Visualization: Images are displayed in 3D either using multiplanar reconstruction (MPR) (Fig. 7.3 left) or volume rendering (VR) (Fig. 7.3 middle). MPR consists in displaying a slice in every orthogonal direction, while VR uses 3D textures with an opacity transfer function and is hardware-accelerated. The opacity transfer function for the VR is controlled by the contrast of the 2D views. The transfer function that controls opacity in VR is given by two parameters, *window* and *level*: any voxel whose scalar value is less than $level - window/2$ will be transparent, and any value above $level + window/2$ will be opaque. The ramp in-between is linear.

Volume Of Interest (VOI) extraction: Manual extraction of a VOI in VR is a desirable feature. Clinicians may want to have an insight into the patient's brain and remove only one block from his head by visually clicking on the screen. This

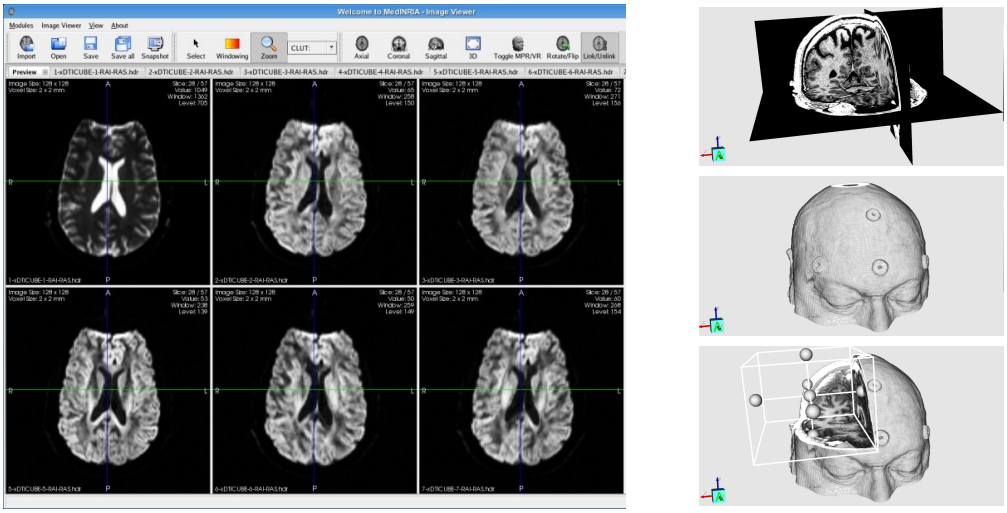


Figure 7.3: **Image visualization with MedINRIA. Left:** The preview screen. **Right**: Image rendering in 3D. Top: Multiplanar reconstruction (MPR) of a T1 image. An axial, sagittal and coronal slices is displayed. Middle: Volume rendering (VR) of the same image. Bottom: The cropping box was used to extract a VOI. By clicking on the handles, it can be resized (face handles) and translated (center handle).

is done in MedINRIA using a *cropping box*: it consists of a white box that can be scaled and translated by clicking on white handles present on each face (Fig. 7.3 right). The volume inside the box is automatically removed and cropping is done in real-time (hardware-accelerated).

7.3 DTI Track: Log-Euclidean DT-MRI Processing

The *DTI Track* application is dedicated to DT-MRI processing and visualization using Log-Euclidean metrics. *DTI Track* guides the user through the whole process, from importing the measures (i.e., the DWI) to the reconstruction of potential nervous fibers. The processing pipeline consists of three steps. First, the diffusion tensor field is estimated from the DWI. We use the standard approach which consists in using a linear version of the diffusion equation (Sec. 6). We have not integrated the non-linear estimation with a Rician noise model of Sec. 6 yet mainly for performance reasons: we need to optimize its implementation a little more before being practically usable. Second, as the standard estimation does not ensure the result to be positive definite, we replace any non-positive tensor by the Log-Euclidean mean of its positive neighbors. This step is fast and satisfactory removes non-positives matrices. An optional step consists in anisotropically smoothing the diffusion tensor field using Log-Euclidean metrics as described in Sec. 4.3.4. Special attention was taken to make these two steps as fast as possible, not to lower

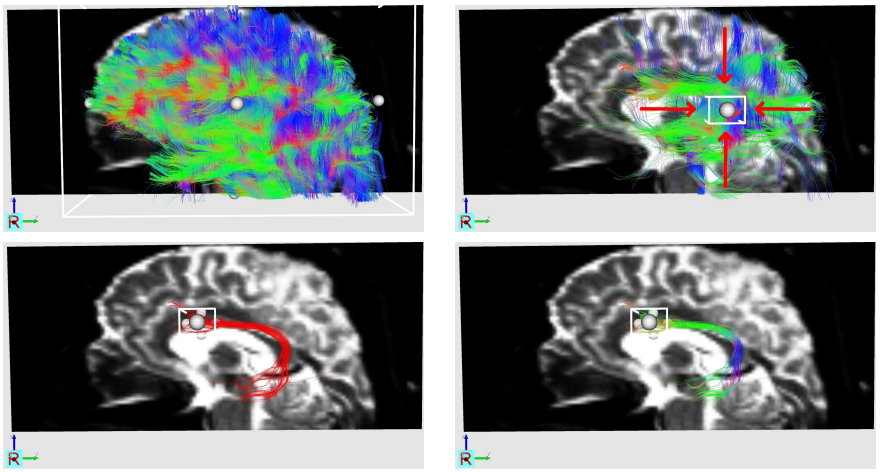


Figure 7.4: **Example of fiber bundling with MedINRIA. Top left:** The complete set of reconstructed fibers. The white box displayed is the *cropping box*. Only fibers that go through that box are shown. **Top right:** The box was shrunk by clicking on the face handles. The red arrows show the direction of the shrinking. The user can choose to navigate only in this sub-sample of fibers. **Bottom right:** The box was moved to another position, thus defining a second VOI. The bundle of interest (BOI) is ready to be tagged (i.e. labeled and saved). **Bottom left:** The BOI is now tagged and colored in red. It can be saved for further analysis.

the overall efficiency of the software. This includes a multithreaded implementation with optimized numerical libraries.

In third comes the tractography (Sec. 6.4.1). Our strategy in *DTI Track* is to start a new fiber at every voxel of the brain whose FA is greater than a user-defined threshold (default is 0.2). Consequently, we extract all possible brain fibers at once and store them in memory for further analysis. This also results in a gain of reactivity when defining ROIs: the user does not have to wait until completion of the tractography seeded by the ROIs, the result is immediate (fibers are already in memory). This last step generates thousands of nervous fibers that requires specific visualization and interaction techniques. In the following, we focus on the manipulation and analysis of fibers.

Fiber Manipulation: Depending on the image resolution, tractography can generate up to 100 000 fibers with an average of 100 points each. The task is not easy: one should display a maximum of information while keeping a real-time rendering. We chose to: 1. display fibers using only lines and not more complex geometric primitive like tubes, and 2. introduce a *cropping box* that intrinsically defines a VOI and only fibers that go through it will be displayed (Fig. 7.4). The cropping box can be scaled and translated with handles on each of its sides. Moving and scaling the box are real-time actions (fibers are pre-computed), giving clinicians a good insight of the brain connectivity. Moreover, the color-coding of fibers plays an important role: we color each fiber coordinate by the tangent at that point, i.e.,

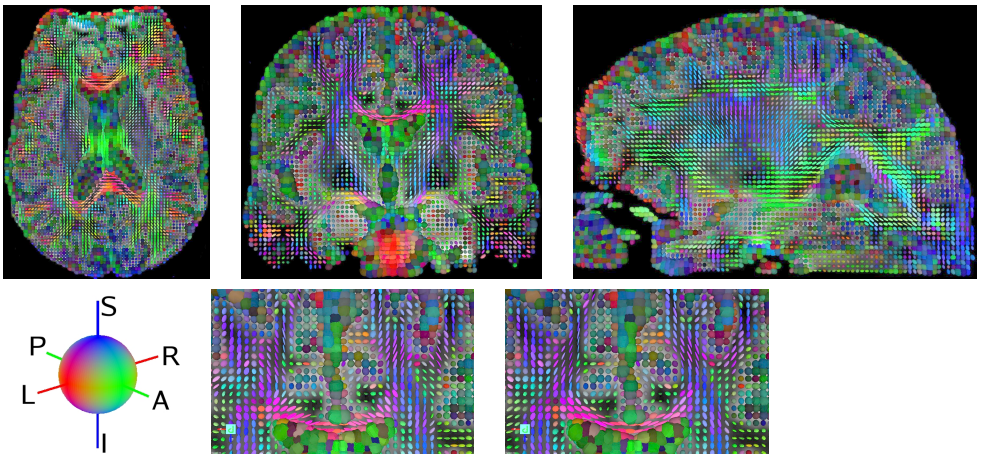


Figure 7.5: **Tensor visualization with MedINRIA. Top:** An axial (left), coronal (middle) and sagittal (right) slice of a 3D tensor field. **Bottom** The sphere indicates the color scheme of tensors (left). L: left, R: Right, I: Inferior, S: Superior, A: Anterior, P: Posterior. The tensor field (middle) looks flipped left-right. We applied a flipping (right) to correct this defect.

the normalized tangent is mapped into a RGB value (if a coordinate is negative, we take the absolute value). Then, if a fiber is left-to-right oriented, it is red, if it is infero-superior, it is blue, and if it is antero-posterior, it is green (see e.g., Fig. 7.4 top left).

Interactive Fiber Bundling: Random navigation into the reconstructed neural fibers may not be satisfactory. Experts might be willing to identify and label a specific fiber bundle of interest (BOI). For this purpose, we offer two solutions. The first one is to manually define a region of interest (ROI) slice by slice on a 3D image. When this image is the patient anatomy (T1), critical structures can be easily spotted and segmented. Accessing the exact bundle that goes through these structures can be of interest. The second method relies on the cropping box. It can be used recursively, i.e., multiple VOIs can be defined by positioning it at various locations. Thus, one can easily isolate a single fiber using this method. Note that both methods can be used alternatively. Once the user is satisfied with the BOI, he can label it, color it and save it for further analysis. Quantitative values of the bundle can be computed: apparent diffusion coefficient, fractional anisotropy, etc.

7.4 TensorViewer: Tensor Fields Visualization

Visualizing diffusion tensor fields produced by *DTI Track* is important for quality-control. In DTI, it is frequent that tensor fields are flipped in the X, Y or Z direction. This is caused by a misalignment of the diffusion gradient coordinate system with the patient acquisition frame. *TensorViewer* allows to visually control and correct for flipped tensors (Fig. 7.5 bottom). In *TensorViewer*, tensor fields

are considered as 3D images, i.e. three slices in three orthogonal directions (namely axial, coronal and sagittal) are displayed (see Fig. 7.5 top). However, one may question whether using ellipsoids for representing tensors is the best choice in terms of speed of refresh rate.

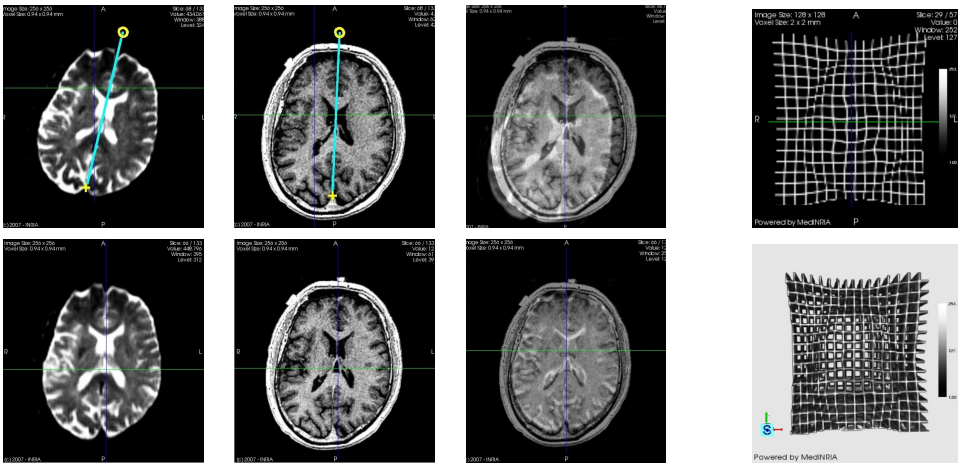
Which tensor representation to choose? The choice of the geometric primitive to represent a tensor is an active area of research (see e.g. [Kindlmann 2004]). As we said earlier, a tensor is fully represented by an ellipsoid: it can be divided into three eigenvectors (axis of the ellipsoid) and eigenvalues (axis magnitude). However, ellipsoids require lots of polygons to be rendered nicely (about 64 polygons). For instance, a dense field of $128 \times 128 \times 60$ voxels (typical resolution of DTI) gives a total of 2 millions polygons to display three orthogonal slices, by far too many for a real-time rendering on a regular computer (P4 2GHz, 1Gb of system memory and 256 Mb of graphic memory). To hasten the rendering, we give the possibility to downsample the field and lower the number of polygons. Moreover, we propose to represent a tensor not only by an ellipsoid, but by any other geometric primitives that can correctly represent the shape (from a sphere to a cigar-shaped ellipsoid) with fewer polygons. For instance, a cube, which has only 6 faces, gives a good approximation of the true tensor shape (the cube is rotated and scaled w.r.t. the eigenvectors/eigenvalues). To go one step further, if we get rid of the smallest eigenvalue (which is meaningless most of the time in DTI), planar shapes are possible too: a disk or a square are potential primitives.

Eventually, one can simply represent the principal eigenvector of a tensor instead of the full tensor itself: in DTI, it is admitted that the main eigenvector is aligned with the underlying oriented structure under specific conditions. These conditions are when the fractional anisotropy is high. In that case, a simple line is an acceptable shape for displaying tensors. MedINRIA proposes a listbox where all geometric shapes discussed above can be selected.

In addition to the shape chosen, we color code a tensor by its main eigenvector (absolute values of the coordinates are considered as RGB) as shown in Fig. 7.5 bottom left: red is for a left-right oriented tensor, green is for antero-posterior and blue is for infero-superior. The color magnitude is controlled by the FA: a dark tensor means absence of fibers. Conversely, a bright tensor means presence of fibers, and the color indicates the principal direction of the structure, according to the color code above. This provides a fast and accurate insight of the locations and directions of brain fibers.

7.5 Image Fusion: Fast and Simple Image Registration Toolkit

Image Fusion is dedicated to image registration. Among all image processing algorithms, registration remains a mandatory pre-processing step before doing statistical analysis, group comparisons or atlas formation. Moreover, it can be used in the con-



(a 1). (a 2). (a 3). (b)

Figure 7.6: **(a)**: Manual registration of a T2 and T1 image of the same patient. **(a) Top:** A point plus a direction were drawn on the moving T2 image (a 1), and fixed T1 image (a 2), (yellow cross plus blue line). The right image (a 3) is a blending of the previous images, showing that they are misaligned. **(a) Bottom:** The moving T2 image (a 1) after manual registration. The blending (a 3) shows that images are better aligned. **(b):** The deformation field is the output of the non-linear registration algorithm. It is displayed as a regular grid wrapped by the deformation field in a axial view (top) and in VR mode (bottom).

text of DT-MRI analysis to align the DWI onto the structural (namely T1) image. Indeed, medical experts prefer to have the T1 image overlapped with the fibers, because structural images generally have a better resolution and reveal finer structure compared to the b_0 image for example. Practically, we do not warp the DWI onto the T1 image (in that case we should also take care of tensor reorientation issue, see [Alexander 2001b] for a detailed description), but one computes a transformation matrix (affine matrix) that matches the geometry of the DWI and the geometry of the T1 image. Then, this matrix can be eventually loaded in *DTI Track* during the DWI importation process. Doing so, extracted fibers will be automatically warped w.r.t. the affine matrix and be aligned with the T1 image.

This application provides a simple interface for several types of registration. It takes two images as input, the "fixed image" and the "moving image", and offers up to now three ways of registering them:

- Manual Rigid Registration: In a few clicks, one can perform a manual rigid registration of two images. Rigid registration consists in a translation and a rotation. This is done intuitively by asking the user to place in both fixed and moving images a point plus a direction. The two points are used to compute the translation, and the directions are used to calculate the angle of rotation (Fig. 7.6).

- Automatic Affine Registration: We interface the affine registration framework of ITK [Ibanez 2003] to achieve this automatic matching of images. It consists in maximizing the mutual information of the two images, and is embedded in a multi-resolution framework. No user input is required. This method gives optimal results when registering two multi-modal images of the same patient acquired at different times. For matching images of different subjects, one should rely on the last method.

- Automatic Diffeomorphic Demons Registration: This non-linear registration method was very recently introduced by Vercauteren et al. [Vercauteren 2007]. It is based on variants of the Thirion's demons algorithm [Thirion 1998] for non-linear mono-modal image registration. It is fast, and insures the invertibility of the transform, which is a very desirable feature in non-linear registration.

Once images are registered, *RegistrationTool* offers two ways of visually evaluating the result: either using a blending technique (Fig. 7.6 (a 3)), or a checkerboard, which consits in creating an image with cubes, one cube out of two containing the moving and the fixed image, alternatively.

7.6 Conclusion on MedINRIA

The software MedINRIA is a collection of applications dedicated to medical image processing and visualization. The goal is to provide to clinicians an intuitive, reactive and powerful processing and visualization system. Along with classical features that any medical imaging software has, MedINRIA introduces quite a few new user-interactions that aims at easing the exploration of medical data by clinicians, who often need a simple, yet powerful, visualization system. For instance, MedINRIA proposes to extract a 3D volume of interest of an image interactively in real-time using a 3D box and volume rendering techniques. In addition, it introduces an interactive method to extract a neural fiber bundle of interest among thousands of fibers, using the same type of 3D box. MedINRIA does not only offer user-interactions for manipulating 3D data, but it also provides state-of-the-art algorithms that are most of the time not accessible to clinicians, as these algorithms are generally available in more research-oriented software. Here, we propose to process DT-MRI using Log-Euclidean metrics, and to visualize tensors using a panel of geometric shapes.

Future work includes the development of new applications: people (graduate students and engineers) from the Asclepios project are currently working on a system to automatically segment lesions in Multiple Sclerosis, a tumor growth simulation software, and a cardiac image processing toolkit. MedINRIA is growing fast, and will remain free. A Windows, Linux and MacOSX version can be downloaded at: http://www-sop.inria.fr/asclepios/software/MedINRIA.

In the next chapter, we illustrate how some of the DT-MRI processing methods developed in Chap. 6, as well as the software MedINRIA, could contribute to

the feasibility of DT-MRI in a clinical environment with a study on the possible indications in spinal cord lesions.

Chapter 8

Diffusion Tensor MRI and Fiber Tracking in Spinal Cord Lesions: Current and Future Indications

Contents

This chapter describes a clinical study realized and is highly medically oriented. It was written in collaboration with Denis Ducreux, MD, a radiologist at the Bicêtre hospital in Paris, and mainly follows reference [Ducreux 2007].

Magnetic resonance (MR) imaging plays a major role in the diagnosis and follow-up of spinal cord lesions. The main objectives of spinal cord imaging are to detect and characterize lesions, to assess the feasibility of surgical resection, and to diagnose recurrences and complications of therapy. Conventional MR imaging using T1- and T2-weighted sequences (in spin or gradient echo) lacks sensitivity in detecting and characterizing cord lesions, such as multiple sclerosis or acute spinal cord infarction. In addition, in patients who have cord tumors, conventional sequences

may not be able to clearly identify the transition between the tumor and the surrounding edema. In the brain, diffusion-weighted (DW) imaging is an established and reliable method that helps to detect and characterize such lesions, and diffusion tensor (DT) imaging is becoming an important technique to identify white matter tracts and the effects of different lesions on them. DW imaging and DT imaging are usually performed using echo planar sequences, which are sensitive to noise, motion, and susceptibility artifacts. These two caveats make it difficult to detect and characterize spinal cord lesions, particularly with DW imaging. In addition, the resolution of most currently clinically used DW imaging sequences is not optimal to image structures as small as the spinal cord and its internal features. Better characterization of white matter lesions (and therefore many cord lesions) may be achieved using DT imaging and fiber tracking (FT) algorithms [LeBihan 1991, Basser 1994a, Basser 1996, Wheeler-Kingshott 2002, Facon 2005]. Recently, several investigators have assessed the feasibility of performing spinal cord DT imaging studies [Facon 2005, Holder 2000, Ries 2000, Clark 2000, Bammer 2000, Cercignani 2003]. DT imaging sequences with computation of fractional anisotropy (FA) are more sensitive than spin echo T2-weighted images in detecting intrinsic abnormalities in acute or chronic spinal cord compression [Facon 2005, Demir 2003]. In lesions that produce involvement of white matter fibers, it has also been reported that DT imaging with FT may help to define abnormal areas that are undetected on routine T2-weighted imaging [Facon 2005]. In our experience, FA and FT maps derived from DT imaging computations may help neurosurgeons to better delineate spinal cord tumors and may contribute important information before tumor resection. In the following, we review the different methods available to obtain DT imaging and FT in the spinal cord and their clinical applications. We discuss novel and dedicated spine FT programs and speculate about the future of DT imaging and FT in spinal cord imaging.

8.1 Diffusion Tensor Imaging and Fiber Tracking Methods

8.1.1 Image Acquisition

DT imaging may be reliably performed on 1.5 T MR imaging systems with actively shielded magnetic field gradients, but strong gradients are needed ($\geq$30 mT/m) for optimal imaging. To decrease magnetic susceptibility artifacts intrinsic to echo planar DT imaging sequences, parallel imaging with the shortest echo time is desirable. The spinal cord is a small organ with less extracellular water than the brain. Consequently, small b values are required to prevent signal attenuation on DW imaging. SENSE or GRAPPA echo planar imaging or multishot fast spin echo sequences obtained with a b value of approximately 500 s/mm2 are most advantageous because they reduce magnetic susceptibility artifacts and result in shorter acquisition times [Renoux 2006]. Several DT imaging gradient directions are needed, and theoret-

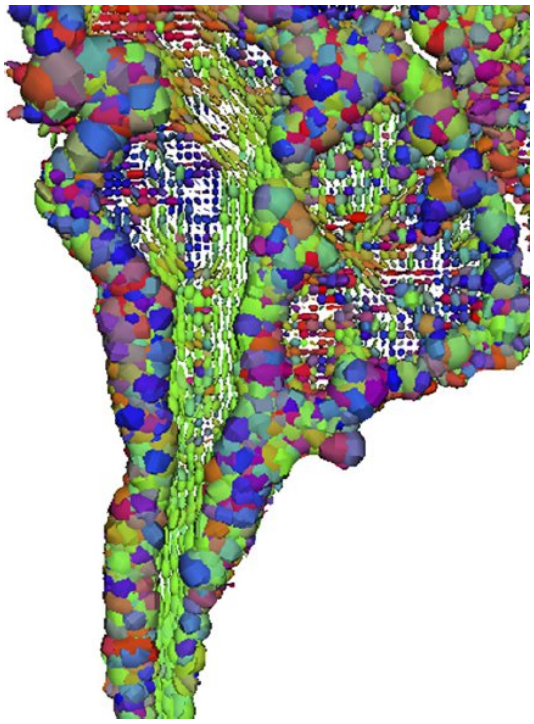


Figure 8.1: Color-coded diffusion tensor field of the spinal cord centered on the cranio-cervical area. Large ellipsoids are seen in the CSF areas because its motion at a microscopical level is not clearly directional, resulting in a mixture of colors, shapes, and sizes. Conversely, the cord shows small longitudinal ellipsoids (mostly green and uniform in shape and size).

ically more directions result in a better signal-to-noise ratio and better diffusion tensors. Investigators have tested 6, 12, 25, and 55 directions and have concluded that the best compromise between acquisition time and image quality is achieved with 25 directions. Data are then resampled in the sagittal plane (resulting in a more complete visualization of the spinal cord than in the axial plane) using a simple matrix rotation. Sagittal DT imaging sequences are most sensitive to water diffusivity along the main spinal cord axis. To decrease cerebrospinal fluid partial flow motion, cardiac gating may be used, especially for thoracic cord imaging, but it is not mandatory if saturation pulses are placed in the region of the heart. Based on these observations, we performed our studies using sagittal, single-shot, spin echo, echo-planar parallel GRAPPA DT imaging with an acceleration factor of 2 and 25 noncollinear gradient directions with two b values (b=0 and 500 s/mm2) (field of view: 180×80 mm; image matrix: 128×128; 12 slices with a thickness of 3 mm, nominal voxel size: 1.4×1.4×3 mm, and TR/TE=4600/83 ms). Acquisition time is slightly over 3 minutes per study, during which time the patients are asked to hold still and to try to avoid swallowing. Other acquisition schemes exist (field of view: 200 mm; axial 2 mm slice thickness, and b value 800 s/mm2), but our experience

showed that they fail to generate adequate FT images needed to visualize the white matter tracts in the spinal cord.

8.1.2 Image Analysis

Image analysis is performed on a voxel-by-voxel basis using dedicated softwares (DPTools http://www.fmritools.org, and MedINRIA http://www-sop.inria.fr/asclepios/software/MedINRIA). Before performing the tensor estimation, an unwrapping algorithm is applied to the DT imaging dataset to correct for distortions related to eddy currents induced by the large diffusion sensitizing gradients. This algorithm relies on a three-parameter distortion model including scale, shear, and linear translation in the phase-encoding direction [Haselgrove 1996]. Optimal parameters are assessed independently for each slice relative to the corresponding T2-weighted image by maximization of an entropy-related similarity measure called "mutual information" [Poupon 2000]. This algorithm has proven to be fast and reliable and is used in many brain applications. We adapted it to the spinal cord to decrease the distortions induced by the echo planar sequence. DT imaging acquisitions can be further processed without distortion corrections, but FA and FT maps may show artifacts at the edges of the field of view. After distortion correction, the diffusion tensor and subsequently the eigen decomposition (with eigenvalues λ_1, λ_2, λ_3) is calculated on a voxel-by-voxel basis. Thus, the diffusion tensor field (Fig. 8.1), the apparent diffusion coefficient ADC $= (\lambda_1 + \lambda_2 + \lambda_3)/3 = \lambda$, and the FA:

$$\mathrm{FA} = \sqrt{\frac{3}{2}}.\frac{\sqrt{(\lambda_1 - \lambda)^2 + (\lambda_2 - \lambda)^2 + (\lambda_3 - \lambda)^2}}{\sqrt{(\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

are calculated. An FA value around 1 indicates a very anisotropic tensor, while an FA value close to 0 means that the tensor is isotropic.

8.1.3 Fiber Tracking Method

In addition to the two-dimensional parametric color maps (also called directionality maps) obtained using the previously mentioned method, three-dimensional white matter fiber tracts maps can be generated. Many FT algorithms are available in the literature, and can be divided into two categories: fiber tractography and connectivity mapping. The former approach basically consists in integrating the field of principal direction of diffusion (main tensor eigenvector) and attempts to reconstruct the fibers that pass through a given point [Basser 2000, Xu 2002, Westin 2002, Mangin 2002]. The latter approach attempts to estimate a connectivity map by simulating the diffusion process based on the data, and look for the most likely pathways between a seed and a target [Hagmann 2000, O'Donnell 2002, Lenglet 2004, Zhang 2007a]. Both approaches have pros and cons, but we will not discuss them as this is not the topic of this work. The algorithm we use (included in MedINRIA) is based on the advection-diffusion scheme proposed by [Weinstein 1999] (streamline fiber tracking), which performs

well in regions with two crossing fibers, and is combined with a Log-Euclidean anisotropic smoothing and interpolation to improve the reliability of fiber tracking in relatively low SNR images typical of our spinal acquisitions. MedINRIA estimates the tensor field of the DT imaging acquisition and then processes the entire study to find all detectable fibers. The user selects a region-of-interest that maps in three dimensions all fibers that pass through it. The process takes less than 30 seconds on a PC workstation. Three-dimensional reconstructions of white matter tracts are color coded. Green denotes cranio-caudal fibers, blue denotes the left-to-right fibers, and red denotes the anterior-posterior fibers. FT reconstructions may falsely depict aberrant fibers that are due to susceptibility artifacts, especially in the edges of the field-of-view. These artifacts do not affect the intrinsic cord fibers and are easy to recognize.

8.1.4 Fractional Anisotropy Measurements

Spinal DT imaging can be used to generate three types of maps (similar to those used in brain DT imaging). In the directionality map, the fibers traveling in different directions (left to right, bottom to top, and anterior to posterior) are assigned different colors that permit their identification. These maps result in a two-dimensional display and contain important information regarding anatomy and anisotropy of white matter tracts. Unfortunately, their resolution is limited. Directionality and FT maps provide anatomic information but cannot provide objective values on FA. FA maps thus result in numeric information that may be useful in many disease processes. As with many other new techniques, we recommend that each center obtain a set of normal FA measurements that can be used for comparison in cases of lesions. Special attention should be paid to avoiding cerebrospinal fluid partial volume effects and magnetic susceptibility and motion artifacts in the selection of each region-of-interest (ROI). For example, we initially performed cord FA measurements in healthy volunteers at three different levels (cervical [C2-C5], high thoracic [T1-T6], and low thoracic [T7-T12]) using ROIs of 20 mm2 (10 voxels), which included gray and white matter. FA values are nearly identical at the different spinal cord levels, implying that anisotropy of white matter fibers is remarkably uniform throughout the cord. For spinal cord lesions, FA measurements should be performed at the site of abnormality using an ROI that is located completely within the lesion. To set the ROI, the b_0 images are used because they permit clear visualization of the lesions. From them, these ROIs are copied to the FA maps to avoid partial volume effects and magnetic susceptibility and motion artifacts.

8.2 Clinical Applications

8.2.1 Normal Anatomy

FT of the spinal cord shows the main white matter tracts: posterior-lateral corticospinal, posterior lemniscal, and spinal-thalamic (Fig. 2). On the fiber tracking

three-dimensional reconstructions, it is also possible to visualize the fibers of the nerve roots (Fig. 8.2).

8.2.2 Spinal Cord Tumors

DT imaging and FT may help to characterize some tumors and to delineate their margins [Xu 2002]. FA values are similar for astrocytomas (0.48±0.02), ependymomas (0.5±0.04), and metastases (0.46±0.04) but are different for hemangioblastomas (0.59). The lowest FA values are seen in metastases, and the highest are seen in hemangioblastomas. Because most tumors of the cord are gliomas, DT imaging may not play an important role in the histologic characterization of these tumors but may play a role in distinguishing hemangioblastomas from metastases (an observation that may be important in patients who have Von Hippel Lindau disease). The margins of tumors as seen on DT imaging match those seen on T2-weighted sequences. Using FA maps, the surrounding edema may be separated from the tumor due to lower values in the former. FT may show fibers that are warped or frankly destroyed by tumor (Figs. 8.3 and 8.4). This may be important when assessing highly infiltrative tumors and delineating their margins before resection. In addition, tumors such as metastases and hemangioblastomas are localized and tend not to infiltrate the surrounding areas (Fig. 8.5). Rotation of the three-dimensional FT maps is needed in some patients to localize the tumors (Figs. 8.4 and 8.5).

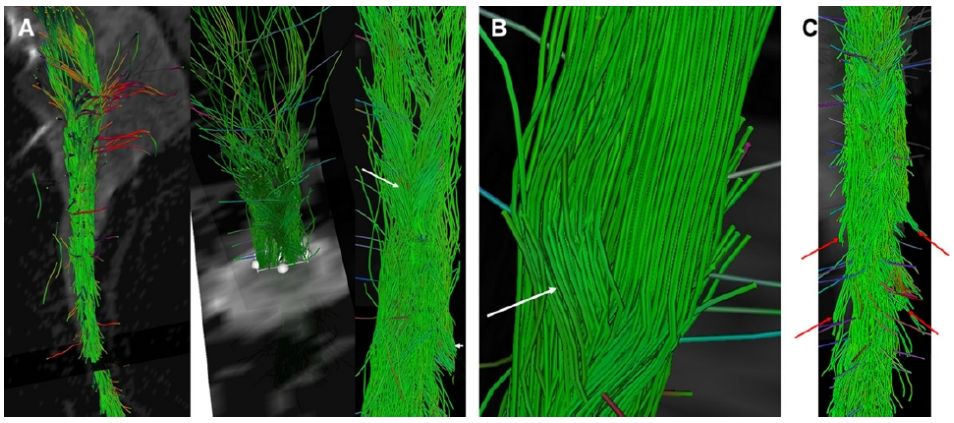


Figure 8.2: Fiber tracking performed on a volunteer's cervical spinal cord. (A) Sagittal view (left) shows green color coded posterior lemniscal tracts. A slightly oblique projection (center) shows the same tracts superimposed on an axial b=0 image. Magnified view (right) shows the decussating fibers (arrow) of these tracts. (B) Magnified view shows decussation of the spino-thalamic tracts (arrow). (C) Frontal view of the cervical spine shows fibers in nerve roots (arrows).

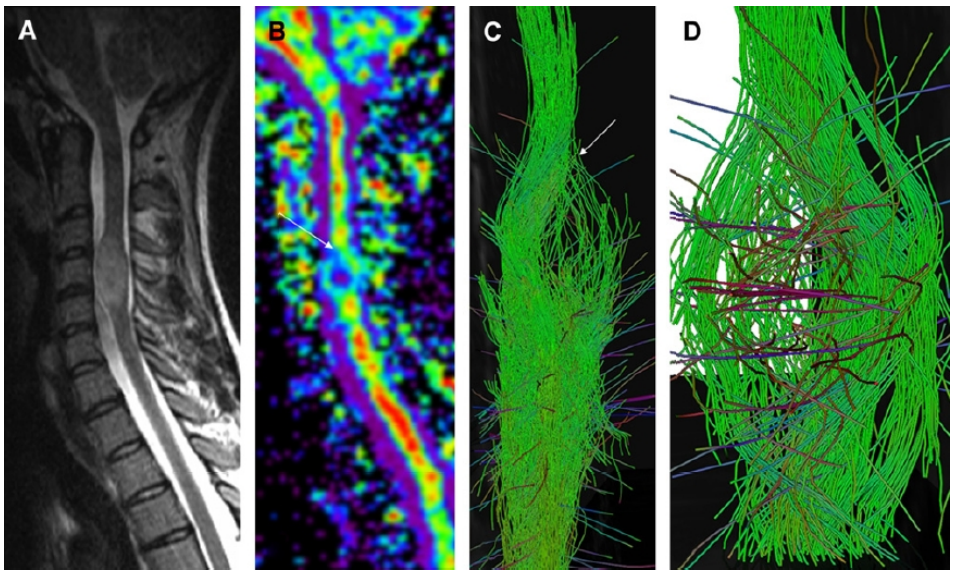


Figure 8.3: Cervical spinal cord astrocytoma. (A) Midsagittal T2-weighted image shows an expansile and mildly hyperintense tumor at the C2-C4 levels. (B) Directionality sagittal map shows loss of anisotropy in the mass (blue color, arrow) that is focal and compatible with diagnosis of WHO grade II astrocytoma. Anisotropy is immediately reestablished along tumor margins due to the lack of edema. (C) FT map shows splaying of fibers by a tumor normalized at the tumor margin (arrow). (D) Magnified FT map shows fiber splaying by a tumor. Most fibers are intact due to the nondestructive nature of tumor.

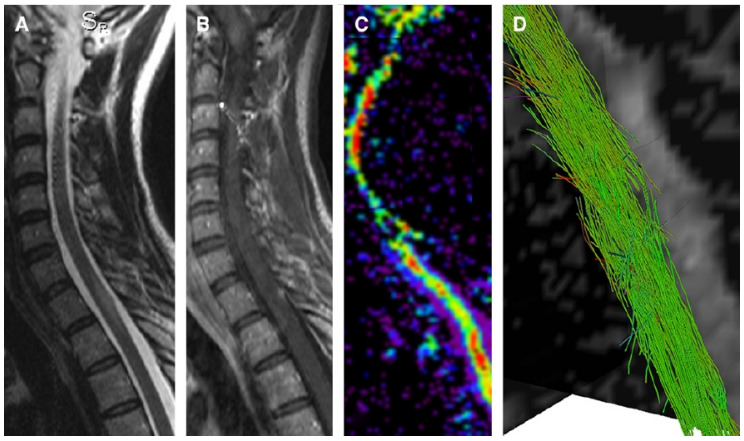


Figure 8.4: Cervico-thoracic spinal cord hemangioblastoma. Midsagittal T2 (A) and postcontrast T1 images (B) show focal enhancing lesion on the posterior aspect of the cord at the T1 level with surrounding edema. (C) Directionality map shows focal loss of anisotropy (blue color) in this noninvasive tumor. Anisotropy recovers immediately outside of the tumor margins. (D) Oblique FT map shows the focal "hole" where the tumor is located. Note splaying of fibers but no significant thinning or invasion.

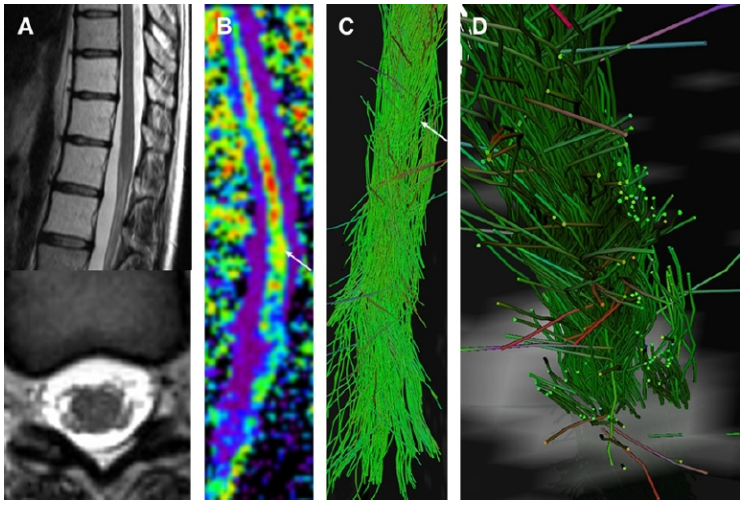


Figure 8.5: Breast metastasis to the conus medullaris. (A) Midsagittal (upper) and axial (lower) T2 image shows no significant abnormality. (B) Corresponding directionality map shows localized loss of anisotropy in this tumor. Metastases are well marginated tumors that generally do not invade the surrounding tissues. (C) Frontal FT map fails to show the lesion as a small "hole" (arrow) amid fibers. Note the nerve roots of the cauda equina inferiorly. (D) Oblique FT map shows to a better extent the focal hole corresponding to the site of the tumor in the inferior and left lateral aspect of the cord. Neighboring fibers are preserved. This case illustrates the significant sensitivity of DT imaging to localized lesions even when T2 images are normal.

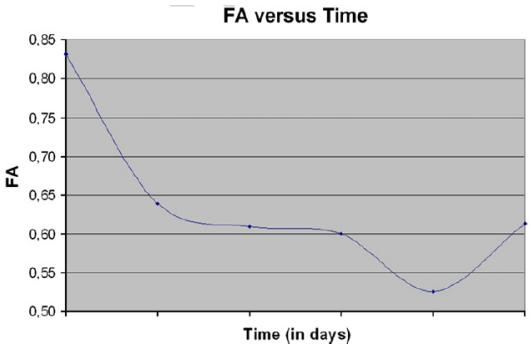


Figure 8.6: Time-course curve from Day 1 to Day 30 of the averaged FA estimated from sites of spinal cord compression. FA values decrease from Day 1 through Day 21 and then slightly increase. This phenomenon is related to extracellular water diffusivity. There is restricted diffusivity in the acute stage and increased diffusivity in the chronic stage. Preservation of diffusivity in the acute stage may imply a better prognosis.

8.2.3 Spinal cord compression

Conventional T2-weighted images may underestimate the effects of compressive lesions on the spinal cord, particularly when no hyperintense signal accompanies cord compression in the hyperacute period (which is the critical time to treat these lesions). DT imaging can detect abnormal areas within a normal-appearing spinal cord on T2-weighted imaging. FA has a better sensitivity (73%) and specificity (100%) in the detection of acute spinal cord abnormalities compared with conventional T2-weighted imaging [Facon 2005]. FA measurements versus the time of injury may also help to predict the patients' outcome (good outcome for FA values >0.6 and worse outcome for FA values <0.6 in the acute period) (Fig. 8.6). FT identifies the sites of compression and aids by depicting mass effect and discontinuity of white matter fibers, which also have a prognostic implication because patients who have the latter generally show little or no improvement (Fig. 8.7).

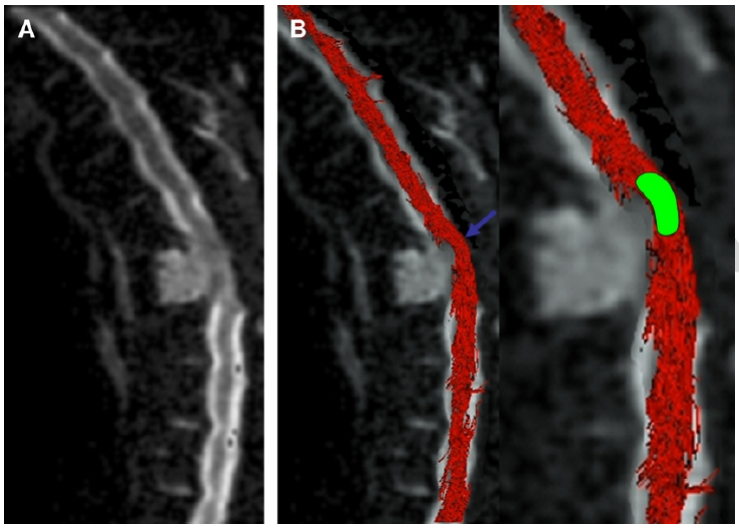


Figure 8.7: Spinal cord compression due to vertebral and epidural metastasis of a breast cancer. (A) Midsagittal b 5 0 image is shown on the left. (B) Color-coded FA map (red) shows a mass-effect on fiber tracts (arrow) but preserved anisotropy in this patient, who had pain but no neurologic symptoms.

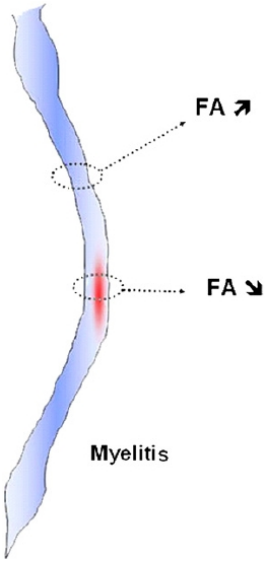


Figure 8.8: Diagram illustrating FA variations seen in inflammatory myelitis. The centers of these lesions show decreased FA values, and their boundaries show increased FA. This pattern may be specific to myelitis.

8.2.4 Myelitis

DT imaging is more sensitive than regular T2-weighted imaging in detecting spinal cord inflammatory lesions. FA maps and FT performed on patients who have suspected myelitis show lesions that are not seen on conventional T2 imaging. In one series, patients who had idiopathic myelitis had one (100%) or multiple (80%) areas of decreased FA regardless of the appearance of the cord at those levels on T2-weighted imaging [Renoux 2006]. Increased FA may be seen in the periphery of the zone of myelitis and in normal-appearing T2-weighted areas. The latter zones tend to be asymptomatic; the significance of this finding is uncertain. Thus, an inflammatory myelitis is characterized by decreased FA values in the region of the T2-weighted

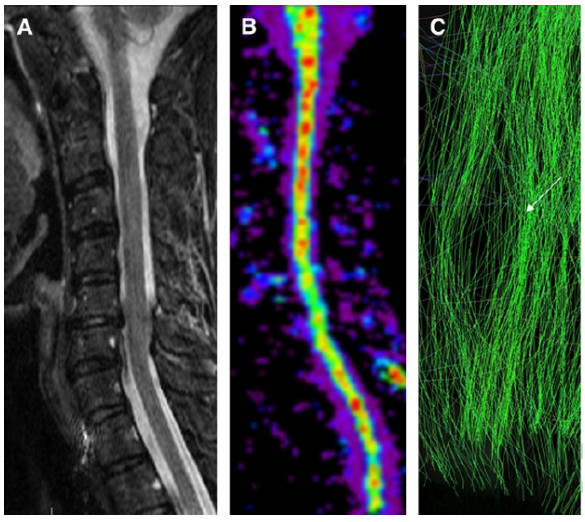


Figure 8.9: Idiopathic inflammatory cervical myelopathy (presumed diagnosis). (A) Midsagittal T2 image shows faint hyperintensity and expansion of the cord at the C5-C6 levels. (B) Corresponding directionality map shows loss of anisotropy greater than the size of the lesion seen on the T2 image. (C) FT map shows splaying and loss of fibers greater than that seen in localized tumors and due to the ill-defined nature of this inflammatory process, which resolves with antiinflammatory drugs.

lesion and increased FA values in the lesion's boundaries (Fig. 8.8). This pattern is different from that seen in invasive tumors, in which FA is low in peripheral regions of edema. Additionally, FT shows that inflammatory lesions spread the fibers of the spinal cord in areas that have an abnormal T2 signal; this pattern may be related to a decrease in extracellular water due to cytotoxic edema, axonal cluster regeneration, or cellular infiltration by inflammatory cells [Cassol 2004, Ciccarelli 2003]. This pattern is not seen in invasive tumors and may be an important marker of inflammatory lesions. Occasionally, fiber thinning is seen in inflammatory lesions and may be due to early axonal involvement (Fig. 8.9).

8.2.5 Arteriovenous Malformations

In spinal cord arteriovenous malformations, DT imaging with FA measurements may help to better understand their pathophysiology [Ozanne 2007, Stein 1992]. Additionally, FA values may improve after embolization and correlate with a better patient outcome. FT shows that at the level of the arteriovenous malformation nidus, the tracts are spread, shifted without spreading, interrupted, or normal. FT shows no fibers running through the nidus, an observation that may become important if surgical resection is contemplated. Distant to the nidus, congestive edema or a "cavitation" pattern may be found. In congestive edema, FT shows spreading of fascicles with global enlargement of the beam of tracts. There can also be a slight rarefaction of tracts, with decreased FA but normal ADC values. In cavitation, FT shows a loss of tracts with global thinning of their beam. Damage due to cavitation may be irreversible. In segments of the cord distant to the nidus without T2-weighted hyperintensity but where draining veins are present, FA values are slightly decreased when compared with segments where no draining veins are present, implying abnormal congestion at a microscopic level and perhaps explaining patients' symptoms when the level of edema seen on T2-weighted images does not match clinical deficits (Fig. 8.10).

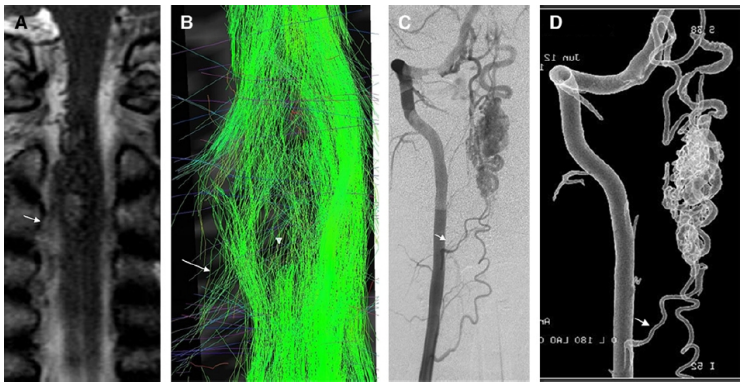


Figure 8.10: Cervical spinal cord arteriovenous malformations in a 23 y.o. man who had sensory symptoms in the right arm of 4 years duration. (A) Coronal T2 image shows an intramedullary lesion at C3-C4 with an associated enlarged blood vessel laterally (arrow). (B) FT map shows segmental interruption of the right posterior fascicle (arrowhead) at the level of the inferior part of the arteriovenous malformation nidus, which also splays the fibers. Lateral fibers are interrupted, and nerve root fibers are thin (arrow). Frontal views of catheter angiogram (C) and three-dimensional reformation (D) from the catheter angiogram show the malformation. The arrow points to an inferior feeding artery that presumably resulted in alteration of nerve root fibers shown on the FT map. DT imaging findings suggest that the symptoms are associated with tract damage and that recovery of neurologic function may not occur after treatment. The zone of fiber disruption may be used for surgical approach to avoid further damage of intact fibers elsewhere.

8.2.6 Metabolic Disorders

MELAS (mitochondrial encephalopathy, lactic acidosis, and stroke-like events) is a disorder that affects the brain and spinal cord. Patients who have spinal cord involvement have a worse prognosis. On the T2-weighted and FLAIR images, multiple, abnormal, high-intensity signal lesions may be seen in the mesencephalon, medulla oblongata, cerebellum, and cervical spinal cord. FA values are decreased within the spinal cord of these patients even when T2 abnormalities are not obvious [Ducreux 2005]. FT in the cervical spinal cord is not sensitive enough to detect abnormalities in MELAS patients. In MELAS, the extracellular water may not alter the shape of the tracts because the edema is not severe enough to warp them (Fig. 8.11) [Ducreux 2005].

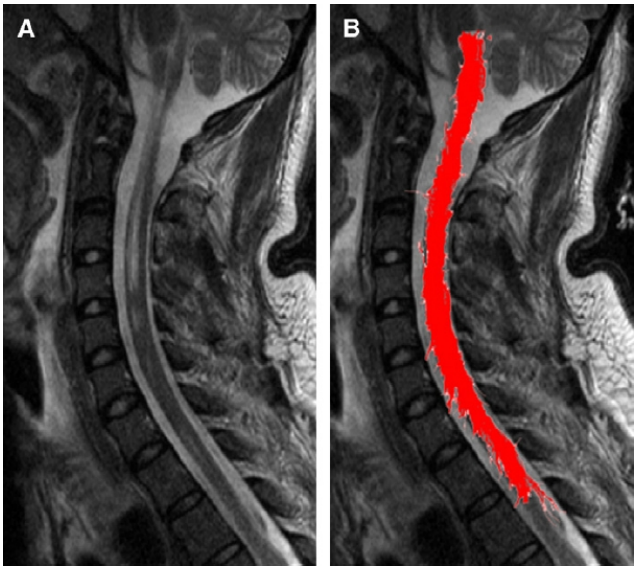


Figure 8.11: A patient who has MELAS and spinal cord involvement. (A) Midsagittal T2 image shows the central area of high signal in the cord from C2-C4. (B) FT map coregistered and projected on T2-weighted image shows an unaltered shape of the white matter tracts. In this disease, accumulation of extracellular water does not warp the fiber.

8.2.7 Syringomyelia

DT imaging may be useful to investigate syringomyelia. A syrinx may involve the spinothalamic tracts and result in temperature and sensory deficits and pain not directly related to the site and size of the lesion. Preliminary work seems to indicate that detailed anatomic evaluations in these patients are possible using FT. FT is useful in identifying the spinothalamic tracts in patients who have syrinx. We hypothesize that if the tracts are present but displaced by the lesion, the patients will have a better prognosis than when FT shows thinning or destruction of these tracts (Fig. 8.11). Identification of these tracts may be useful to the surgeon before placement of electrodes to control pain (Fig. 8.12).

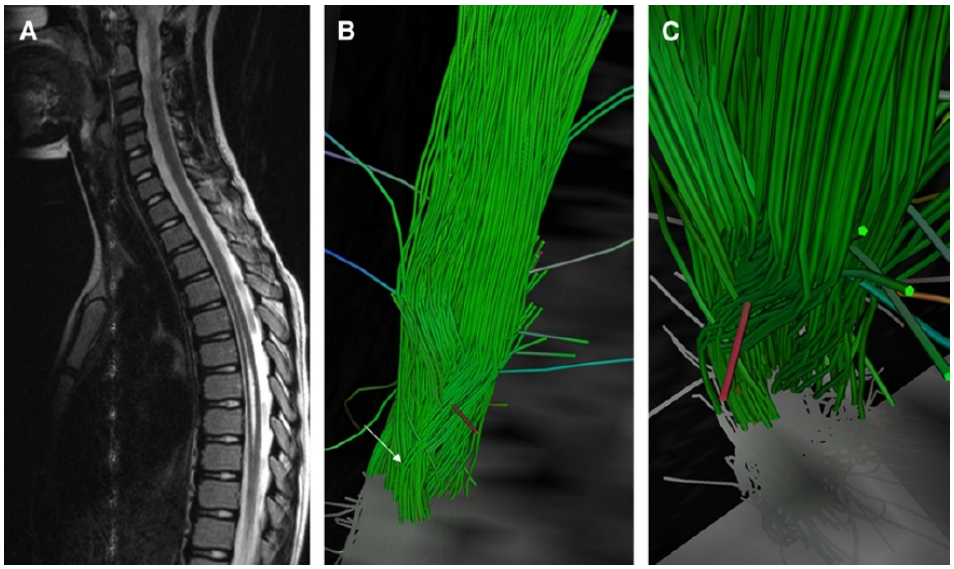


Figure 8.12: Fluid-filled cavity in the thoracic spinal cord. (A) Midsagittal T2 image shows syrinx in the lower thoracic cord. (B) FT map shows that the syrinx alters the shape of white matter fibers especially the spinothalamic tracts (arrow), which may explain symptoms in some patients. (C) Magnified FT map demonstrates the decussating tracts warped by the cavitation (arrow). Objective warping or tracts may lead to decompression even in patients who have small syrinxes.

8.2.8 Spinal Cord Injuries

Patients who have had prior spinal cord injuries may benefit from DT imaging and FT. Preliminary results after peripheral nerve grafting seem promising (Tadie, personal communication, 2002). Patients treated this way experienced a 40% motor recovery. Our early experience shows that FT may be used to assure the anatomic presence of intact fibers, a factor needed for successful grafting (Fig. 8.13). Fibers destroyed by the initial injury or secondarily to Wallerian degeneration do not respond to grafting.

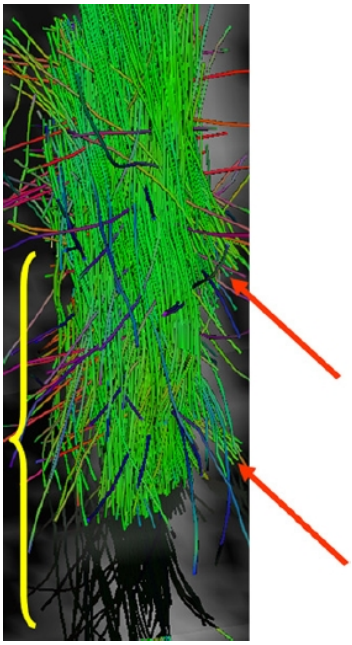


Figure 8.13: A patient who has paraplegia after a spinal cord injury. Frontal view of the FT map shows a relative absence of right sided-fibers (yellow bracket) when compared with the opposite side due to Wallerian degeneration. Note the normal nerve roots (arrows), which are not identified in the area of Wallerian degeneration. The nerve roots on the opposite side could be used for a nerve bypass to the remaining functional spinal cord.

8.3 Conclusions

DT imaging and FA may be more sensitive than other conventional MR imaging techniques to detect, characterize and map the extent of spinal cord lesions. Moreover, Fiber tracking offers the possibility of visualizing integrity of white matter tracts surrounding some lesions, and this indirect information may help in formulating a differential diagnosis and in planning biopsies or resection. FA measurements may also play a role in predicting the outcome of patients who have spinal cord lesions.

We have shown through this study that the diffusion tensor can effectively be used in clinical applications, and is still promised to a great future. The availability of free methods and softwares (like MedINRIA) helps clinicians analyze their own data. Conversely, methodological research is often driven by clinical applications. This symbiosis between these two worlds is the key for a fertile development of each

of them, and softwares appear as one of the possibility to help them communicate.

In the next chapter, we conclude the part with a summary of the contributions and a discussion on the future directions.

Chapter 9

Discussion

In this part, we have developed, implemented and applied in real conditions a set of methods for the processing of DT-MRI typical of clinical applications. In particular, the specific constraints of the clinical world (short scanning time, 1.5T scanners) lead us to develop a joint diffusion tensor estimation and smoothing procedure that is robust enough to handle a low number of encoding gradients (i.e., a low number of diffusion weighted images) and rather low signal-to-noise ratios. In order to be used by medical experts who are not necessarily strong computer scientists, we believe that such method should be embedded in easy-to-use softwares. For this purpose, we have developed MedINRIA, a collection of medical image processing software packages dedicated for use by clinicians. It includes applications for 3D image visualization, a complete processing pipeline for DT-MRI analysis, a tensor visualizer for quality control of diffusion tensor fields, and an image registration toolkit for image alignment. Finally, we have demonstrated the use of MedINRIA with a survey of possible applications of spinal cord DT-MRI. We have shown that DTI could play a role in characterizing tumors, detecting compressions and inflammatory lesions, or as a follow up in nerve grafting.

In the future, the questions of validation and reproducibility of our joint diffusion tensor estimation and smoothing procedure have to be answered. For this purpose, we could think of repeating scans of the same patient in various orientations in the scanner, and in various scanners. One also could think of using phantoms or histological data as in [Perrin 2005, Savadjiev 2006]. Furthermore, the observed qualitative impact on the tracking could be quantified, using for instance a dispersion measure of the fibers. In addition, this estimation method should be integrated into MedINRIA. Several steps have to be done prior to its integration: first, one needs to optimize its implementation to make its execution as quick as possible. Second, parameters (like noise variance, amount of smoothing) should be set automatically, and this task is far from being straightforward (some groups are specifically working on it, in particular for the estimation of the noise variance [Aja-Fernández 2007]). Finally, the practical input of DT-MRI in clinical applications should be further investigated with other clinical studies. We showed how DT-MRI could be used in spinal cord imaging, but most applications are studying the brain: DTI could be used to characterize tumors, to follow evolving pathologies like multiple sclerosis, or better understand neurological pathologies like autism, schizophrenia or depression. Further applications using DTI include the modeling of tumor growth: there are some evidence that tumors grow principally

in the directions of the neural fibers and disseminate cancer cells in these directions [Clatz 2005, Konukoglu 2007]. This would be extremely useful in radiotherapy to more precisely delineate a tumor along with the regions where potential cancer cells are located, even if they cannot be seen with MRI.

One last application apart from the clinical world is the modeling of the white matter fiber architecture. Modeling white matter fibers is important 1) to better understand brain and brain maturation (which implies a modeling at different ages) and 2) to be able to perform group studies and consequently understand effect of pathologies on these fibers. This is a computational anatomy problem which aims at evaluating what a normal fiber architecture is in terms of average representation and variability around this average within a given population. One would need first to put diffusion weighted images of different patients in correspondences, which raises the question of non-linear registration of DTI. This field is very active, and promising work include [Goodlett 2006, Ziyan 2007, Zhang 2007b]. Moreover, we believe that DT-MRI could be used to improve registration of white matter in T_1 images for instance, as shown in [Studholme 2007]: white matter in T_1 images appears uniformly white, while DT-MRI reveals structures (fibers) which could drive the registration. Second, we need to be able to reliably extract corresponding fiber bundles among patients. Fiber bundling is an active area of research, and several groups are heading towards that direction [O'Donnell 2006]. Finally, one should fuse all these fiber maps into a single average representation, and then determine its variability. Interesting work on this include [Corouge 2006, Maddah 2007, O'Donnel 2007].

The final part of this thesis is dedicated to the computational anatomy of the brain. Unlike the present part, we are not dealing with diffusion MRI and fiber tracts anymore, but we focus on building a second order model of the cortex. The cortex is the principal locus of neural activity (while fibers can be seen as the "links" between neurons), and is a very complex structure and variable among individuals. Thus, measuring and modeling its variability within a population is one step towards a more complete model of the brain, which is the main goal of this thesis.

Part III

Statistical Analysis of the Human Brain Cortex Anatomy

Chapter 10

Introduction

Contents

"Aristotle was famous for knowing everything. He taught that the brain exists merely to cool the blood and is not involved in the process of thinking. This is true only of certain persons."
Will Cuppy.

Brain structures differ greatly in shape and size even among normal subjects (see e.g., Fig. 10.1), and these variations make it difficult to identify abnormal differences due to disease. Understanding the degree and quality of brain variations is vital for distinguishing early signs of disease from normal variations. Neuroscientists are also interested in identifying the causes of brain variability at a genetic or environmental level. An efficient, parsimonious model of the complex patterns of brain variation would help in identifying factors that contribute to it. Furthermore, measuring brain asymmetry (i.e., differences between hemispheres) is of special interest as it sheds light on how the functions of the two hemispheres become specialized [Toga 2003]. Improved modeling of the range of variations in brain structure could make it easier to isolate specific effects of genetic polymorphisms on these normal variations and asymmetries [Cannon 2005, Geschwind 2002, Thompson 2001a]. Finally, geometric variability of anatomy also makes the automated segmentation and labeling of brain structures difficult. Statistical information on brain variability would make this task easier [Pitiot 2004, Fischl 2002], and could be used in Bayesian approaches for inter-subject nonlinear registration ([Mangin 2004a, Gee 1998, Ashburner 2005]) in order to adjust for anatomical variations across subjects prior to group analysis of brain function or metabolism.

10.1 Definition of Brain Variability

A major class of anatomical variations can be thought of as arising from the smooth deformation of a reference anatomy, where the deformation is represented

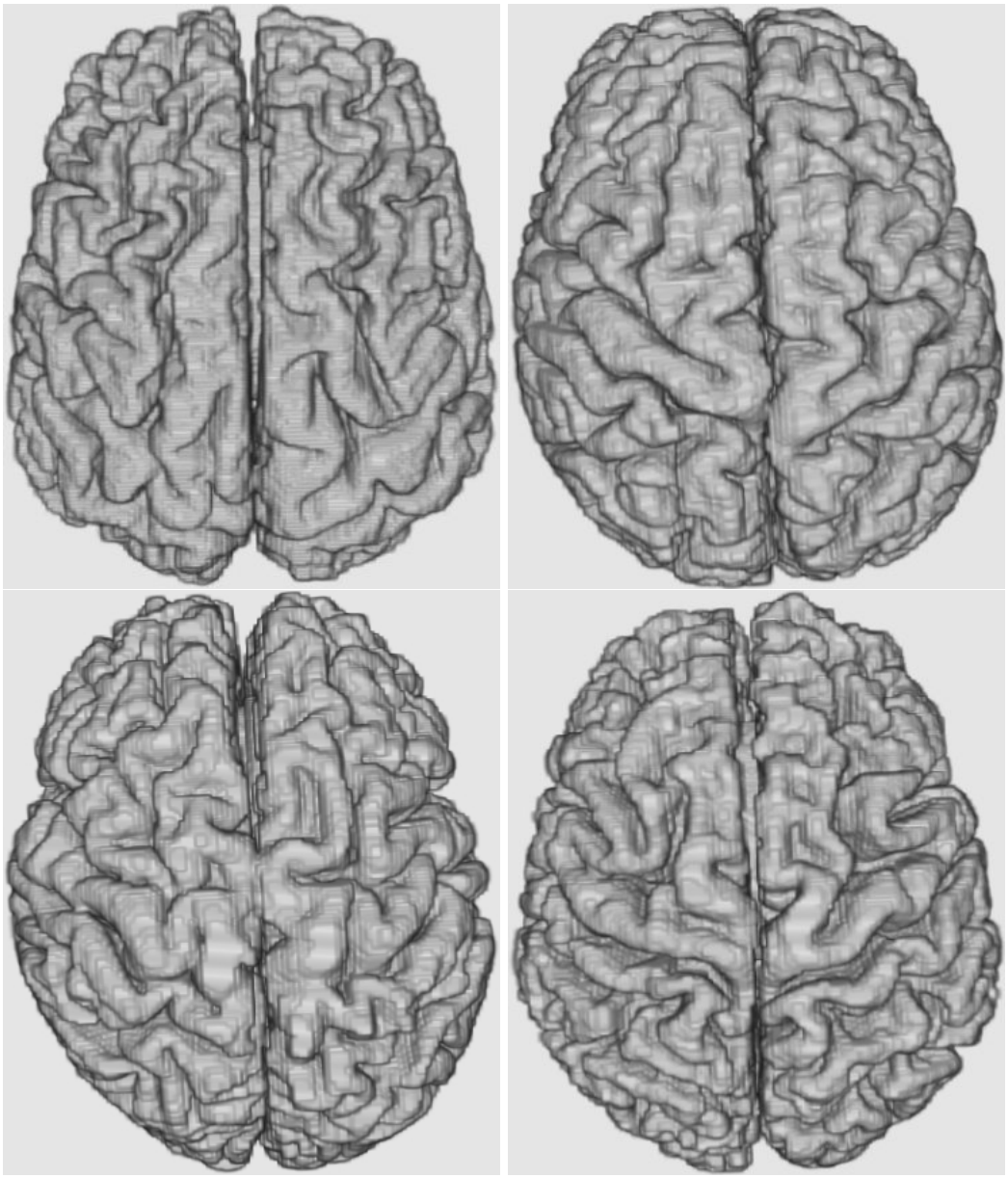


Figure 10.1: **The cortical surfaces of four different subjects.** It is hard to detect common patterns, and consequently to detect abnormalities, when looking at cortical surfaces. Computational anatomy is a field whose goal is to model organs like the brain to be able to distinguish automatically between normal controls and pathological subjects.

as a 3D displacement field, after affine differences are factored out [Roche 2000]. These variations can be measured via some statistical tools, and in particular the covariance matrix. This matrix tells not only how much an anatomical position varies (i.e., a quantitative value in centimeters), but also in which direction it

varies the most (there is some evidence that structural variation is greatest along certain preferred directions [Thompson 2001b]). Having this information at every anatomical position is what we call a *second order model of brain variability* (by analogy, a first-order model would be the atlas, or the "mean anatomy").

To go further in the modeling of brain variability, we could look at the possible joint variability of all pairs of points to see how the displacement of one point (in a population) with respect to the reference anatomy covaries with the displacement of neighboring or distant points in the brain (e.g., symmetric ones in the opposite hemisphere). In other words, measuring the variability of each individual point might not be sufficient since it does not reveal a potential correlation between two positions. These correlations may, however, reveal crucial information on the relationships between brain structures. This information is complementary to the variability, and could be seen as a new measure of brain connectivity.

Before going into the main material of this part, there is something important that we have not discussed: the data. The choice of the type of data (and their dimensionality) for this study is not straightforward and data can take various forms, like images, surfaces or lines. In the next section, we discuss what type of data can be considered for this work, and we motivate our choice.

10.2 What Type of Data do we Need?

One initial idea for measuring inter-subject brain variability might be to gather statistics on 3D displacement fields computed between a reference anatomy and many individuals. Such data could be obtained using an inter-subject registration algorithm. However, we would need to estimate the influence of the chosen registration method, as well as the influence of the target chosen for registration. Furthermore, the image intensity only constrains the registration in the direction of the image gradient. In other words, matching is generally uncertain in directions parallel to edges. Only by adding some constraints of regularity on the displacement field, one can obtain a smooth deformation field and hope to recover part of the missing displacement components. To provide information that is completely independent of volumetric image registration algorithms, we chose in this work to rely on lower dimensional structures, such as cortical landmarks identified by expert neuroscientists following a formalized protocol, with known inter- and intra-rater reliability. We could have taken surfaces, e.g., the hippocampus, corpus callosum, or even the whole cortex as in [Thompson 2000]. However, even for surface-based deformations, we would still face the problem of finding appropriate correspondences between surfaces (this is currently an active area of research, especially for the brain [Pitiot 2003, Tosun 2005, Wang 2005, Vaillant 2005, Vaillant 2006]). Moreover, these surfaces may be difficult to extract accurately. Thus, we chose to focus on anatomically well defined 3D curves that could be manually delineated by neu-

roanatomists and considered as ground truth data. This choice naturally led us to the primary anatomical landmarks on the cortex: the sulci. A large number of of sulcal landmarks consistently appear in all normal individuals and allow a consistent subdivision of the cortex into major lobes and gyri [Mangin 2004b]. In the absence of individual functional imaging data, sulci also provide an approximate guide to the functional subdivisions of the cortex, for all of the lobes. They are also used to guide intersubject functional activations registration [Corouge 2003].

The LONI Sulcal Lines: We use a dataset of sulcal lines manually delineated in 98 subject images by expert neuroanatomists according to a precise protocol[1] (an example of tracing is shown in Fig. 10.2). The dataset consists of 47 men and 53 women (age: 51.8 +/- 6.2 years), all normal controls. The lines are traced in 3D on the cortical surface, using an interface that allows curves to be traced interactively on surfaces. In Fig. 10.2, the sulcal lines appear to be traced out as curved lines in the intrasulcal CSF, near the exterior of the cortex. This is because the cortical surfaces we use are regularized and they clip across the sulcal CSF. The cortical surfaces are extracted with an active surface algorithm [MacDonald 1998] that creates a surface whose geometry is well adapted to a single intensity isovalue in the image while at the same time minimizing a curvature penalty that prevents complete penetration of the surface into the sulci. As such the sulcal lines are more like the gyral divisions on the exterior cortical hull as seen in the classical cytoarchitectonic maps of Brodmann [Brodmann 1905]. Using these as landmarks in prior studies for the normalization of fMRI and other cortical signals [Thompson 2004, Sowell 2004] was shown to improve the registration of gyral crests although it has to be admitted that the surface does not represent the full depths of the cortical sulci. In the following, we abusively call these sulcal lines *sulci* to simplify the description.

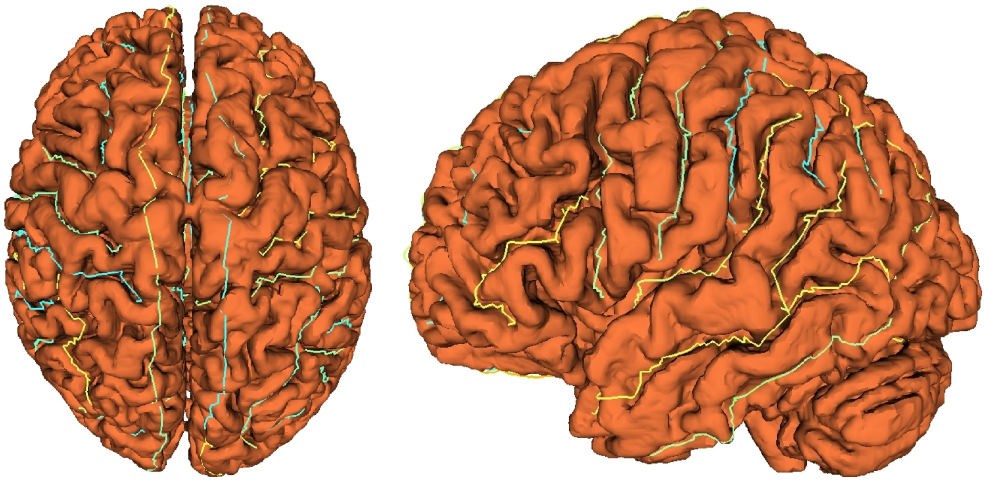


Figure 10.2: **Example of sulcal lines drawn on the cortical surface.**

[1] http://www.loni.ucla.edu/~khayashi/Public/medial_surface/

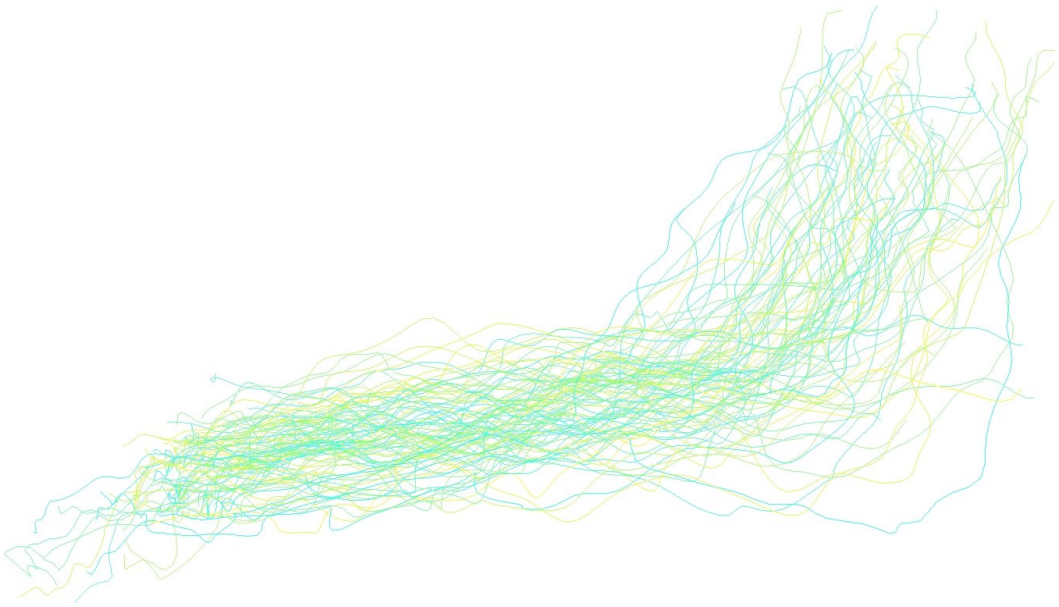


Figure 10.3: **The Sylvian Fissure. Lines of 98 subjects are displayed in two colors.**

We included the maximal subset of all sulcal curves that consistently appear in all normal subjects, 72 in total. Image analysts blind to subject sex and age traced the sulci on the lateral brain surface (including the Sylvian fissure, and central, pre-central, post-central, superior temporal sulcus (STS) main body, STS ascending branch, STS posterior branch, primary intermediate sulcus, secondary intermediate sulcus, inferior temporal, superior frontal, inferior frontal, intraparietal, transverse occipital, olfactory, occipito-temporal, and collateral sulci) in each hemisphere on the surface of each subject's brain. An additional set of sulci were outlined on interhemispheric surfaces (including the callosal sulcus, inferior callosal outline, superior rostral sulcus, inferior rostral sulcus, paracentral sulcus, anterior and posterior segments of the cingulate sulcus, outer segment double parallel cingulate sulcus when present, parieto-occipital sulcus, anterior and posterior segments of the calcarine sulcus, and the subparietal sulcus). In addition to contouring the major sulci, a set of 6 midline landmark curves bordering the longitudinal fissure were outlined in each hemisphere to establish hemispheric gyral limits. Spatially registered gray-scale image volumes in coronal, axial, and sagittal planes were available simultaneously to help disambiguate brain anatomy. Detailed criteria have been developed for delineating the cortical lines, and for starting and stopping points for each sulcus using brain surface atlases as references. By repeated training on test sets of brain images, the maximum observed inter- and intra-rater error (reliability) was ensured to be better than 2mm everywhere, in terms of r.m.s. distance, and in most regions less than 1mm, far less than the intersubject

anatomical variance. MR images used for delineations were linearly aligned to the ICBM stereotactic space [Collins 1995], thus providing a common coordinate system for all traced curves.

The rest of the part is organized as follows. In Chapter 11, we show how to create a second-order model of brain variability from the LONI sulcal lines. We also perform statistical tests, including leave-one-(sulcus)-out test, to evaluate the predictive power of such model. In Chapter 12, we no longer focus on measuring the variability of each individual brain position, but we look for the joint variability of (i.e., correlations between) any pair of points. In particular, we study the correlations between two major sulcal lines and the rest of the brain. This first study leads us to explicitly target the correlations between symmetric positions of the brain,. We finally conclude in Chapter 13.

Chapter 11

Building a Second-Order Model of Cortical Variability

Contents

The second-order model of cortical variability as we define it consists in representing the variability of each individual point of the cortex via its covariance matrix, or *variability tensor*. We call this matrix a tensor by analogy to the diffusion tensor, which is nothing other than the covariance matrix of the Brownian motion of water molecules. Both matrices have exactly the same properties (real, symmetric and positive definite). This is the reason why we call such model a "second-order" model. Our strategy consists of three steps. First, the modeling of the sulcal lines, and the extraction of variability tensors (Sec. 11.1). Here, we define what an average sulcal line is, as well as the correspondences between the average line and each subject's instance. Then, we can compute covariance matrices along the mean curves. Second, the model simplification using tensor interpolation (Sec. 11.2). In this second step, a subset of meaningful (we will define what meaningful is) tensors are extracted to effectively create the 2^{nd} order model of variability: a sparse field of variability tensors. We show also how to extrapolate this sparse field to obtain a dense representation. Third, the evaluation of the model (Sec. 11.3). In this last section, we judge the predictive power of the model through several statistical tests, including a leave-one-(sulcus)-out test.

11.1 Learning Local Variability from Sulcal Lines

In this first section, we detail the methodology to model the sulcal curve, i.e., how we define the mean curves and the correspondences between the means and each instance, and how we extract variability tensors along them.

11.1.1 Sulcal Curve Modeling

Statistical models have frequently been constructed for objects such as open or closed curves and surfaces [Bookstein 1991, Cootes 1995, Trouve 2000, Paulsen 2003] to cite only a few examples. In each of these, the *aperture problem occurs*: as we do not have point landmarks, the point-to-point correspondences between instances of a surface or a curve cannot be recovered exactly. For instance, correspondences of two sulcal lines are intrinsically subject to error, with a greater uncertainty in the tangential direction than in the direction orthogonal to the sulci. In this work, we propose a one-to-one correspondence mapping that explicitly minimizes the amount of tangential variability. By doing this, we clearly underestimate the true variability (we only recover the normal component of the variability), but we don't introduce artificial variability due to erroneous correspondences.

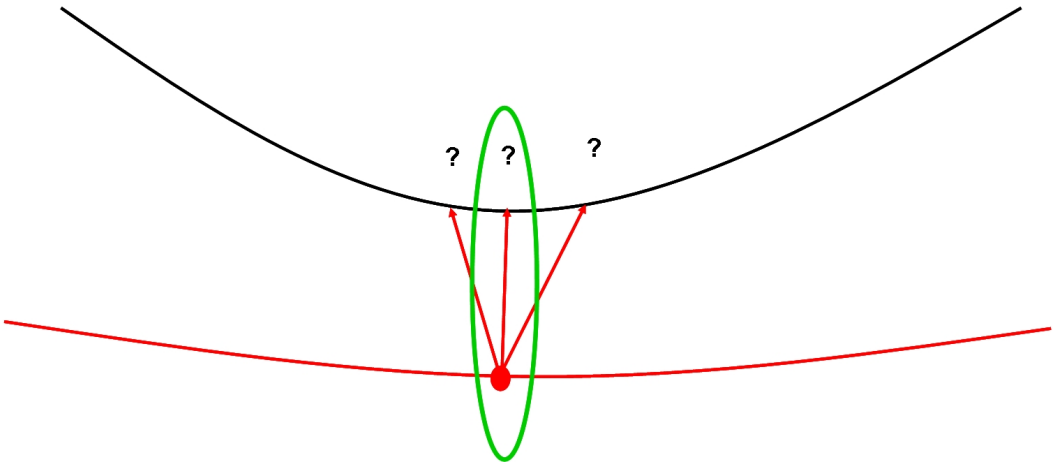


Figure 11.1: **Illustration of the aperture problem.** Matching the red curve onto the black one results in uncertainty in the tangential direction, when no extra information is available. Three possible matchings are highlighted with question marks, though an infinity is possible. In our sulcal matching procedure, we chose to minimize the impact of uncertain matchings by reducing the most possible the tangential component, which lead us to choose in this case the closest point (green circle).

First, we denoise the sample lines by approximating them with B-splines (Fig. 11.2): the manual sampling of 3-dimensional curves is only precise up to the voxel size (about 1mm^3), which is lower than the inter-rater variability of 2mm. In this

continuous setting, the number of degrees of freedom can be adjusted to increase robustness to noise while avoiding resampling problems [Baumberg 1994]. Typically, we reduce the number of control points to one third of the original sampling points, with a mean distance of 0.25mm and a maximum error of 2.7mm.

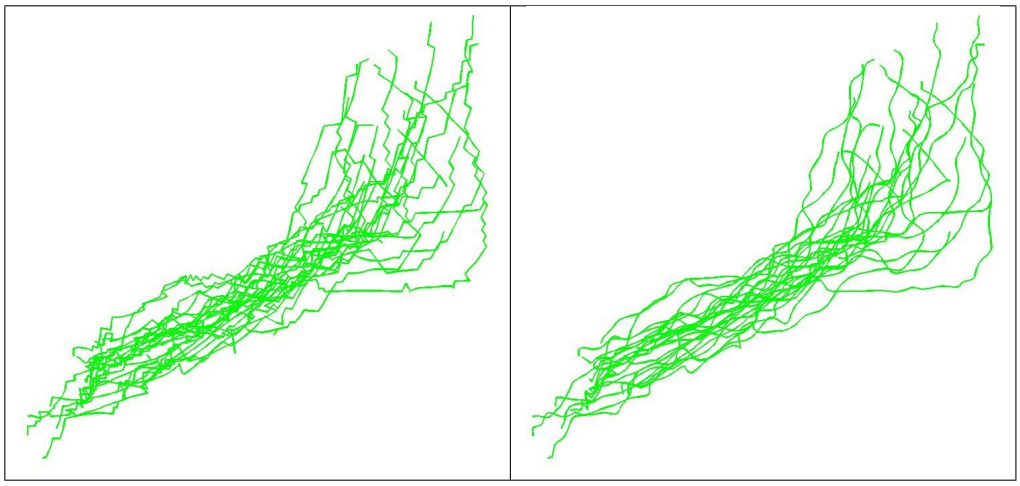


Figure 11.2: **Sulcal lines before and after B-Spline parameterization. Left:** Original lines from the LONI. The "stair" effect is due to sampling noise. **Right:** The same set of lines after our B-Splines parameterization. The high frequency noise is removed while keeping a very good approximation to the data. Moreover, the lines have now a continuous formulation thanks to the splines.

Many criteria have been proposed in the literature to evaluate the mean curve for a set of curves and to assess the appropriateness of one-to-one correspondences between geometric objects. They usually invoke local differential characteristics such as the tangent space, curvature, the local Frenet frame for a curve on a surface [Guéziec 1994, Bakircioglu 1998], regional shape information [Pitiot 2003]. In our case, the variability is so large (see e.g. Fig. 10.3), that using such refined measures is difficult. In general, sulcal curves do not have internal geometric features, along their length, that occur consistently from subject to subject. Therefore, we simply use the total variance of curve models as a criterion:

$$C(z) = \frac{1}{N-1} \sum_{k=1}^{N} \int_0^1 \|y^k(\phi^k(s)) - z(s)\|^2 \|z\prime(s)\| ds, \tag{11.1}$$

where y^k is the given sulcus of subject k, N the number of subjects, z the mean sulcus and ϕ^k the correspondence function between the subject's curve and the mean curve and s the curve length.

Minimizing this variance greatly reduces the variability due to inadequate correspondences. Practically, we alternately improve the correspondences between the mean curve and each sample by dynamic programming and optimize the average curve position by a first-order gradient descent with an adaptive step. This optimization strategy converges toward the mean curve after a few iterations

(dashed curve in Fig. 11.4).

For each of the 72 sulci, we end up with the mean curve $z(s)$, and one-to-one mappings $\phi^k(s)$ that give the corresponding position $y^k(\phi^k(s))$ in each subject k. The variability tensor $\Sigma(s)$ is then given by:

$$\Sigma(s) = \frac{1}{N-1} \sum_{k=1}^{N} \left[y^k \left(\phi^k(s) \right) - z(s) \right] \left[y^k \left(\phi^k(s) \right) - z(s) \right]^\top. \tag{11.2}$$

Results of covariance tensors estimated along the 72 sulci are shown in Fig. 11.3. A first and immediate remark is that the variability tensors looks much bigger at the extremities of the curves than in-between. This can be explained in two ways. A first hypothesis is that these points are actual landmarks identified by neuroanatomists (they know exactly where a sulcus starts and where it ends). During our sulcal matching procedure, we force the very first and end points of all sulci to correspond in order to initialize the dynamic programming procedure. Consequently, these points are the only locations where the aperture problem does not take place. Thus, they can carry an extra-information not available elsewhere: the variability tangential to the mean sulci. A second hypothesis, which is the one we are more confident with, is that the extreme points are very hard to identify: a sulcus may end because a gyrus is cutting it into two pieces, or there is a large uncertainty due to partial voluming, a wrong surface extraction at these location, etc. Consequently, these points exhibits a large variability due to a relatively poor accuracy of their localization. Although the truth is surely a trade-off between these two hypothesis, we chose in this work to remove the variability information at the extremities of the sulci from our model, and focus only on the inner parts of the sulci.

11.1.2 Estimation of the Affine Transformation from Correspondences

Initially, images were affinely registered onto a common reference image (in our case the ICBM305 space). To remove the influence of the chosen target and build unbiased atlases, one often needs to use more elaborate strategies, like in [Guimond 2000], or [Kochunov 2002]. With our curves, one can simply rely for each subject on the established correspondences between its sulci and the mean curves to refine the affine transformation. This will also further reduce the overall sulcal variability.

Let (A^k, t^k) be the affine transformation of subject k (A^k being the linear part, and t^k being the translation. Let us assume that the mean curves and the mappings between each subject's curve and the mean are known. We are looking for the optimal affine transformation $(A^k_{(opt)}, t^k_{(opt)})$ in the least-squares sense, i.e., the ones that minimizes the sum of the squared differences between the transformed sulci and the mean curves (as transformations are independent between different subjects, one

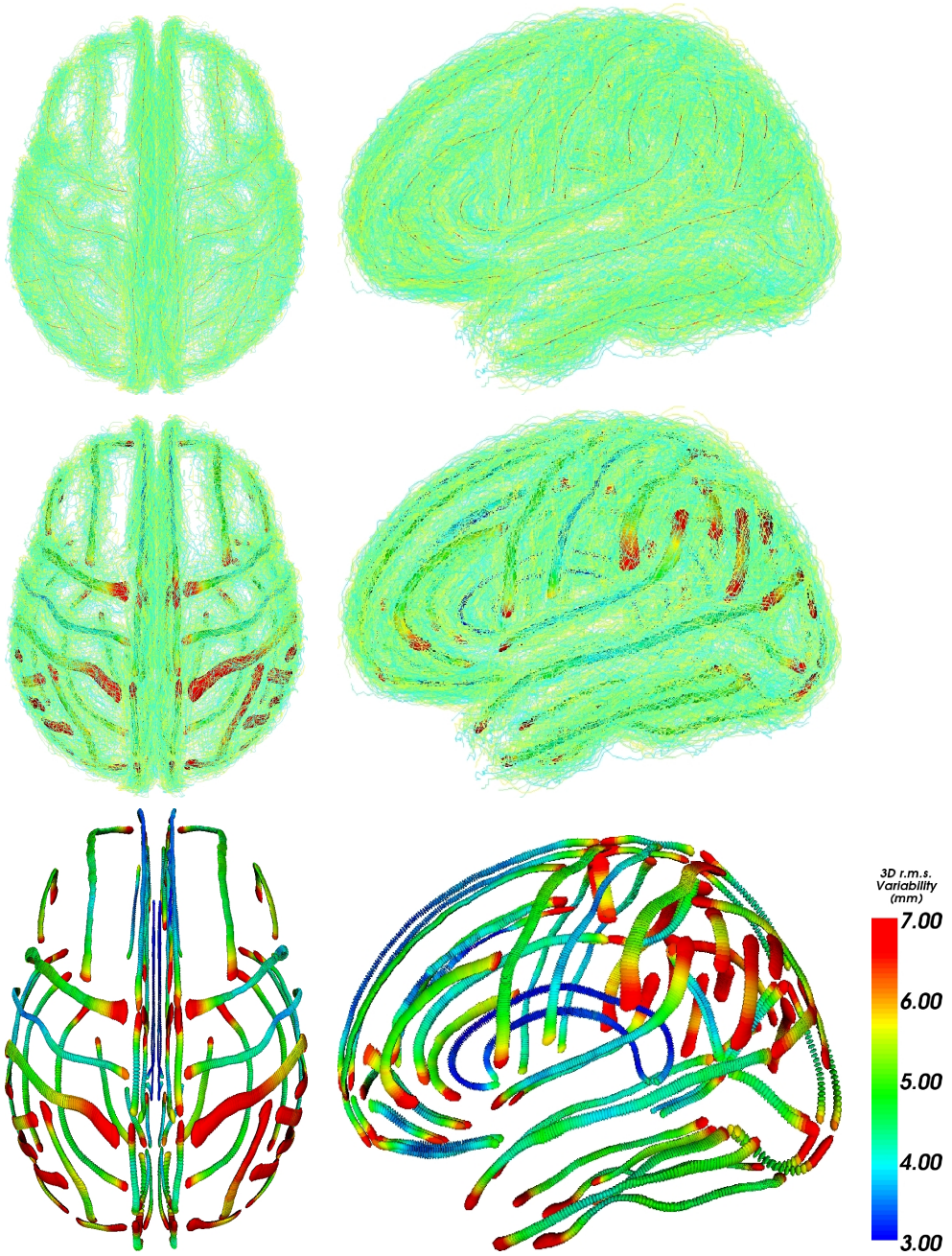


Figure 11.3: **Sulcal variability**. **Top:** The sulcal lines alone. Mean curves are shown in red, and traces from 98 healthy normal individuals are displayed in green and yellow. **Bottom:** Covariance matrices (ellipsoids at one standard deviation) are overlapped at regularly sampled spatial positions along the mean sulci. The color codes for the trace: Note that tensors close to the sulci extremities are larger.

can deal with each system at a time):

$$C(A^k, t^k) = \sum_{i=1}^{N} \int_0^1 \|A^k y_i^k \left(\phi_i^k(s)\right) + t^k - z_i(s)\|^2 ds \tag{11.3}$$

To optimize this energy, we first fix A^k and look for the optimal translation. We find that the minimum value is reached for:

$$t_{(opt)}^k = \bar{z} - A^k \bar{y}^k, \tag{11.4}$$

where $\bar{z} = 1/N \sum_{i=1}^{N} \int_0^1 z_i(s) ds$ and $\bar{y}^k = 1/N \sum_{i=1}^{N} \int_0^1 y_i^k(\phi_i^k(s)) ds$. Second, one introduces $t_{(opt)}^k$ back into Eq. 11.3. Calling $\tilde{y}_i^k = y_i^k - \bar{y}^k$ and $\tilde{z}_i = z_i - \bar{z}$, one can rewrite Eq. 11.3 as:

$$C(A^k) = \sum_{i=1}^{N} \int_0^1 \|A^k \tilde{y}_i^k(s) - \tilde{z}_i(s)\|^2 ds.$$

In Appendix B, we show that the solution is given by:

$$A_{(opt)}^k = \left(\sum_{i=1}^{N} \int_0^1 \tilde{z}_i(s) \tilde{y}_i^k(s)^\top ds \right) \left(\sum_{i=1}^{N} \int_0^1 \tilde{y}_i^k(s) \tilde{y}_i^k(s)^\top ds \right)^{-1}. \tag{11.5}$$

The overall framework now consists in alternating the positioning of the mean curve, the computation of the correspondence functions, and the evaluation of the optimal affine transformation for each subject using Eq. 11.4 and 11.5. Curves in Fig. 11.4 represent the energy of Eq. 11.1 averaged over the 72 sulci and for each iteration. The optimization of mean curves and correspondences reduces the amount of variability to 70% of the initial value. Reestimating the affine transformation further reduces the amount of variability to 60% of its initial value.

Figure 11.5 shows the resulting variability tensors computed using Eq. 11.2 and overlapped along the mean sulcus. One notices that the correspondence optimization greatly reduces the tangential components of the variability. The affine correction reduces the variability more globally.

11.2 Model Simplification and Extrapolation

We are now able to compute variability tensors at any continuous position along the 72 mean sulci of our datasets, thanks to the B-Spline parameterization and the correspondence mappings obtained in the previous section. The goal of this second step is to optimally choose a certain number of covariance matrices and their positions to simplify the sparse field of variability tensors. Then, we will show how to extrapolate this sparse field to obtain a dense representation, which will be considered as the actual second order model of brain variability. This will be used for visualization purposes and further analysis in a third step.

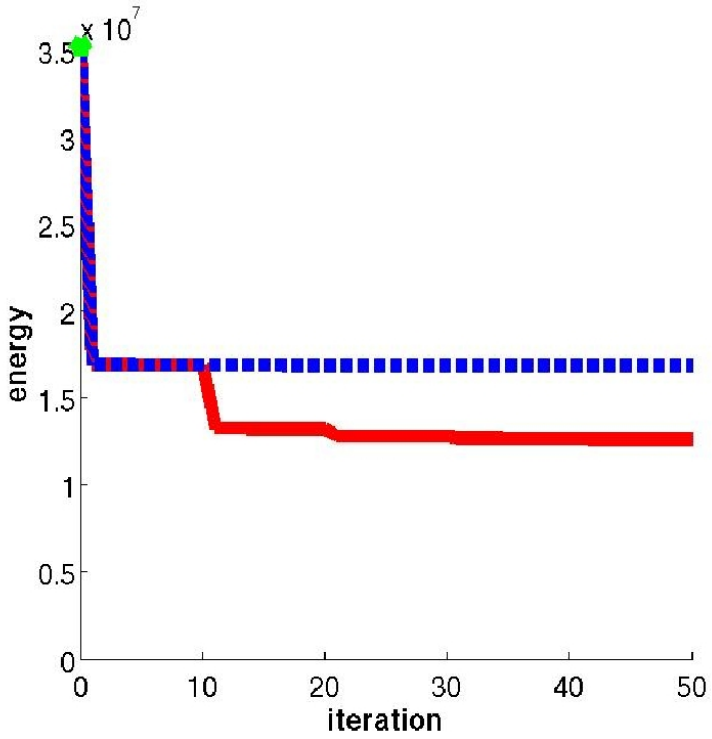


Figure 11.4: **Mean curves calculation: Energy vs. iterations. Dashed curve:** Without affine refinement. **Solid curve:** With affine refinement every 10 iterations. Note that the process has converged after 40 iterations.

11.2.1 Model Simplification using Tensor Interpolation

In the interior part of the sulci, tensors are highly regular in size and shape (see Fig. 11.3). Some of this information is therefore redundant and could be simplified by selecting only a few tensors at specific points along the mean sulcus, and interpolating in between them. Moreover, to avoid corrections for multiple testing problem, we had to downsample the representation of the continuous curves to a point where each tensor is independent of its neighbors. A second computational reason is that the tensor extrapolation process is very time consuming, and reducing the number of seed tensors decreases the computational time drastically. We use geodesic interpolation (see Sec. 4.2.1) to interpolate between successive tensors, because it preserves the monotone (i.e., consistently increasing or decreasing) evolution of the determinant. This is crucial when interpolating two probability densities and is not possible in general with standard Euclidean interpolation. For efficiency reasons, we also selected the tensor values among the observed data rather than optimizing them as free parameters. This operation has been automated in an algorithm called *tensor picking*.

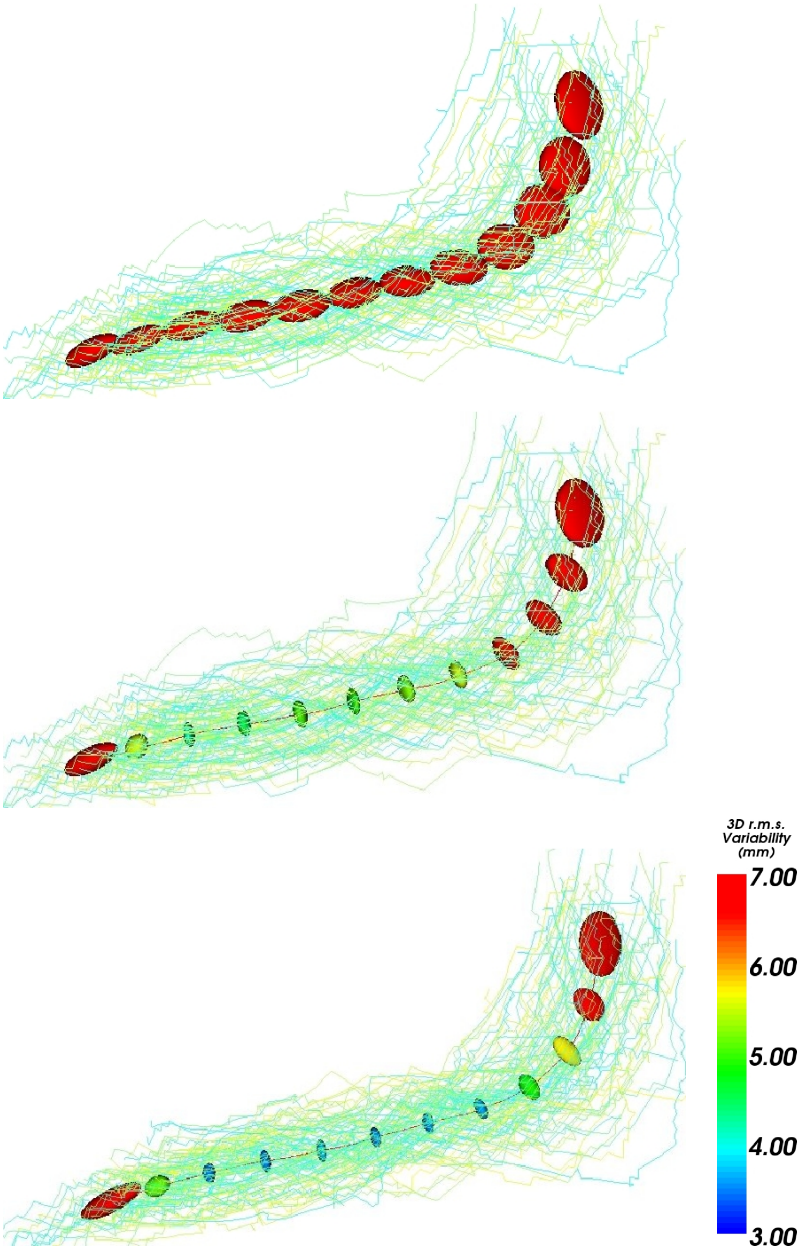


Figure 11.5: **Variability tensors for the Sylvian Fissure. Top:** Without correspondence optimization. **Middle:** After correspondence optimization. **Bottom:** After correspondence optimization and affine transformation optimization. Note that the curves get slightly more concentrated around the mean, showing less variability. This is reflected by the reduction of covariance matrices in the direction orthogonal to the mean curve. The reduction of variation in the tangential direction is mainly due to the optimization of correspondences between curves.

Tensor Picking: Let $(\Sigma_i)_{1 \leq i \leq N}$ be a set of N variability tensors defined at abscissa s_i along a mean sulcus. The geodesic interpolation between 2 successive

tensors is given by: $\tilde{\Sigma}(s) = \exp_{\Sigma_i}[(s-s_i)/(s_{i+1}-s_i)\log_{\Sigma_i}(\Sigma_{i+1})]$ for $s_i \leq s < s_{i+1}$. The mappings $\exp_{\Sigma_i}$ and $\log_{\Sigma_i}$ are the exponential and logarithmic maps, using either the affine-invariant (Sec. 3.3.4) or Log-Euclidean (Sec. 3.4) metrics. As we are working only on the interior of the sulcus, s takes its values between s_2 and s_{N-1} (we remove the extremities), so that the *interpolated variability* $\tilde{\Sigma}(s)$ is always defined. The tensor picking operation consists in finding the optimal set of N (Σ_i) such that the geodesic distance between the measured (initial) and interpolated variability tensors is minimized: $C(\Sigma_1,\ldots,\Sigma_N) = \int_{s_2}^{s_{N-1}} \mathrm{dist}^2\big(\Sigma(s),\tilde{\Sigma}(s)\big)ds$. The geodesic distance dist() can be either the affine-invariant or Log-Euclidean distance between tensors, but should be consistent with the choice of the metric for the geodesic interpolation. To minimize this energy, we first set N equal to 2. Then, an exhaustive search for the optimal set of N tensors is done. If the energy obtained is below a threshold t (set empirically to 0.2 in our experiments, see section 11.3.1 for an in-depth analysis of this parameter influence), the tensors are picked. Otherwise the number N is increased and the search is reiterated.

Results of this operation are presented in Fig. 11.6 (middle panel): by choosing tensors at adequate positions, one can qualitatively reconstruct the full variability of each sulcus using 4 to 10 covariance matrices, depending on its length and shape. The variability of all the major sulci can be adequately represented by 366 variability tensors only out of 2000 initially.

11.2.2 Extrapolating the Variability to the Full Brain

The next step consists in extrapolating these selected tensors to the full brain. Extrapolation produces a dense field of covariance matrices and is useful for two reasons. First, one can access the variability tensor at any continuous position of the brain. Obviously, modeling brain variability just by looking at a limited set of sulcal lines is not enough to infer the variability of deeper structure. If a greater number of structures were chosen (including subcortical structures such as the ventricles), we could better model the variability of the full brain. Second, it facilitates the exploration of the results by visually identifying large areas that vary similarly. To extrapolate the sparse field of tensors, we use the framework detailed in Sec. 4.2.3, which combines a radial basis function (RBF) interpolation and a diffusion PDE with a data attachment term. The RBF interpolation is used as initialization of the diffusion PDE, which converges in a very few iterations.

Fig. 11.6 presents the result extrapolating our 366 tensors on a discrete grid of size $91 \times 109 \times 91$ and with a spacing of $2\times2\times2\,mm^3$ (ICBM 305 space). We used the parameter values $\alpha = 20$ and $\gamma = 0.95$ for the RBF interpolation and $\sigma = 2$ for the discretization of the data attachment term in the extrapolation (Eq. (4.18)). We derived 2 scalar measures from the extrapolation. The first one (Fig. 11.7 left column) is a variability map given by the 3D root mean square (r.m.s.) of the trace of each covariance matrix: $rms = \sqrt{\mathrm{trace}(\Sigma(x))}$. One can see highly variable regions (such as the parietal cortex and Broca's area) with hot colors, and more stable areas (such as the primary sensorimotor cortex) with cold colors. Note

that the maximum allowed inter- and intra-rater error for the manual tracings of the curves was ensured to be less than 2mm everywhere, less than the inter-patient variability that is observed (the smallest value is about 3mm). The second map shows the principal direction of each tensor (i.e., the eigenvector associated to the largest eigenvalue), whose coordinates are mapped on the RGB sphere, as presented in Fig. 11.7 right column (same color as in DTI - Part II). This map confirms the anatomical intuition that there are sets of sulci in certain cortical areas that tend to vary in a consistent way in the top view, the principal direction of variation is lateral-to-medial for the superior frontal and parietal sulci, but the central and precentral sulci tend to vary more along an anterior-posterior direction. The temporal lobe sulci also tend to be consistent in varying with the same principal direction.

The spatial pattern of variability agrees with established neuroanatomical data. For instance, [Thompson 2000] computed the variability of the cortex surface in an independent normal sample (15 controls) using a non-linear surface registration algorithm. Fig. 11.8 compares his variability map with ours. Our model of variability presents the same high values in the temporo-parietal cortex (red and purple area, marked "A" in Fig. 11.8) and low values in the superior frontal gyrus (marked "B" in Fig. 11.8), Broca's area, and the lower limits of the primary sensorimotor cortices in the central and precentral gyri. Phylogenetically older areas (e.g., orbitofrontal cortex), and primary cortices that myelinate earliest during development (e.g., primary somatosensory and auditory cortex) exhibit the least variability. The planum parietale (marked "A" in Fig. 11.8) consistently shows the highest variance of any cortical area, consistent with the complex pattern of secondary fissures surrounding the supramarginal and angular gyri (the perisylvian language cortex). It is also reasonable that the temporo-parietal areas around the Sylvian fissures are the most variable: they specialize and develop in different ways in each hemisphere, and are also the most asymmetric in terms of gyral patterning and volumes [Toga 2003].

11.3 Model Evaluation

Evaluating our extrapolated variability model is a difficult. Obviously, using the information given by the sulci is not enough to infer the variability of the full brain, particularly in the depth the brain (e.g., in the white matter, ventricles and deep gray matter nuclei). Moreover, we have no ground truth in these areas to validate the predicted variability. Thus, we restrict the evaluation of the predictive power of our model to the places where we have enough data: on the cortex. The first idea is to see how well our interpolation and extrapolation models fit the observed variability along each sulcus. This yields a root mean square error (RMSE) assessing the fidelity of the approximation. Then, we can perform a "leave one sulcus out" test to see if a group of sulci can correctly predict the variability of another sulcus in their neighborhood. This would mean that the model could effectively find missing data (i.e., the measures are dependent) and somehow predict the variability of missing structures in our datasets.

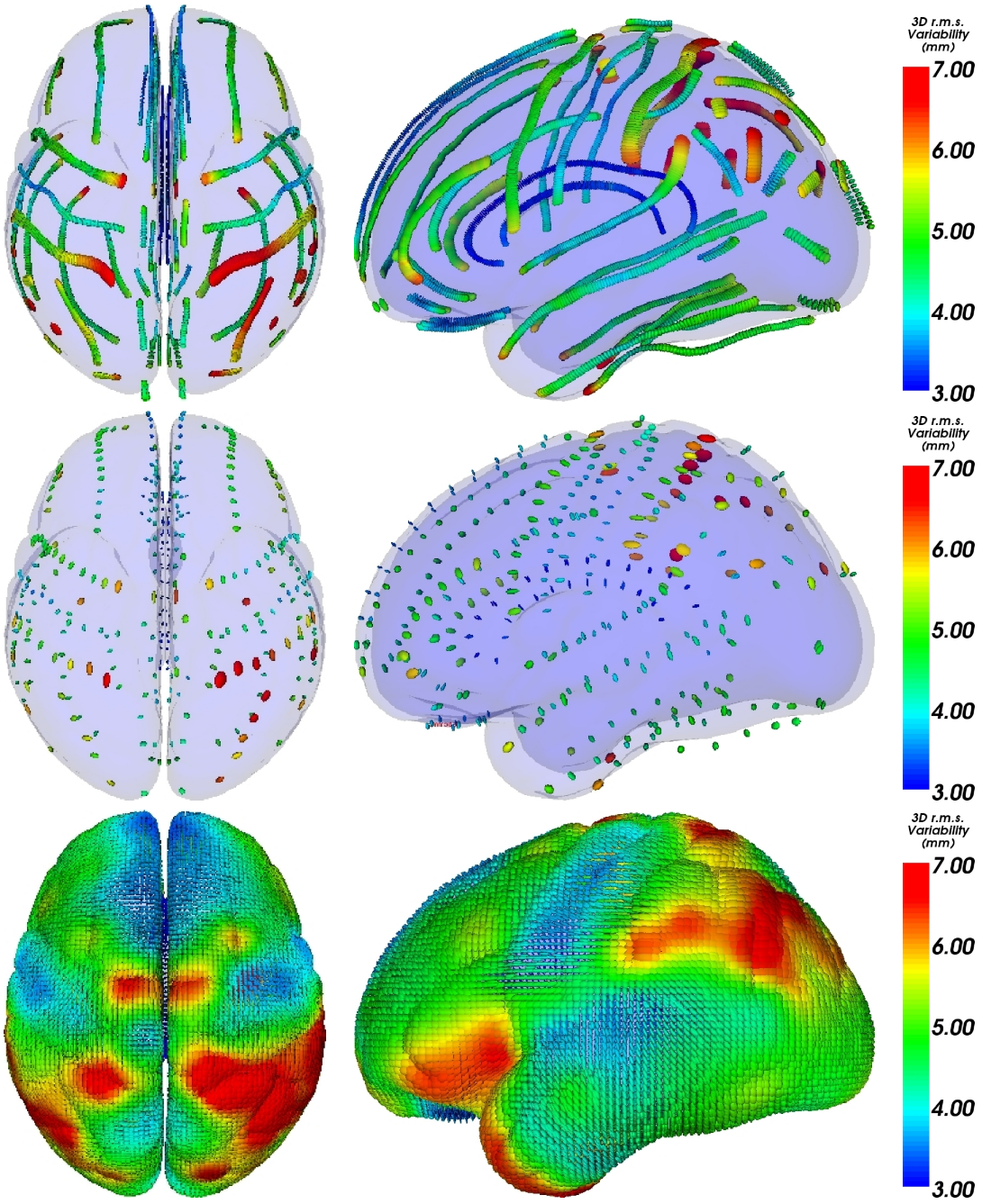


Figure 11.6: **Accessing the full map of cortical surface variability step by step**. **Top:** Covariance matrices calculated along mean sulci. **Middle:** Matrices selected by the tensor picking operation. **Bottom:** Result of the extrapolation.

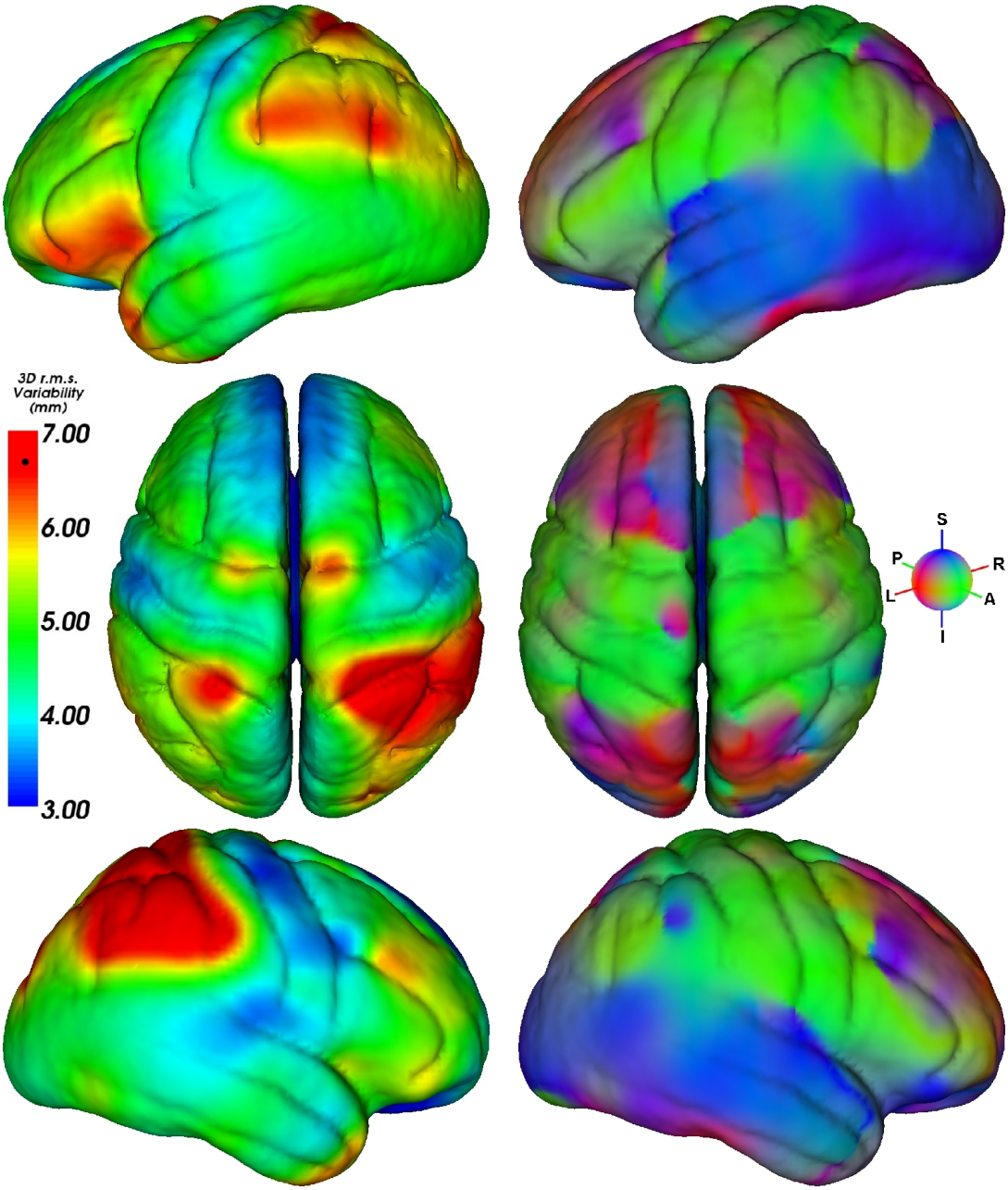


Figure 11.7: **Left column: Variability map derived from the dense variability tensor field obtained by extrapolation.** The color codes for the 3D rms variability (mm). Hot colors mean high variations among subjects. **Right column: Maps showing the main direction of variability.** The color codes for the main direction of the variability tensor at each point. Red: left-right oriented tensor, Green: posterior-anterior oriented, Blue: inferior-superior oriented.

11.3.1 Intra-Sulcus Variability Recovery

At each sampling point x of the sulcal curves, we computed the "difference" vector $Z(x)$ between the observed variability tensor (the reference $\Sigma(x)$) and the tensor in-

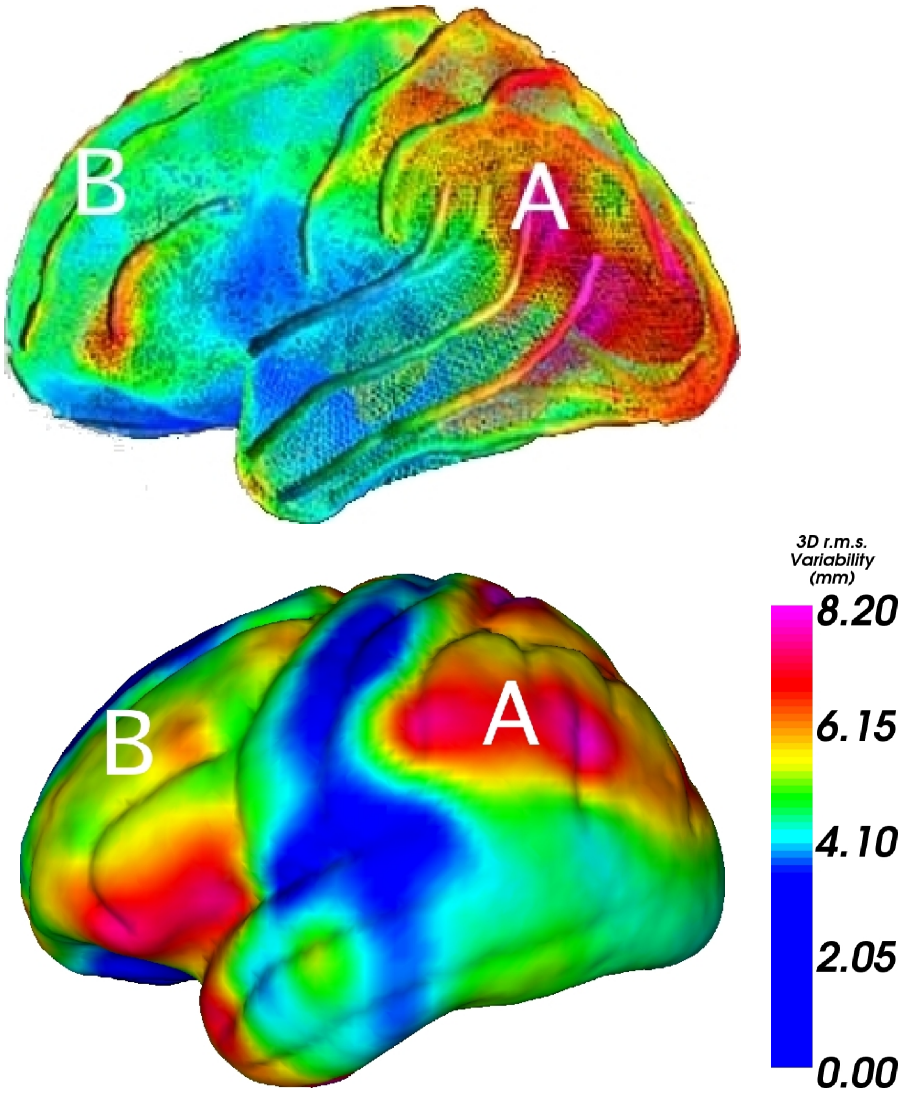


Figure 11.8: **Comparison of two independent models of brain variability.** The scalar value mapped on the mean cortex is the Trace of the tensors (the variance). **Left:** Cortical variability map from [Thompson 2000]. **Right:** Extrapolation of our simplified sulci variability model to the full brain. Note the similarity in the temporo-parietal cortex [shown in red colors (A)] and the superior frontal gyrus (B).

ferred by our model ($\tilde{\Sigma}(x)$), either using our interpolation (interpolation in-between two consecutive tensors) or our extrapolation method:

$$Z(x) = \log\left(\tilde{\Sigma}(x)\right) - \log\left(\Sigma(x)\right).$$

$Z(x)$ is expressed in the Log-Euclidean framework. For completeness, we also give the affine-invariant expression, though we won't use it in the following due to its higher complexity (and our experience shows that result are almost identical with

the LE formulation):

$$Z(x) = \log\left(\tilde{\Sigma}(x)^{-1/2}\Sigma(x)\tilde{\Sigma}(x)^{-1/2}\right).$$

For all mean sulcal positions x_i, we obtain two difference vectors $Z_i = Z(x_i)$: one obtained by interpolation and one by extrapolation. Then, we tested if the set of vectors Z_i had a null mean by using Hotelling's T_2 test. This test is the multivariate extension of the Student's t-test. Its aim is to test the significance of the mean estimation under the normal assumption. For both strategies, we found that the mean error was not significantly different from zero (p-value of 0.25), which means that using either interpolation or extrapolation does not induce a bias in the reconstruction.

Second, we found a standard deviation: $\sigma_{ref} = \sqrt{1/N \sum_{i=1}^{N} \|Z_i\|_2^2}$ of 0.15 for the interpolation error. This value gives us a lower bound on the range of the reconstruction error, and this is why we call it the reference error: an error of reconstruction (i.e., norm of the difference vector) lower than $3\sigma_{ref}$ we will be considered as not significant. The extrapolation gives a slightly higher standard deviation of 0.21. It is mainly attributable to the aperture problem: in regions with orthogonal sulci, the normal component of one tensor influences the tangential part of its perpendicular neighbors and vice versa, and the reconstruction error will be greater.

To further illustrate the prediction power of our extrapolation method, we compared the results of the extrapolation using 2000, 1000, 366 and 144 tensors, respectively. The first value corresponds to retaining all tensors. The second value is obtained by taking one tensor out of two. The third value is the number of tensors retained by the tensor picking operation. The last value is the minimum number of tensors that the tensor picking operation can produce (2 measures for each of the 72 sulci, i.e. 144 tensors). Indeed, at least two tensor values per sulcus are required to be able to perform a linear interpolation between them. Figure 11.9 summarizes the experience results. One notices that even with very few tensors (366 compared to 2000 initially), the model is able to recover a correct estimation of the variability in almost all areas. Local errors arise when the correlation of variability between neighboring sulci is too low (see regions with hot colors in Fig. 11.9, right column).

11.3.2 Leave One Sulcus Out

To evaluate further our model, we now perform the "leave one sulcus out" test. This test removes one sulcus and its variability tensors from the model and extrapolates the rest of the data to the full brain. Then, the prediction error made on this specific sulcus is compared to the interpolation and extrapolation errors. As the measures are independent, an error below $3\,\sigma_{ref}$ will be considered as not significant and shows that our extrapolation model recovers the missing variability information up to the intrinsic reconstruction uncertainty. By contrast, a RMSE larger than $3\sigma_{ref}$ means that we do not recover a comparable variability *in at least one direction.* We

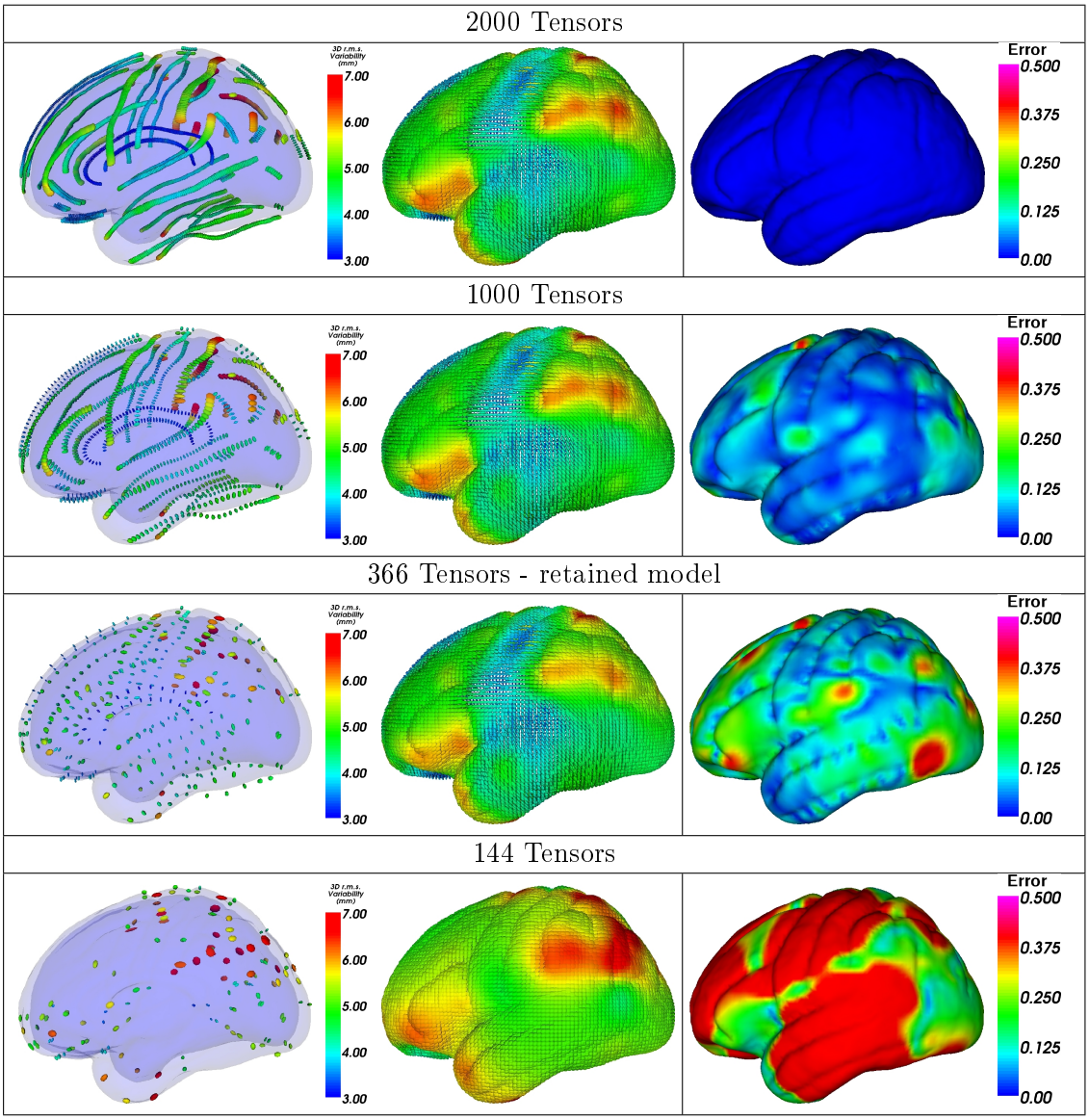


Figure 11.9: **Decreasing the number of picked tensors. Left:** Tensors retained for the model. **Middle:** Results of the extrapolation. **Right:** Error (LE distance) between the extrapolated and initial covariance matrices field.

know that an uncertainty in the tangent of the mean sulcus could be induced by the aperture problem. To remove this effect, we "project" the error vector onto the plane perpendicular to the tangent of the mean sulcus. Thus, the error component in this direction is zeroed out. We will call this error the "partial error".

This test is performed on 3 sulci: the Sylvian Fissure, the Superior Temporal Sulcus and the Inferior Temporal Sulcus. Fig. 11.10 displays the variability of the reconstructed sulci after extrapolation with and without their tensors, while Table 11.1 summarizes the global RMSE statistics. The prediction error with

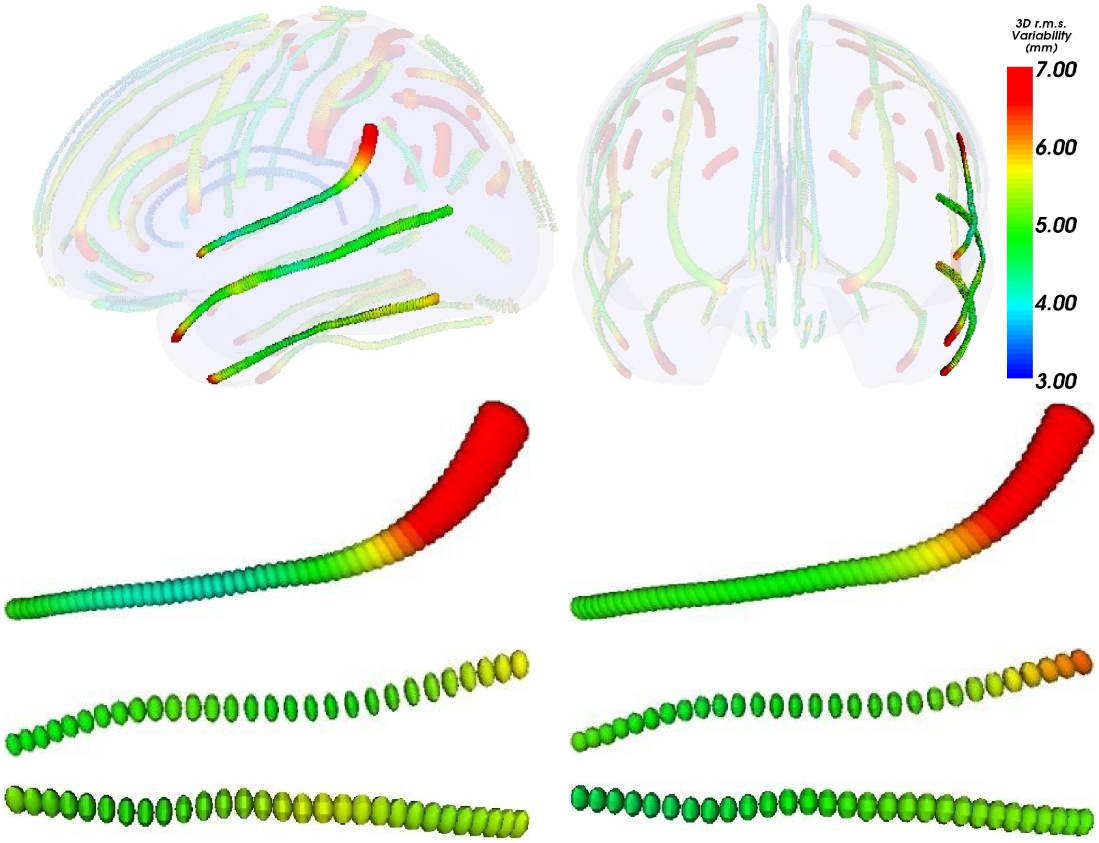


Figure 11.10: **Results of the "leave one sulcus out" test**. **Top:** Positions of the 3 tested sulci in the ICBM305 space. **Bottom left:** variability of each sulcus after extrapolation of the complete model (from top to bottom: the Sylvian Fissure, the Superior Temporal Sulcus, the Inferior Temporal Sulcus). The color bar is the same as in Fig. 11.8. **Bottom right:** extrapolated variability from the neighboring sulci only.

missing sulci is globally 2 to 3 times larger than that incurred by interpolating or extrapolating the full model, but the difference is not high enough to be significant. However, errors are locally significant. In some places, like for the Sylvian Fissure, the prediction errors occur primarily in the tangential direction to the mean sulcus. Indeed, three main sulci (the Pre-Central, Central and Post-Central Sulcus), are orthogonally adjacent to the Sylvian Fissure (Fig. 11.11), even though they do not actually merge with it, they influence the estimates of the tangential component of the variability as discussed in Sec. 11.3.1. Such behavior is confirmed by the "partial" error that is much lower than the standard one (Table 11.1). By contrast, the variability of some sulci like the Central Sulcus (Fig. 11.11) cannot be correctly recovered from neighboring sulci: the error is not only due to the aperture problem but spatial correlations between adjacent sulci may be lower in some brain regions, making variations more difficult to predict.

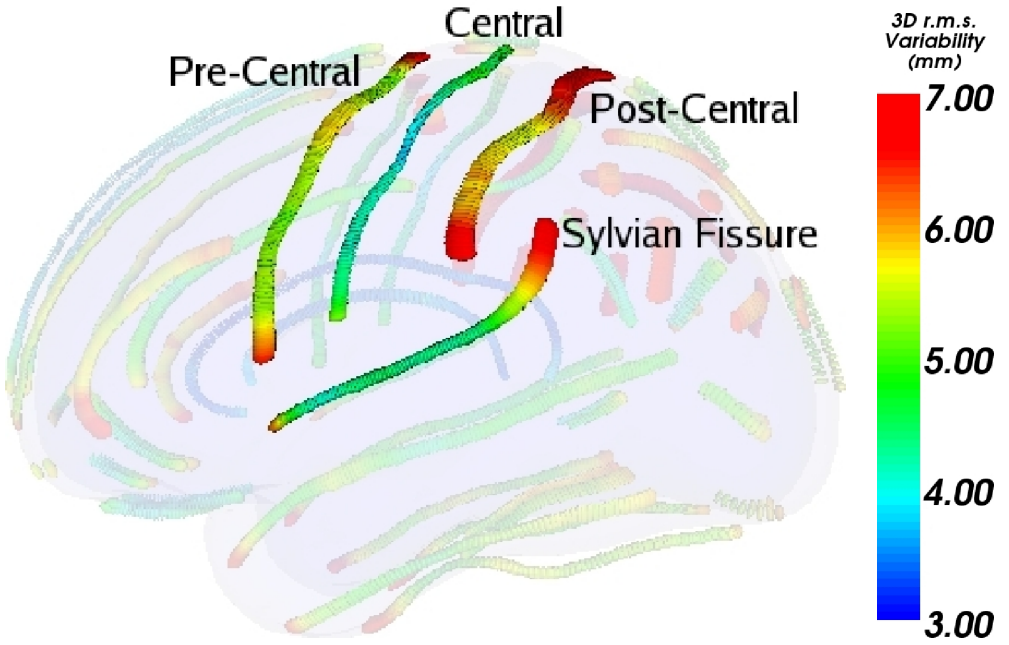


Figure 11.11: **Illustration of orthogonal sulci.** The Central, Pre-Central and Post-Central sulci have an orientation that is somewhat orthogonal to the Sylvian Fissure. This causes a difficulty in predicting the variability of the Sylvian Fissure from the other 3 somewhat orthogonal sulci. Another difficulty (less critical) is the prediction of the Central Sulcus from Post- and Pre-central sulci, because the Central Sulcus has a rather low variability compared to its immediate neighbors. Thus, the prediction will be difficult due to the lack of correlation in this region.

In conclusion, our model is able to recover intra-sulcus variability from the selected tensors, and to predict the variability in regions that are locally correlated. Nevertheless, our evaluation method is limited for two reasons. First, the aperture problem will cause an underestimation of the variability along the direction of the mean curve, and this feature is in fact observed in regions with orthogonal sulci. However, there are not so many such regions in our database, so this effect remains largely unnoticed. Second, the variability of some sulci is not correlated with that of their neighbors (this is the case for the Central Sulcus). These sulci carry some independent variability information, and should definitely be part of any brain variability model. One could consider including features as constraints for nonlinear registration of brain data when they contribute the greatest information on anatomical variation, i.e., information that cannot readily be predicted or interpolated from information on adjacent structures.

In the next Section, we investigate another problem: the asymmetry in cortical variability. In other words, we are interested in measuring whether a position on

Sulcus	Sylvian Fiss.	Sup. Temporal	Inf. Temporal.
Interpolation	0.17 - 0.15*	0.17 - 0.15*	0.17 - 0.14*
Extrapolation	0.19 - 0.16*	0.21 - 0.19*	0.20 - 0.17*
Extrapolation w/o sulcus	0.56 - 0.34*	0.42 - 0.32*	0.38 - 0.32*

Table 11.1: RMSE of reconstruction of 3 sulci with the interpolation, extrapolation and leave one-sulcus out extrapolation methods. * indicates the "partial error" (Sec. 11.3.2).

one hemisphere has a different variability than its symmetric positions in the other hemisphere.

11.4 Hemispheric Differences in Variability

The study of asymmetry of brain variability is of great interest for neuroscience [Toga 2003], and measures of structural and functional lateralization are of interest in mapping brain development, and disorders such as dyslexia, epilepsy, and schizophrenia. The two brain hemispheres develop according to slightly different genetic programs, and the right hemisphere is torqued forward relative to the left, with greatest volume asymmetries in the planum temporale and language cortex surrounding the Sylvian fissures (typically larger in the left hemisphere). If the types of variation in the two hemispheres could be differentiated, their genetic basis would be easier to investigate. It could also help understanding whether there is an asymmetry in the power to detect consistent task-related or group-specific activation patterns in functional brain imaging studies, due to structural variance asymmetries.

We measured the asymmetry of brain variability using our extrapolation model. The principle is to compute the (Riemannian) distance between the variability tensor at one point and the (symmetrized) tensor at the symmetric point in the brain. To define the symmetric point, a first geometric method is to simply use the 3D position that is symmetric with respect to the mid-sagittal plane in the stereotaxic space (mid-sagittal symmetry). In that case, we compute a dense asymmetry map from the extrapolated tensor values at each 3D point of a hemisphere (Fig. 11.12, left).

Another anatomical possibility is to measure brain asymmetry on sulcal lines, and extrapolate those measures to the whole brain (sulcal symmetry). First, all curves are mapped into a common hemisphere (e.g. the left hemisphere). Then, for each sulcus, a global mean is computed, as well as the left and right means (obtained by taking only the left (or respectively, the right) instances). Second, we compute the correspondences between this global mean and the left and right means. Thus, we define a common reference curve, the global mean, to compare left and right variability tensors. This prevents us from introducing a bias in the results, such as what might happen if we had chosen either the left or right mean

as the reference curve. Finally, difference vectors (Eq. 3.8) between left and right tensors are measured along the reference curve and extrapolated to the full brain using our framework. Note that as tangent vectors are elements of a vector space, we could use the vectorial and not the Riemannian (either AI or LE) version of our extrapolation. We end up with another dense asymmetry map, whose color is proportional to the distance between left-right tensors (Fig. 11.12 right).

A very interesting feature is that the regions with greatest asymmetries in variability include the one of main language areas, Broca's speech area (labeled A in Fig. 11.12) as well as the parietal cortex, which exhibits the greatest gross anatomical variation of any cortical area (labeled B in Fig. 11.12). As expected, these areas are more structurally variable in the left hemisphere which is typically dominant for language in normal right-handers. One surprise is that the tips of the Sylvian fissures do not show the greatest difference in variability between the hemispheres, as these are regions with highly variably branching patterns that have been studied extensively in the language and dyslexia literature. Also as expected, the primary sensorimotor areas (central and pre-central gyri) are relatively symmetric in their variance, as the absolute variability is lower, as is their degree of hemispheric specialization (i.e. they perform analogous functions in each hemisphere, but innervate opposite sides of the body).

11.5 Discussion

We presented an original methodology to model the profile of structural variability in cortical landmarks. We began with a dataset of 72 expertly delineated sulcal lines in 98 subjects, and proposed an original mean sulcal landmark computation strategy. The approach consists in alternately optimizing the position of the mean curve, computing the correspondences by dynamic programming and refining each subject's affine transformation based on the obtained matching. Then, variability tensors are measured along sulcal lines. Second, we applied our strategy for sparse tensor fields extrapolation detailed in Sec. 4.4.3 to obtain a dense representation of variability tensors on the cortex. The variability maps obtained agree with the notion that there is a natural hierarchy of variability in the brain, with the variations of structures within a lobe or set of gyri being predictable from their immediate neighbors with relatively high accuracy. The resulting framework is very flexible in the sense that other sources of information can be easily incorporated to improve the accuracy of the variability model.

The results are also interesting neuroscientifically. The amplitude and asymmetry of variability are greatest in the most phylogenetically recent areas in the cortex: the frontal lobe, the dorsolateral prefrontal cortex, and in the more dorsal areas of the parietal association cortices. The language areas, in particular, have fundamentally different developmental programs in each brain hemisphere, leading to volumetric and functional asymmetries (e.g. left hemisphere language

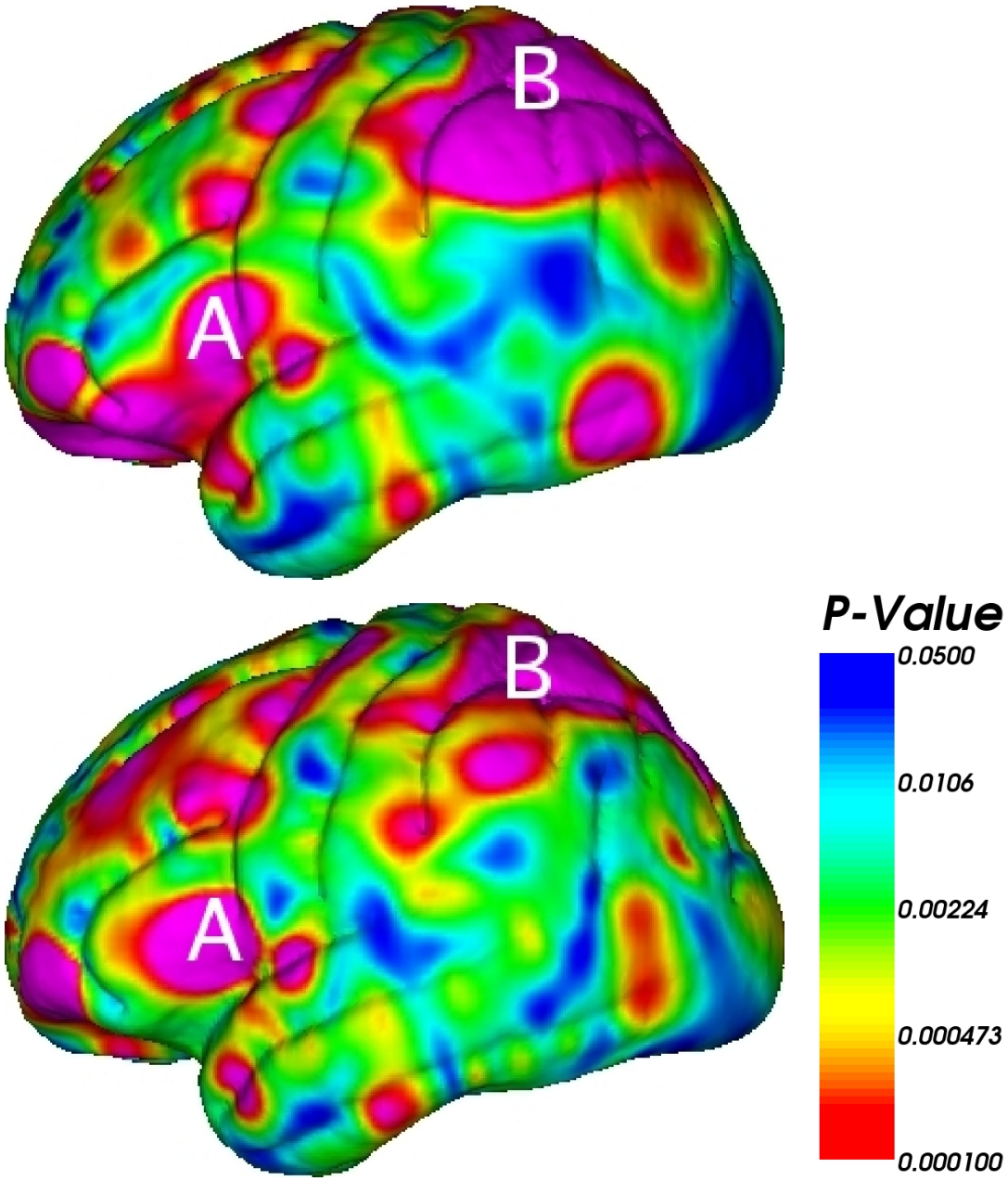


Figure 11.12: **Maps of brain variability asymmetry obtained using the methodology described in Sec. 11.4. Left:** The mid-sagittal symmetry. **Right:** The sulcal symmetry.

dominance). This variance asymmetry is seen in Broca's area, which is specialized in the left hemisphere for producing speech, but is less commonly associated with structural asymmetries. Lower variance was seen in cortical regions subserving primary brain functions (e.g., touch, motor function, hearing) and these areas are the earliest to mature in utero. It would be interesting to hypothesize that the

areas of the brain that emerged most recently in human evolution are the same ones that have greatest variation between hemispheres, reflecting the drive towards hemispheric specialization of function in higher primates and man. The modeling of variance is practically valuable for understanding and discovering genetic and disease related factors that affect brain structure, which are currently hard to identify given the extremely complex patterns of variation in normal anatomy. For example, many neurodevelopmental disorders are associated with subtle variations in the patterning of the cortex, and new computational anatomy techniques are making these features easier to identify (e.g., increased cortical complexity in Williams syndrome [Thompson 2005]).

This concludes our first contribution on the statistical analysis of the brain. The second chapter is dedicated to the analysis of the potential correlations between anatomical positions: we are now considering the relationships between any pair of positions via a statistical tool called the *total covariance matrix*.

Chapter 12

Accessing Close- and Long-Range Correlations via the Total Covariance Matrix

Contents

Understanding structural correlations between brain structures is a challenging problem in neuroscience. Most computational anatomic studies of development and disease study deficits or changes by modeling individual brain structures independently, or create voxel-based maps of group anatomical differences. This reveals factors (e.g., age or disease) that influence each brain structure individually, but may miss important supra-regional correlations, such as brain regions that develop or fail to develop together, or correlations between structures in opposite brain hemispheres. Mechelli et al. [Mechelli 2005] studied gray matter density, examining 12 brain subregions in 172 normal subjects. They discovered (1) spatial correlations between corresponding regions in opposite brain hemispheres, except in the visual cortex, and (2) some negative correlations between functionally distinct regions in the same hemisphere. Neuroscientists are interested in identifying the reasons of such long-range brain correlations, and what causes them at a genetic and environmental level. Normally coordinated systems may be disrupted in neuropsychiatric disorders such as schizophrenia, autism, and Williams syndrome, and systems with known fiber or functional connectivity are thought to exert mutual influences on each other's growth, via activation-dependent plasticity. Structures in different lobes also influence each other's degeneration, for instance when loss of neuronal input from one brain region induces degeneration in another. Despite many hypotheses, few tools allow such long-range correlations to be measured, and thus studied empirically.

An efficient, parsimonious model of the complex patterns of brain correlations should help to identify factors that influence them. Inter-hemispheric correlations (i.e., correlations between point in anatomically homologous structures in both hemispheres) are of interest as they shed light on how their functions become specialized or depend on each other. Improved modeling of correlations of brain structures could make it easier to isolate specific effects of genetic polymorphisms on these normal correlations. Furthermore, information on statistical correlations could reduce the difficulty of automated segmentation and labeling of brain structures. Accessing anatomical correlations also opens up a broad range of studies and group comparisons (e.g., disease versus normal) and a new path to generate hypotheses regarding patterns of brain growth.

First-order models (mean anatomical templates), and second-order models (variability models - see Sec. 11) of the brain have been previously built to capture the 3D variations of each anatomical point independently around a mean anatomy, after registration of multi-subject anatomical images to a common reference space. These variations are often represented by covariance matrices, or variability tensors, as variations may not be the same in all directions. Here, we go one step further and model the joint variability of all pairs of anatomical points, to see how the displacement of any point in a specific subject w.r.t. a reference anatomy covaries with the displacement of neighboring or distant points.

In Section 12.1, we introduce the main tool of our analysis: the total covariance matrix (TCM) between two vector variates, and we recall how to extract from it some matrix and scalar measures to test if these two variables are correlated. In Section 12.2, we experiment this framework on TCMs defined from anatomical landmarks (sulcal curves). We start by studying the TCMs of 6 sulcal positions to the rest of the brain, which eventually lead us to analyze the TCMs of all sulcal positions of one hemisphere with their homologous positions in the opposite hemisphere.

12.1 Methodology

12.1.1 The Total Covariance Matrix

Let $X = \{X_i\}_{i=1..N}$ and $Y = \{Y_i\}_{i=1..N}$ be the sets of N measures of two random vectors whose dimensionality is d. Computing the correlation between X and Y requires to know not only the variability of each vector (i.e., its covariance matrix), but also their cross-covariance. We therefore define the total covariance matrix of X and Y that contains this information as $\Lambda(X,Y)$:

$$\Lambda(X,Y) = \frac{1}{N-1}\sum_{i=1}^{N} \begin{pmatrix} X_i - \bar{X} \\ Y_i - \bar{Y} \end{pmatrix} \begin{pmatrix} X_i - \bar{X} \\ Y_i - \bar{Y} \end{pmatrix}^{\top}, \tag{12.1}$$

where $\bar{X} = (\sum_1^N X_i)/N$ and $\bar{Y} = (\sum_1^N Y_i)/N$. After term by term multiplication, one obtains:

$$\Lambda(X,Y) = \frac{1}{N-1}\sum_{i=1}^{N}\begin{bmatrix} (X_i-\bar{X})(X_i-\bar{X})^\top & (X_i-\bar{X})(Y_i-\bar{Y})^\top \\ (Y_i-\bar{Y})(X_i-\bar{X})^\top & (Y_i-\bar{Y})(Y_i-\bar{Y})^\top \end{bmatrix}. \quad (12.2)$$

The diagonal terms are in fact the covariance matrices of X and Y. We denote by Σ_{XX} (resp. Σ_{YY}) the covariance matrix of X (resp. Y): $\Sigma_{XX} = \mathrm{E}[(X - \bar{X})(X - \bar{X})^\top]$. The off-diagonal elements are the cross-covariances of X and Y: $\Sigma_{XY} = \mathrm{E}[(X - \bar{X})(Y - \bar{Y})^\top]$. We have the property that $\Sigma_{YX} = \Sigma_{XY}^\top$. By further analyzing Eq. 12.1, one can write $\Lambda(X,Y)$ in a simpler way:

$$\Lambda(X,Y) = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}. \quad (12.3)$$

Λ is a $2d \times 2d$ matrix. It has the same properties as a classical covariance matrix: it is symmetric and positive definite. Then, we may also call Λ a tensor. In the 3D case, Λ is a 6×6 tensor.

12.1.2 Analysis of Total Covariance Matrices

In its current form, it is difficult to appreciate the meaning of the TCM and it cannot be easily represented (it is an ellipsoid in 6D). However, several matrix, vector and scalar measures may be derived from it. Here, we will focus on quantifying the correlation of X and Y through the Canonical Correlation Analysis.

Canonical Correlation Analysis (CCA): CCA [Rao 2006] refers to the method of finding vector bases that maximize the correlation between two vector variates, and is the generalization of the correlation coefficient to multivariate data. In the scalar case, we define the correlation coefficient between x and y as: $\rho = \sigma_{xy}/\sqrt{\sigma_{xx}.\sigma_{yy}}$, where σ_{xx} (resp. σ_{yy}) is the variance of x (resp. y), and σ_{xy} is the cross-variance of x and y. Similarly, the correlation matrix Γ in the multivariate case is defined as:

$$\Gamma(X,Y) = \Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_{YY}^{-1/2}. \quad (12.4)$$

We have the property that $\Gamma(Y,X) = \Gamma(X,Y)^\top$. Taking the mean trace of the Γ gives us an average correlation coefficient $\bar{\rho}$.

$$\bar{\rho} = \frac{1}{d}\mathrm{trace}\left(\mathrm{CCA}(X,Y)\right).$$

The range of $\bar{\rho}$ lies between 0 (absence of correlation - X and Y are independent) and 1 (correlation: X and Y vary in a linear relationship). In multivariate statistics, however, a low average correlation coefficient may hide a potentially high correlation. This is the case, for instance, when only one component of X is correlated with one component of Y. Taking the average correlation coefficient may hide this relationship. To distinguish between correlations along potentially different axes, one needs

to run a canonical correlation analysis, which is nothing else than decomposing Γ in singular values:

$$\Gamma = U.S.V^\top = \begin{bmatrix} | & & | \\ \overrightarrow{u}_1 & \dots & \overrightarrow{u}_d \\ | & & | \end{bmatrix} \begin{bmatrix} \rho_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \rho_d \end{bmatrix} \begin{bmatrix} - & \overrightarrow{v}_1^\top & - \\ & \vdots & \\ - & \overrightarrow{v}_d^\top & - \end{bmatrix}.$$

where U and V are orthogonal matrices of correlation vectors, and S is a diagonal matrix of correlation coefficients ρ_i (ρ_i is the correlation coefficient between vectors $\overrightarrow{u}_i$ and $\overrightarrow{v}_i$).

Significance Testing: To test the statistical significance of correlations, [Fujikoshi 1979] proposed testing the dimensionality of the correlation matrix. If its rank is zero, then there is no correlation ($\rho_i = 0, \forall i$). If we reject this hypothesis, then the rank is at least one, which means that at least two directions in space are correlated. We use the Bartlett-Lawley test [Fujikoshi 1979] with the null hypothesis: $H_0 : \text{rank}(\Gamma) = 0$: $L(\Gamma) = -\left(N - d + \frac{1}{2}\right) \sum_{j=1}^{d} \log\left(1 - \rho_j^2\right)$. L's distribution is chi-squared under the Gaussian assumption on X and Y with d^2 degrees of freedom. We can consequently derive a p-value for testing the significance of correlations.

12.2 Sulcal Correlations

We used the same dataset of sulcal landmark curves as in section 11. It is composed of 72 sulcal curves manually delineated in 98 subjects by expert anatomists, according to a precise protocol with established reliability within and across raters. The dataset consists of 47 men and 51 women (age: 51.8 +/- 6.2 years), all normal controls. MR images used for delineations were first linearly aligned to the ICBM stereotactic space [Collins 1995].

We used the methodology outlined in Sec. 11.1 to determine the mean curve for each sulcal line by modeling samples as deformations of a single average curve. Mean curve computation involves filtering each sample by B-spline parameterization, minimization of total variance, and sulcal matching by dynamic programming. For each of the 72 sulci, we obtain a mean curve, $z(s)$, and a set of instances $y^k(s)$ along with correspondence functions $\phi^k(s)$, where k is the subject's number, so that $z(s)$ corresponds to $y^k \circ \phi^k(s)$ (each curve is parameterized by the normalized arc-length $s \in [0, 1]$, provided by the B-spline representation).

In the following, we investigate the potential correlations between locations on different sulci. First, we study the correlation between particular sulcal lines and other cortical points not belonging to the same structure: we call this study *sulcal correlation*. Second, we assess *inter-hemispheric correlations* between corresponding anatomical points in the two hemispheres.

12.2.1 Sulcal Correlation for 6 Specific Positions

Methodology: Obviously, the study of correlations is a combinatorial challenge. We sampled the 72 mean sulci with approximately 500 points (average of 7 points per sulcus), which gives a total of about 250000 pairs of points to process. To limit the number of pairs investigated, we focused on two major sulcal lines: the Central Sulcus (CS) and the Superior Temporal Sulcus (STS). The CS is a relatively stable landmark, dividing the frontal and parietal lobes, and separates primary sensory and motor cortices. The STS is an important landmark in studies of facial imagery processing, as it separates the fusiform gyrus from the Sylvian fissure. These sulci lie in different lobes, develop at different times during gyrogenesis (CS developing earlier) and are distant in terms of fiber and functional connectivity, so they are good candidates for assessing inter-structure correlation, as little correlation is expected *a priori*. For each of these lines, three reference positions are picked: the beginning, middle, and end point. First, for each of the three reference positions, we extract the set of corresponding sulcal positions in each of the 98 subjects. Second, we compute the TCM of Eq. 12.3 with each of the remaining 999 average sulcal positions. We end up with a sparse field of TCMs. However, we would be more comfortable with a dense field of TCMs, as we could map those onto an average cortex to facilitate the visual interpretation of the results. We use Log-Euclidean (LE) metrics and the methodology exposed in Sec. 4.4.3 (combination of a radial basis function interpolation with an harmonic partial differential equation) for extrapolating TCMs on a mean cortical surface. This type of interpolation was shown to preserve all the features of a covariance matrix, and has desirable properties like absence of swelling effect, and a smooth interpolation of the eigenvectors. Moreover, leave-one-out tests showed that this type of interpolation is able to predict missing data in regions locally correlated. This interpolation is consequently well adapted for TCMs. The correlation matrix and the p-value derived from the CCA can be computed at any point of this mean cortex. Notice that even if we only focused on the p-value defined in Sec. 12.1.2, other measures are potentially interesting, such as the principal vectors of correlation which are currently under investigation. This is why we need to extrapolate the full TCMs and not just the p-value.
The main problem for the curve matching procedure is the *aperture problem*: correspondences in the direction tangent to the curve are almost impossible to retrieve without additional expert knowledge. To keep our results independent from this, we need to cancel the contribution in this tangential direction. The method proposed in this paper is the following. We define at each position of the mean sulci the Frenet frame, which gives us the plane orthogonal to the curve. Then, we project the sulcal positions onto this plane, which zeroes out the tangential component. Note that we lose one degree of freedom in the dimensionality of the data: vectors no longer have three degrees of freedom but two. This must be accounted for in the statistical tests of Sec. 12.1.2.
Results: p-value derived from the CCA are shown in Fig. 12.2 (the significance

level was set to 0.0001 to correct for multiple comparisons [1]). A large area around the reference points shows high p-values: as expected, points that are anatomically close to the reference do have a correlated distribution among individuals. More interestingly, regions with high p-values most often include the structures' opposite hemisphere counterparts, but not always: the first (upper) and middle CS positions are highly correlated in the other hemisphere (Fig. 12.2 top panel, 1st and 2nd row), while the most inferior position is not, most likely because its variability across subjects is extremely low (Fig. 12.2 top panel, 3rd row). In right-handed subjects, we know that some measures of motor skill (maximum finger tapping rate) correlate with gray matter volume positively in the left CS, but negatively in the right CS [Herve 2005]. Logically, such functional specializations may promote correlations between the two hemispheres in these regions of the CS, where tissue volumes directly correlate with skill levels. The posterior part of the STS shows lowest correlation with its opposite hemisphere counterpart. Unlike the bottom of the CS, the posterior tip of the STS is highly variable and asymmetric in structure and function - it is specialized for understanding the semantics of language in the left hemisphere, but for understanding prosodic aspects of language in the right hemisphere. This may suggest partially independent developmental programs for these functionally specialized structures. The long-range correlation between the back of the STS and the left and right intra-parietal sulci is of interest, as the planum temporale and planum parietale are the two distinct areas most widely studied in neuroscience for their very high hemispheric asymmetry.

Nevertheless, it is intriguing that 5 of the 6 sulcal positions studied reflect a correlation with their symmetric counterpart in the opposite hemisphere. In the following, we test if this observation can be generalized to all sulcal positions.

12.2.2 Special Case of Hemispheric Correlations

In this study, we specifically target the correlation between all points of one hemisphere and their homologous region in the other hemisphere. To do so, we use a strategy similar to Sec. 11. We first map all sulci of the right hemisphere onto the left. Then, we define a global mean, i.e. an average sulcal curve computed from the 98 left and right samples. Global means provide a common reference curve to compare left and right positions. Correspondences between global means and left and right average curves are computed using the same framework as for the samples. For any given position on the global mean, we obtain corresponding points on left and right average curves, giving in turn correspondences between left and right sulcal positions in the 98 subjects (the choice of the left hemisphere is arbitrary, and we obtained the same results when using the right hemisphere instead). Finally, we compute the TCM of Eq. 12.3 between left and right positions. As for the sulcal correlation study, we extrapolate this sparse field of TCMs to an average left hemisphere surface, and cancel the tangential component which is uncertain. Then, we extract p-values with the CCA and map those on the surface (Fig. 12.5).

[1] This corresponds to a standard p-value of 0.05 as there are around 500 free variables.

As observed previously on a few specific sulcal positions, most points are correlated with their symmetric counterparts. Regions with lowest correlations include Broca's and Wernicke's areas, which were already shown to exhibit the greatest asymmetries in variability (Sec. 11.4). These cortical regions are specialized for language production and comprehension respectively, but most right-handers show a greater reliance on the left hemisphere for language processing, and the volumes of these regions are highly asymmetric between hemispheres.

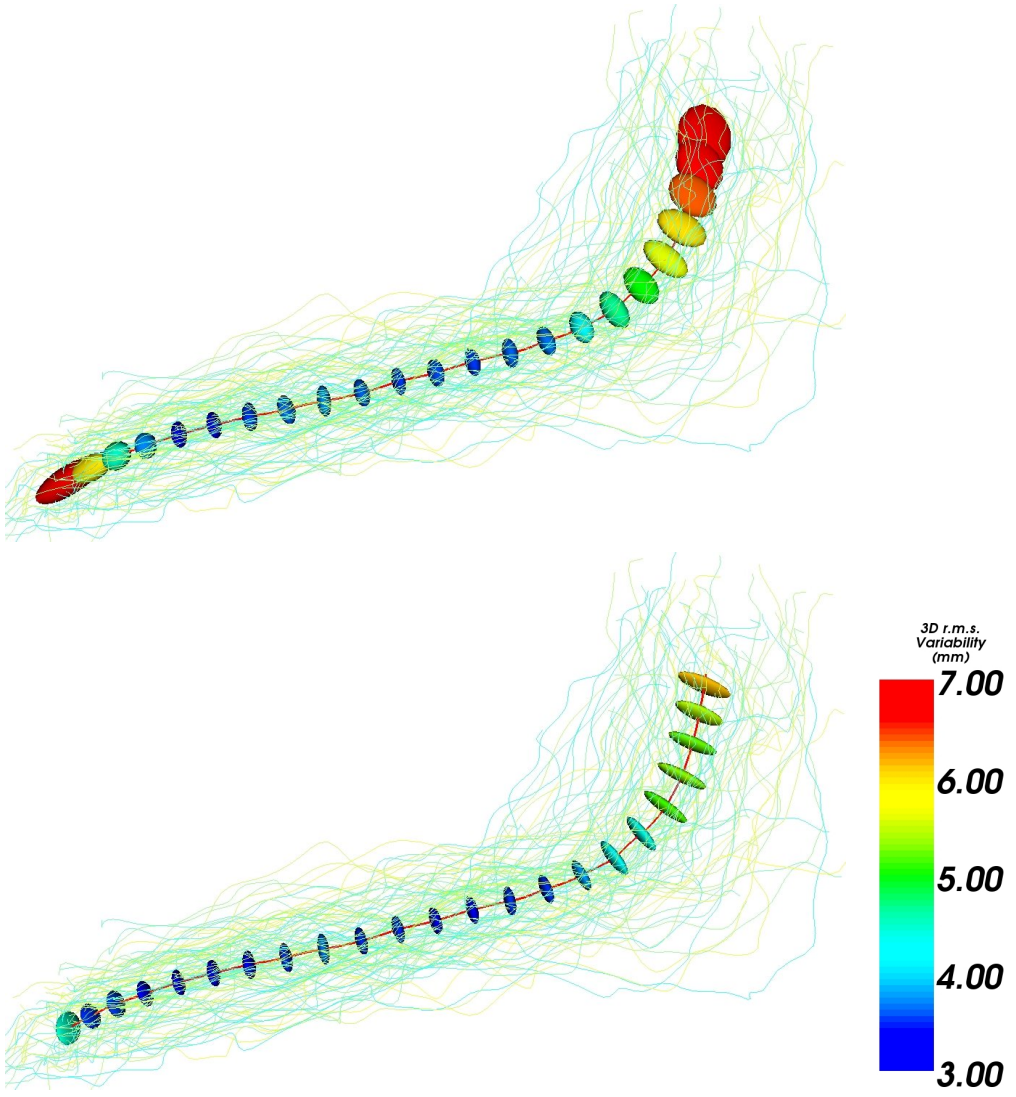


Figure 12.1: **Canceling the tangential component. Top:** Covariance matrices extracted along the mean Sylvian Fissure as in section 11. **Bottom:** The same tensors after canceling the tangential component. A small tangential value is kept for visualization purposes (tensors should be flat).

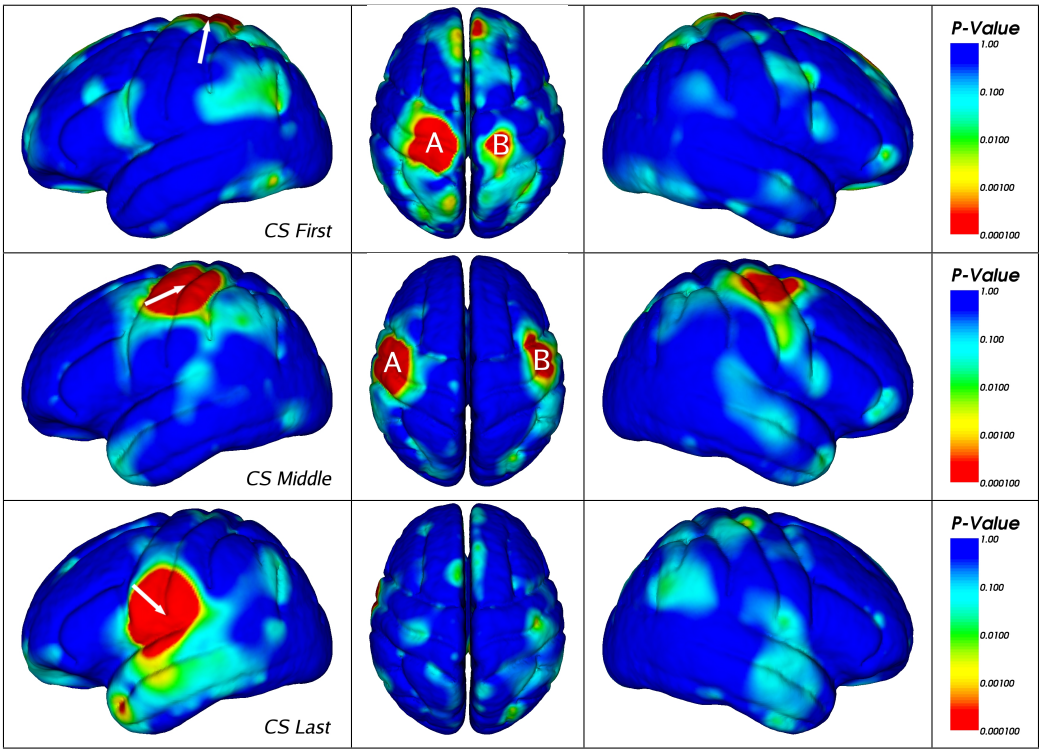


Figure 12.2: **Correlation Maps between Specific Reference Points and Other Brain Regions: the Central Sulcus.** A white arrow in each row indicates a reference landmark; correlations with the reference landmark are plotted. Correlations for 3 reference landmarks on the CS are shown: the first (top row), the middle (second row), and the last, i.e. most inferior, position (third row) on the sulcal trace. Corresponding regions in the opposite hemisphere are highly correlated for the top and middle points (marked A and B). The lower end of the sulcus, however, exhibits low correlation with its symmetric contralateral counterpart. Towards the bottom of the central sulcus, intersubject variation is as low as 4mm (3D r.m.s. after linear registration) which is not much higher than the inter-rater reliability of the tracing protocol, so low correlations with other landmarks are expected here.

12.3 Conclusions

In this section, we represent the cross-covariance between one point and any other point of the brain by a total covariance matrix describing not only the variability of the two points, but also their cross-covariance. Canonical correlation analysis allows us to test for the significance of these correlations. As TCMs have the same properties as classical covariance matrices (symmetric and positive-definite), we used the Log-Euclidean extrapolation to obtain a dense representation of initially sparsely-defined measures. This extrapolation was indeed previously shown to have good properties, contrary to the Euclidean extrapolation of the matrix coefficients.

We apply this method to study sulcal and hemispheric correlations. We showed

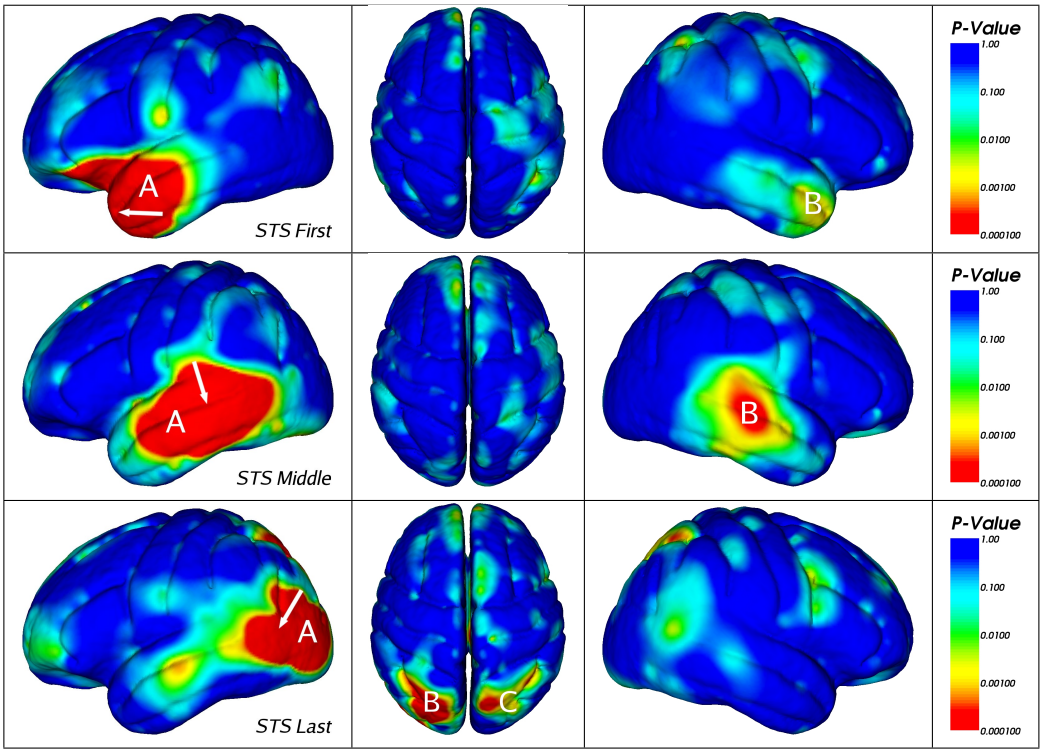


Figure 12.3: **Correlation Maps between Specific Reference Points and Other Brain Regions: The Superior Temporal Sulcus.** The same 3 positions as for the CS are analyzed. The first (top row) and middle (second row) positions are symmetrically correlated (marks A and B). The last position (third row) correlates less with its opposite hemisphere counterpart, than with the intra-parietal sulci (marked B and C), probably because it lies in a region of extremely low anatomical variability.

that the central sulcus was highly correlated with its symmetric counterpart, except in its inferior part which is not highly variable. For the central sulcus, where motor skill is correlated with volume and is also lateralized, a strong hand preference for motor skills is likely to promote negative correlations between hemispheres for volumes in opposite regions. The Superior Temporal Sulcus shows similar intriguing correlations, and its low correlation with its opposite counterpart may reflect their different developmental programs and functions.

Corresponding brain regions in each hemisphere are highly correlated, except for regions including Wernicke's and Broca's areas, which are known to be functionally specialized in one hemisphere. Any longer-range correlation - such as that found between the intra-parietal sulci and superior temporal sulci - is in itself an interesting neuroscience finding. The planum parietale and temporale are distinct highly asymmetric systems in each of these regions, and the long-range correlations may reflect common factors driving programmed asymmetries for both regions.

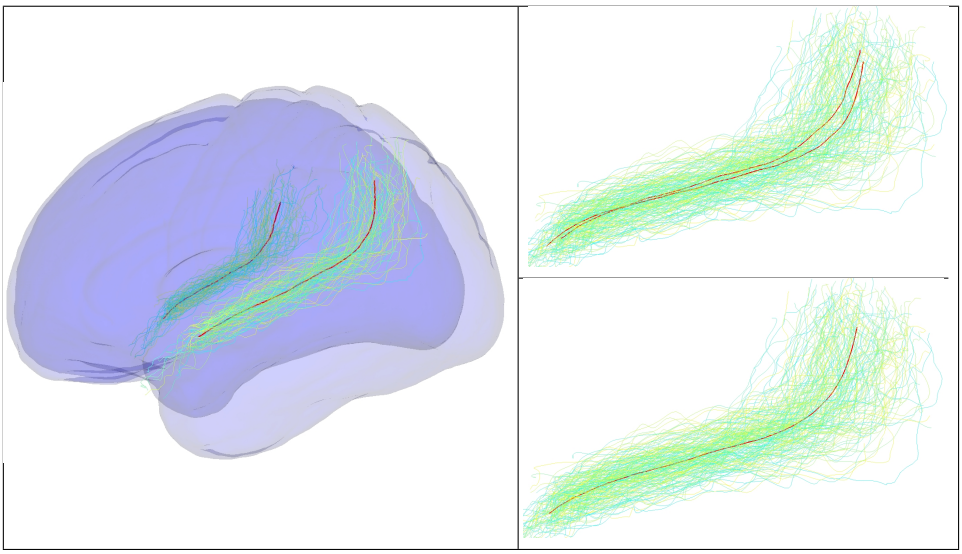


Figure 12.4: **Finding hemispheric correspondences. Left:** Positions of left and right Sylvian fissures (SF), in an oblique view, with a transparent cortical surface overlaid (in blue colors). **Top Right:** All right hemisphere sulci are mapped to the **Bottom Right:** Same set of sulci, whose global mean was computed (in red). This is now used as a reference to compare left and right sulcal positions.

This concludes our second contribution on the statistical analysis of the brain. In the next (and final) chapter of this part, we summarize the contributions and discuss about the future work and the possible applications of such brain models.

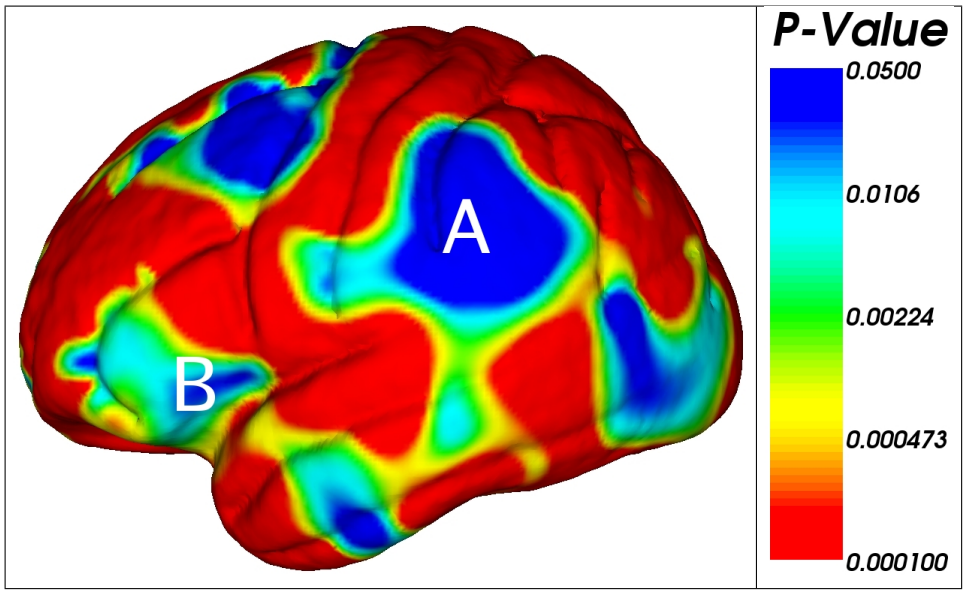


Figure 12.5: **Hemispheric correlations.** Hot colors (red) correspond to p-value$\leq$0.0001. Most of the cortex shows anatomical variations that are correlated with their counterparts in the opposite hemisphere. Uncorrelated regions include Wernicke's (marked A) and Broca's areas (marked B), which, in most subjects, are known to be more heavily specialized for language processing in the left hemisphere.

Chapter 13

Discussion

In this part, we have explored the statistical analysis of the human cortex. We have proposed to: 1) work on sulcal lines as they are precise anatomical landmarks identified by experts and of great interest in neuroscience, 2) model independently the variability of each sulcal positions by their covariance matrices (so called variability tensors) and extrapolate this sparse information to obtain a dense map of variability, and 3) evaluate close and long range anatomical correlation via the total covariance matrix.

As expected, the amplitude and asymmetry of variability are greatest in the most phylogenetically recent areas in the cortex (the frontal lobe, the dorsolateral prefrontal cortex, and in the more dorsal areas of the parietal association cortices). Greatest asymmetry of variability were also found in the language areas, and in particular the Broca's area, which have different developmental programs in each brain hemisphere. Lower variability were found in cortical regions supporting primary brain functions. As these areas are the ones that mature the earliest in utero, we could hypothesize that the areas of the brain that emerged the latest in human evolution are the ones that have greatest variations between hemispheres, and conversely, regions that develop first are very likely to be stable (i.e., less variable) among individuals.

The second analysis showed that anatomical correlations are highly present among individuals, and not only locally, which tends to prove that brain regions do not develop independently, but are connected. The most striking result is the inter-hemispheric correlation which was shown to be dominant. Other longer-range correlation, as that found between the intra-parietal sulci and superior temporal sulci, is in itself a neuroscience finding: the long-range correlations may reflect common factors driving programmed asymmetries for both regions.

The major drawback of working only on sulcal lines is that we only get variability information at specific locations on the cortical surface. A more complete model of brain variations should contain information on deeper structures like the ventricles, grey nuclei or even the white matter. However, this framework is very flexible in a sense that other source of information could (and should) be included to refine the variability model. One could think of including variability tensors calculated from matchings of surfaces of interest like the hypocampi, amygdala, putamen or caudate. Moreover, for a more in-depth analysis of the white matter, statistical analysis of white matter fiber tracts as those obtained in diffusion tensor

MRI (see Chap. II) is of great importance. If we could model some of the major fiber tracts, i.e., define what is the average fiber bundle of a given population, one could also extract meaningful variability tensors from them and feed our model.

The study of anatomical correlations as we have done it, while being very promising, is exploratory and much work still needs to be done. First, one would need to simplify the representation of these correlations. We could store, for each point of the brain, a minimal set of correlated positions, such as the local neighborhood and the set of most correlated distant points. This information could be used as a prior to guide inter-subject non-linear registration algorithms: if one position is moving towards another, then its set of correlated positions should move as well, in a direction given by their correlation vector. Furthermore, a detailed study of all possible correlations between cortical landmarks could help understanding the effects of genes on brain maturation. Validations of long-range correlations could be made using other information, such as functional MRI: Are jointly activated regions, or causal models (achieved through structural equation modeling or dynamic causal modeling) for a given task related to anatomical correlations as well as connectivity (fiber bundle)? In other words, does a physical link, via neural fiber bundles that can be traced with diffusion tensor MRI, exists between functionally correlated areas? All this information, if it converges to the same outcome, could contribute to understand the functional organization of the brain. Finally, these results should be compared with those of other methods, such as surface-based versus volumetric registration algorithms; this comparison is is currently underway for generating second-order models of brain variability [Durrlemann 2007].

Chapter 14

Conclusions and Future Work

Contents

14.1 Summary of Contributions

The central topic of this thesis is the processing of symmetric, positive definite matrices, so-called *tensors*. We were interesting in three classes of problems. First, we investigated the theoretical aspect of tensor processing, and proposed two frameworks (the affine-invariant and Log-Euclidean metrics) which replace the standard Euclidean calculus that suffers from several major drawbacks. Second, we focused on several issues that come with diffusion tensor MRI. In particular, the theoretical framework proposed here is applied to address a clinical problem: how can we optimally exploit diffusion tensor images typical of clinical acquisitions? These data are indeed of a lower quality compared to data coming from a research environment, and experts still want the same quality of results. Thanks to the use of Log-Euclidean metrics, we were able to rigorously model the problem mathematically. We were also interested in another aspect of this problem, which is the transfer of methods to the clinical world. For this purpose, we developed a software called MedINRIA

targeting explicitly the medical experts. Finally, we explored an emerging field: computational anatomy of the brain. Computational anatomy aims at modeling organs in terms of average shape and variations. In the case of the brain, we proposed to model the cortical variations via *variability tensors* computed at specific locations and to extrapolate this sparse information to obtain a dense field. Once again, the framework for tensor processing is intensively used. In a second step, we modeled the anatomical correlations between two cortical positions thanks to the total covariance matrix, another "type" of tensor that fits within our theoretical frameworks. This more exploratory research lead to interesting new neuroscience findings, and opens the door to a large panels of applications. These classes of problems share one common goal of great importance: the modeling and comprehension of the human brain. In the following, we give a more detailed description of our contributions.

14.1.1 Two alternatives to Euclidean Calculus on Tensors

We developed two theoretical frameworks that overcome the limitations of Euclidean calculus on tensors. In Euclidean calculus, non-convex operations are not stable and eventually lead out of the tensor space, i.e., matrices with null or negative eigenvalues appear. As the tensor space is a manifold, we decided to use the tools from differential geometry to overcome these limitations. First, we endowed the tensor space with an affine-invariant Riemannian metric. Thanks to this metric, we could derive two diffeomorphisms, the exponential and logarithmic maps, that are the keys to the generalization of any vector-processing algorithm of tensors. The logarithmic map has the role of "flattening" the tensor space (i.e., it turns it into a vector space where calculation can easily be done), while the exponential map "folds back" this flat space to form the tensor space. As a result, the tensor space is replaced by a smooth manifold where matrices with null and negative eigenvalues are at an infinite distance from any tensor. Although this Riemannian metric has very nice theoretical and practical properties, it can be a burden in practice since it makes an intensive use of the matrix exponential, logarithm, square root and inverse. We finally came up with a second family of metrics, called Log-Euclidean, which have almost the same properties as the affine-invariant metric but with a computational cost close to Euclidean calculus. Their principle is straightforward: one simply needs to take the logarithm of tensors once, process them, and take the matrix exponential to obtain the final result. We show that it is almost impossible to detect differences between these two frameworks with many examples of image processing algorithms (filtering, interpolation, etc.), making the Log-Euclidean metrics a better candidate to replace the Euclidean calculus.

14.1.2 Optimal Processing of Clinical DT-MRI

In this second part, we built a processing pipeline adapted to diffusion tensor data acquired with clinical constraints. We first revisited the estimation step. Diffu-

sion tensor estimation is generally done by linearizing the Stejskal&Tanner diffusion equation, which implicitly assumes additive noise on the logarithm of measured signals (we call this noise log-Gaussian). Although this is reasonable for high SNRs, in our case we chose to rely on the true nature of MR noise which is Rician (independent centered Gaussian noise on the real and imaginary part of the complex MR signal). We showed that estimating tensors with the wrong noise prior leads to a *shrinking effect*: tensor volume tends to be underestimated due to the positive bias induced by the Rician noise. Thanks to Log-Euclidean metrics, we were able to define and minimize an objective function with a Rician prior for the data fidelity term, and an anisotropic regularization term for enforcing the spatial regularity of the tensor field. Finally, we showed quantitative and qualitative improvements of the tensor estimation on both synthetic and real datasets, and a major positive impact on the overall quality of fiber reconstruction of a brain and spinal cord DTI dataset. For a valuable use of these methods in clinics, we started the MedINRIA project. MedINRIA is a general-purpose medical imaging software, which contains some of the advances in diffusion tensor MRI proposed in this thesis. We implemented a full pipeline for DT-MRI processing with Log-Euclidean metrics, from estimation to fiber reconstruction: this includes an anisotropic tensor smoothing algorithm, a tractography method using LE tri-linear interpolation, and intuitive 3D manipulation and exploration of the results. We are willing to make this software as user-friendly and as intuitive as possible. Indeed, medical experts cannot in general spend hours training on new software. Moreover, it should also fit the clinical environment, which in particular means being compatible with the DICOM file format. To demonstrate the practical use of MedINRIA, we achieved in collaboration with Denis Ducreux, MD., at the Bicêtre Hospital in Paris, a study on the current and future indications of DT-MRI and tractography in spinal cord lesions. Very promising results showed that tractography is able to detect certain pathologies earlier than T_2 images, which is currently done. It also proved that DT-MRI can effectively be used in clinics to detect and in the follow-up of pathologies.

14.1.3 New Tools for Statistical Shape Analysis of the Brain

This last application takes place in the context of computational anatomy of the brain. In this work, we relied on anatomical well-defined landmarks: the sulcal lines. We used a dataset of 72 sulcal lines manually delineated in 98 subjects by experts neuroanatomists following a precise protocol. We motivate the choice of these lines by the fact that finding correspondences among individuals of low dimensional structures is less hazardous than for higher dimensional objects like surfaces or volumes. However, the *aperture problem* is still present: without additional knowledge, we are not able to recover correspondences in the direction tangent to the curves. Knowing this, we developed an original sulcal matching procedure which minimizes the influence of this aperture problem. The average representation for each of the 72 sulci is done by alternately optimizing the position of the current estimate of the mean curve so that it minimizes the total variance, and computing optimal corre-

spondences by dynamic programming. We ended up with 72 mean sulci, and a set of point-to-point correspondences between each sulcal instance and the mean line. From this, we extracted along the mean shapes a set of *variability tensors* which are nothing other than the covariance matrix of each of the mean positions. Then, we showed how to select only a subset of meaningful tensors and extrapolate this sparse field to obtain a dense representation. Statistical tests were done to evaluate the predictive power of such model. We first evaluated the intra-sulcus recovery rate: are we able to predict the variability of a full sulcal line just with a few tensors? Second, we performed leave-one-out tests to determine if the variability of one sulcal line could be predicted with surrounding neighbors. Results showed that missing information was nicely retrieved in regions where the correlations between sulci were high, while failing in uncorrelated regions. This lead us to specifically study the anatomical correlations between any pairs of sulcal positions. Thanks to a tool called the *total covariance matrix*, a 6×6 tensor, we were able to build maps of correlations between a point of reference and the rest of the brain. This is obviously a combinatorial study (one 3D map per reference point), so we restricted our analysis to 6 positions taken on two major sulcal lines. Results showed that most of the sulcal positions are strongly correlated locally, as well as with the corresponding position in the opposite hemisphere. By more precisely studying correlations between one hemisphere and the other, we demonstrated that most of the brain is indeed symmetrically correlated. Regions not correlated include the parietal cortex and the language area, which are known to have lateralized functions.

14.2 Future Work

Many aspects of the work presented in this thesis can be further extended and applied to several other studies. In this section, we depict the perspectives that are opened by the different contributions of our research. We will focus only on the aspects that are of special interest to us. There are four main future directions. First, the question of validation of the variability models of the brain needs to be addressed. Validation of such models is a tough issue, although we believe that data fusion holds the key to such validation. Second, we have the tools to perform an in-depth statistical analysis of DT-MRI. In particular, diffusion tensor atlases need to be built: these are already done by some groups but DT-MRI atlases are not widespread in the community, and even less so in the clinical world, where they are actually needed (to compare populations for instance). Third, variability models can be considered as a prior in inter-subject non-linear registration and could greatly improve the results by guiding algorithms. Finally, we believe that more effort should be made to have a systematic transfer of new methodologies towards end-users. In the following, we detail each of the points raised above.

14.2.1 Validation of Variability Models with Data Fusion

The main problem of validation of our variability models is the lack of ground truth. To overcome this severe limitation, we could think of fusing all sources of brain variability we can collect: not only anatomical MRI (our case) can give information on brain variability, but also diffusion tensor MRI and functional MRI have some secrets to reveal. We could think for instance of learning variability from fiber tracts: this implies to be able to reliably reconstruct and identify (label) the fiber bundles of one subject (like this is done with the sulcal lines in the BrainVisa software). Then, we would need to model fiber bundles in order to learn their inter-subject variability. Fiber bundles are more complex than sulcal lines: some of them can be modeled with a center line (e.g., the splenium and genu tracts), others as 2D surfaces (e.g., corona radiata). We need to find an optimal representation for these entities. Functional MRI is a bit different. Variations in the localization of functions of the brain is not equivalent to anatomical variability. However, it would be very interesting to identify regions where functions and anatomical variability are linked: this could show that function plays a role in determining the shape of brain structures.

14.2.2 Statistical Analysis of DT-MRI

This second perspective is strongly linked to the previous one. The statistical analysis of DT-MRI encompasses DT-MRI atlas construction, study of group differences, and modeling white matter fiber tracts.

14.2.2.1 DT-MRI Atlases

Building diffusion tensor atlases of the brain is a major step towards statistical analysis of DT-MRI. Thanks to the Riemannian frameworks proposed in this thesis, we are now able to compute the average and covariance matrix of a set of tensors. Furthermore, the information from the covariance matrix allows to extend some tools for statistical comparisons such as the Mahalanobis distance (the z-score). The Riemannian metrics could be used to compute a similarity measure between two tensor fields and to optimize this objective function. In particular, the simplicity of Log-Euclidean metrics makes them a good candidate for tensor-tensor registration. We started investigating this with Thye Yeo Boon, a graduate student from MIT.

14.2.2.2 Modeling of White Matter Fiber Bundles

As we said previously, modeling white matter fiber bundles is a big challenge. It is so a necessary step before accessing their variability. Our intuition is that fiber bundles should not be modeled as a collection of single fibers. A single fiber is almost meaningless since there is a lot of uncertainty, for instance on the trajectory of the fiber: due to partial voluming, low spatial resolutions, etc. we are not able to reconstruct meaningful physical fibers. Instead, all fibers composing a bundle

should be considered as a region (volume), which can be modeled for instance with a medial representation, a surface, or a current. In other words, we believe that we should consider fiber bundles as regular structures of the brain (like the amygdala or the hipocampi). Once we are able to reliability extract the same fiber bundle of interest in individuals, we could model its variability.

This raises a second question: how can we extract the same bundles across individuals? The DTI atlases that we discussed previously would greatly help us in doing this. Indeed, if we could build an atlas, we could extract the fiber bundles and represent them as a label map. Then, by projecting this map onto each individual DT image, we could automatically segment a brain with bundles.

14.2.3 Inter-Subject Non-Linear Registration with a Prior on Brain Variability

This perspective aims at incorporating variability models that we have discussed previously back into the registration. The idea is that variability models could act as a prior on the registration: it could tell "how much this specific position is variable among individuals, and in which direction should we look for a corresponding point". This type of prior would "guide" non-linear registration algorithms and could help recovering correspondences. A second idea it to incorporate the results of the correlation analysis presented in Chapter 12 back into the registration as well. The main idea would be to tell the registration algorithm that if a position is moved in a direction, these other positions, spotted as "highly correlated", should be moved as well. This implies to model the correlations to use them in practice. We could think of keeping for each sulcal positions the kernel of locally correlated points, plus the set of mostly-correlated distant points.

14.2.4 More Technological Transfer towards Clinics

Finally, we would like to add a few words about the transfer of technologies developed at INRIA (and in any other research institutes) towards end-users, and in particular to the clinicians. This is a critical aspect of science since both worlds, methodological and clinical, need each others. Clinicians are experts in image acquisition and interpretation. This is a very difficult task and one of the role of methodological research, from our point of view, is to automatize some complicated tasks. Thus, much research aims at addressing specific clinical problems, like segmentation of vital structures for radiotherapy planning, pre- and post-operative registration of images for augmented reality, etc. While our research is fed by clinical challenges, few results are actually transfered back to the clinics (meaning every hospital and not only the one we collaborate with) in a convenient and ready-to-use software package. By developing MedINRIA, we are hoping that more fruits of our research will become available to anyone. We are continuing the development of this software by adding new features: new registration algorithms are underway, a full set of applications for multiple sclerosis segmentation and

tracking over time is being developed. The "DTI Track" module could be extended with automatic fiber bundling algorithms, DT-MRI registration and statistical analysis.

All these research directions will be explored in a near future, hopefully leading to new findings in neuroscience, to a better comprehension of the brain, and finally to new clinical indications.

"The world is round and the place which may seem like the end may also be the beginning."
Ivy Baker Priest.

To be continued...

14.3 Publications of the Author

14.3.1 Methodological Articles

Journal Papers

- P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. *Clinical DT-MRI Estimation, Smoothing and Fiber Tracking with Log-Euclidean Metrics.* IEEE Transactions on Medical Imaging, 2007. Note: In Press.

- P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache. *Measuring Brain Variability by Extrapolating Sparse Tensor Fields Measured on Sulcal Lines.* Neuroimage, 34(2):639-650, January 2007. Note: Also as INRIA Research Report 5887, April 2006.

- V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. *Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices.* SIAM Journal on Matrix Analysis and Applications, 29(1):328-347, 2006.

- V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. *Log-Euclidean Metrics for Fast and Simple Calculus on Diffusion Tensors. Magnetic Resonance in Medicine.* MRM, 56(2):411-421, August 2006.

- X. Pennec, P. Fillard, and N. Ayache. *A Riemannian Framework for Tensor Computing.* International Journal of Computer Vision, 66(1):41-66, January 2006. Note: A preliminary version appeared as INRIA Research Report 5255, July 2004.

Peer-Reviewed Conference Papers

- P. Fillard, X. Pennec, P.M. Thompson, and N. Ayache. *Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines.* In Proc. of MICCAI'07 Workshop on Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation, Brisbane, Australia, 2007.

- P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. *Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics.* In Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006), Crystal Gateway Marriott, Arlington, Virginia, USA, pages 786-789, April 2006.

- V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. *Fast and Simple Calculus on Tensors in the Log-Euclidean Framework.* In J. Duncan and G. Gerig, editors, Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part I, volume 3749 of LNCS, Palm Springs, CA, USA, October 26-29, pages 115-122, 2005. Springer Verlag.

- O. Commowick, R. Stefanescu, P. Fillard, V. Arsigny, N. Ayache, X. Pennec, and G. Malandain. *Incorporating Statistical Measures of Anatomical Variability in Atlas-to-Subject Registration for Conformal Brain Radiotherapy.* In J. Duncan and G. Gerig, editors, Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part II, volume 3750 of LNCS, Palm Springs, CA, USA, October 26-29, pages 927-934, 2005. Springer Verlag.

- P. Fillard, V. Arsigny, N. Ayache, and X. Pennec. *A Riemannian Framework for the Processing of Tensor-Valued Images.* In Ole Fogh Olsen, Luc Florak, and Arjan Kuijper, editors, Deep Structure, Singularities, and Computer Vision (DSSCV), number 3753 of LNCS, pages 112-123, June 2005. Springer Verlag.

- P. Fillard, V. Arsigny, X. Pennec, P. M. Thompson, and N. Ayache. *Extrapolation of Sparse Tensor Fields: Application to the Modeling of Brain Variability.* In Gary Christensen and Milan Sonka, editors, Proc. of Information Processing in Medical Imaging 2005 (IPMI'05), volume 3565 of LNCS, Glenwood springs, Colorado, USA, pages 27-38, July 2005. Springer.

- X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache. *Riemannian Elasticity: A statistical regularization framework for non-linear registration.* In J. Duncan and G. Gerig, editors, Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part II, volume 3750 of LNCS, Palm Springs, CA, USA, October 26-29, pages 943-950, 2005. Springer Verlag.

Technical Reports

- P. Fillard, X. Pennec, P. M. Thompson, and N. Ayache. *Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines.* Research Report 6241, INRIA, July 2007.

- P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache. *Measuring Brain Variability by Extrapolating Sparse Tensor Fields Measured on Sulcal Lines.* Research Report 5887, INRIA, April 2006.

- V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. *Fast and Simple Computations on Tensors with Log-Euclidean Metrics.* Research Report RR-5584, INRIA, Sophia-Antipolis, France, May 2005.

- P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. *Joint Estimation and Smoothing of Clinical DT-MRI with a Log-Euclidean Metric.* Research Report RR-5607, INRIA, Sophia-Antipolis, France, June 2005.

- X. Pennec, P. Fillard, and N. Ayache. *A Riemannian Framework for Tensor Computing.* Research Report 5255, INRIA, July 2004. Note: Published in Int. Journal of Computer Vision 65(1), October 2005.

14.3.2 Medical Articles

Journal Papers

- D. Ducreux, P. Fillard, D. Facon, A. Ozanne, J.-F. Lepeintre, J. Renoux, M. Tadié and P. Lasjaunias. *Diffusion Tensor Magnetic Resonance Imaging and Fiber Tracking in Spinal Cord Lesions: Current and Future Indications.* Neuroimaging Clinics of North America, 17(1):137-147, February 2007.

- D. Ducreux, F. Dhermain, and P. Fillard. *Functional MRI in Brain Tumours.* Cancer Radiotherapy, September 2006.

- D. Ducreux, J.-F. Lepeintre, P. Fillard, C. Loureiro, M. Tadié and P. Lasjaunias. *MR Diffusion Tensor Imaging and Fiber Tracking in 5 Spinal Cord Astrocytomas.* American Journal of Neuroradiology (AJNR), 27:214-216, 2006.

- M. Liberatore, C. Denier, P. Fillard, M.C. Petit-Lacour, F. Benoudiba, P. Lasjaunias, and D. Ducreux. *Diffusion Tensor Imaging and Tractography of Central Pontine Myelinolysis.* Journal of Neuroradiology, 33(3):189-193, June 2006.

- D. Ducreux, I. Huynh, P. Fillard, J. Renoux, M.C. Petit-Lacour, K. Marsot-Dupuch, and P. Lasjaunias. *Brain MR Diffusion Tensor Imaging and Fibre Tracking to Differentiate Between Two Diffuse Axonal Injuries.* Neuroradiology, 47(8):604-608, August 2005.

- D. Facon, A. Ozanne, P. Fillard, J.-F. Lepeintre, C. Tournoux-Facon, and D. Ducreux. *MR Diffusion Tensor Imaging and Fiber Tracking in Spinal Cord Compression.* American Journal of Neuroradiology (AJNR), 26:1587-1594, 2005.

14.3.3 Programming & Software Articles

Peer-Reviewed Workshop Papers

- N. Toussaint, M. Sermesant, and P. Fillard. *vtkINRIA3D: A VTK Extension for Spatiotemporal Data Synchronization, Visualization and Management.* In Proc. of MICCAI'07 Workshop on Open Source and Open Data for MICCAI, Brisbane, Australia, October 2007.

- N. Toussaint, J.C. Souplet, and P. Fillard. *MedINRIA: Medical Image Navigation and Research Tool by INRIA.* In Proc. of MICCAI'07 Workshop on Interaction in medical image analysis and visualization, Brisbane, Australia, 2007.

Other Workshop Papers and Abstracts

- P. Fillard, N. Toussaint, and X. Pennec. *MedINRIA: DT-MRI Processing and Visualization Software.* November 2006. Note: Guest paper at the Similar Tensor Workshop, Las Palmas, Spain.

- J.-C. Souplet, C. Lebrun, P. Fillard, N. Ayache, P. Clavelou, and G. Malandain. *SepINRIA : un outil d'évaluation de la charge lésionnelle chez des patients atteints de sclérose en plaques.* Journées Internationales de la SFN : Sclérose en Plaques (Multiple Sclerosis), Paris, France, June 2007. Note: Abstract in Revue Neurologique (Paris) 2007; 163: 6-7, 725.

14.3.4 Patent

- V. Arsigny, X. Pennec, P. Fillard, and N. Ayache. *Dispositif perfectionné de traitement ou de production d'images de tenseurs.* French patent filing number 0503483, April 2005.

Appendix A

Tensor regularization: the Laplace-Beltrami operator

A.1 Gradient of the L^2 regularization of a tensor field

Let $\Sigma(x) \in \mathcal{S}ym_n^+$ be a tensor field over $\mathbb{R}^d$, and $(x_1, \dots x_d)$ be an orthonormal coordinate system. To simplify the notations, we use in this section ∂_i for the spatial derivative $\partial/(\partial x_i)$ and we do not specify the (spatial) integration variable x. The L^2 regularization criterion is:

$$Reg(\Sigma) = \int_\Omega \|\nabla\Sigma(x)\|^2_{\Sigma(x)}\ dx = \sum_{i=1}^d \int_\Omega \|\partial_i\Sigma\|^2_\Sigma = \sum_{i=1}^d \int_\Omega \mathrm{Tr}\Big((\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\,\Sigma^{(-1)}\Big).$$

Using the Taylor expansion $(\Sigma+\varepsilon W)^{(-1)} = \Sigma^{(-1)} - \varepsilon\Sigma^{(-1)}\,W\,\Sigma^{(-1)} + O(\varepsilon^2)$ in the the Taylor expansion of our regularization criterion and identifying the first order term to $Reg(\Sigma+\varepsilon W) = Reg(\Sigma) + \varepsilon\,\partial_W Reg + O(\varepsilon^2)$, we get the directional derivative:

$$\partial_W Reg = 2\sum_{i=1}^d \int_\Omega \mathrm{Tr}\Big((\partial_i W)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\,\Sigma^{(-1)} - (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\,\Sigma^{(-1)}\,W\Sigma^{(-1)}\Big)$$

The main goal is to find out the field of tangent vectors $\nabla Reg(x) \in T_{\Sigma(x)}\mathcal{S}ym_n^+$ such that, by definition of the gradient, we have the equality: $\partial_W Reg = \int_\Omega \langle\, W \mid \nabla Reg\rangle_\Sigma\ dx$ for every field of tangent vectors $W(x) \in T_{\Sigma(x)}\mathcal{S}ym_n^+$. As the above expression of $\partial_W Reg$ is in the standard Euclidean chart (matrix coefficients), we shall safely use the computations of the previous sections. Notice that we are using the matrix coefficients only as a chart and not as a metric. Let $\Lambda_i = \Sigma^{(-1)}\,(\partial_i\Sigma)\,\Sigma^{(-1)}$. We get:

$$\partial_W Reg = 2\sum_{i=1}^d \int_\Omega \Big(\mathrm{Tr}((\partial_i W)\,\Lambda_i) - \big\langle\, W \;\big|\; (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\big\rangle_\Sigma\Big)$$

Now, assuming the proper Neumann boundary conditions, we can apply the previous integration by part formula $\int_\Omega \mathrm{Tr}((\partial_i W)\,\Lambda_i) = -\int_\Omega \mathrm{Tr}(W\,(\partial_i\Lambda_i))$ to the first term:

$$\begin{aligned}
\partial_W Reg &= -2\sum_{i=1}^d \int_\Omega \Big(\mathrm{Tr}\Big(W\Sigma^{(-1)}\,(\Sigma(\partial_i\Lambda_i)\Sigma)\,\Sigma^{(-1)}\Big) + \big\langle\, W \;\big|\; (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\big\rangle_\Sigma\Big) \\
&= -2\sum_{i=1}^d \int_\Omega \big\langle\, W \;\big|\; \Sigma(\partial_i\Lambda_i)\Sigma + (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\big\rangle_\Sigma
\end{aligned}$$

We have obtained the expression that defines the gradient of our regularization criterion:

$$\nabla Reg = -2\sum_{i=1}^{d}\Big(\Sigma\,(\partial_i\Lambda_i)\,\Sigma + (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)\Big)$$

To compute explicitly its value, let us observe first that $\partial_i\Sigma^{(-1)} = -\Sigma^{(-1)}\,(\partial_i\Sigma)\,\Sigma^{(-1)}$ because $\partial_i(\Sigma^{(-1)}\,\Sigma) = 0$. Thus, thanks to the chain rule, we have:

$$\Sigma\,(\partial_i\Lambda_i)\,\Sigma = \Sigma\,\partial_i\left(\Sigma^{(-1)}\,(\partial_i\Sigma)\,\Sigma^{(-1)}\right)\,\Sigma = \partial_i^2\Sigma - 2\,(\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)$$

Eventually, we end up with $\nabla Reg(\Sigma) = -2\Delta\Sigma$, where Δ is the Laplace-Beltrami operator on our manifold:

$$\Delta\Sigma = \sum_{i=1}^{d}\Delta_i\Sigma \qquad \text{with} \qquad \Delta_i\Sigma = \partial_i^2\Sigma - (\partial_i\Sigma)\,\Sigma^{(-1)}\,(\partial_i\Sigma)$$

A.2 Numerical Implementation of the Laplace-Beltrami Operator

From the Taylor expansion of a tensor field (considered as a matrix field) Σ at x, we have $\Sigma(x+\varepsilon u) = \Sigma(x) + \varepsilon\,\partial_u\Sigma(x) + \varepsilon^2\,\partial_u^2\Sigma(x)/2 + \varepsilon^3\,\partial_u^3\Sigma(x)/6 + O(\varepsilon^4)$. Thus, we may approximate the first and second order tensor derivatives by their Euclidean derivatives:

$$\begin{aligned}\partial_u\Sigma(x) &= \frac{1}{2}\big(\Sigma(x+u) - \Sigma(x-u)\big) + O(\|u\|^3)\\ \partial_u^2\Sigma(x) &= \big(\Sigma(x+u) - \Sigma(x)\big) + \big(\Sigma(x-u) - \Sigma(x)\big) + O(\|u\|^4)\end{aligned}$$

This finally gives us a fourth order approximation of the Laplace-Beltrami operator in the spatial direction u:

$$\begin{aligned}\Delta_u\Sigma(x) &= \partial_u^2\Sigma - 2\,(\partial_u\Sigma)\,\Sigma^{(-1)}\,(\partial_u\Sigma)\\ &= \Sigma(x+u) + \Sigma(x-u) - 2\Sigma(x)\\ &\quad -\frac{1}{2}\big(\Sigma(x+u) - \Sigma(x-u)\big)\Sigma^{(-1)}\big(\Sigma(x+u) - \Sigma(x-u)\big) + O(\|u\|^4)\end{aligned}$$

However, this numerical scheme is extrinsic since it is based on (Euclidean) differences of tensors. We propose here an intrinsic scheme based on the exponential chart at the current point: we claim that $\overrightarrow{\Sigma(x)\Sigma(x+u)} + \overrightarrow{\Sigma(x)\Sigma(x-u)}$ is a forth order approximation of the Laplace Beltrami operator in the direction u. Indeed, we have

$$\begin{aligned}\overrightarrow{\Sigma(x)\Sigma(x+\varepsilon u)} &= \Sigma^{\frac{1}{2}}(x)\,\log\left(\Sigma^{-\frac{1}{2}}(x)\,\Sigma(x+\varepsilon u)\,\Sigma^{-\frac{1}{2}}(x)\right)\Sigma^{\frac{1}{2}}(x)\\ &= \Sigma^{\frac{1}{2}}(x)\,\log\left(\mathrm{Id} + \varepsilon W + \frac{\varepsilon^2}{2}H + O(\varepsilon^3)\right)\Sigma^{\frac{1}{2}}(x)\end{aligned}$$

where we put $W = \Sigma^{-\frac{1}{2}}\,\partial_u\Sigma\,\Sigma^{-\frac{1}{2}}$ and $H = \Sigma^{-\frac{1}{2}}\,\partial_u^2\Sigma\,\Sigma^{-\frac{1}{2}}$. From the Log series (Eq. 3.1), we get:

$$\overrightarrow{\Sigma(x)\Sigma(x+\varepsilon u)} \;=\; \Sigma^{\frac{1}{2}}\left[+\varepsilon W + \frac{\varepsilon^2}{2}H - \frac{1}{2}\left(\varepsilon^2 W^2 + \frac{\varepsilon^3}{2}(WH+HW)\right) + \frac{\varepsilon^3}{3}W^3 + O(\varepsilon^4)\right]\Sigma^{\frac{1}{2}}$$

The Taylor expansion of $\overrightarrow{\Sigma(x)\Sigma(x-\varepsilon u)}$ is obtained by replacing ε by $-\varepsilon$, so that we finally end up with $\overrightarrow{\Sigma(x)\Sigma(x+\varepsilon u)} + \overrightarrow{\Sigma(x)\Sigma(x-\varepsilon u)} = \Sigma^{\frac{1}{2}}\left[\varepsilon^2 H - \varepsilon^2 W^2 + O(\varepsilon^4)\right]\Sigma^{\frac{1}{2}}$, which proves that

$$\overrightarrow{\Sigma(x)\Sigma(x+u)} + \overrightarrow{\Sigma(x)\Sigma(x-u)} = \partial_u^2\Sigma - (\partial_u\Sigma)\Sigma^{(-1)}(\partial_u\Sigma) + O(\|u\|^4) = \Delta_u\Sigma + O(\|u\|^4)$$

Appendix B

Estimation of the Affine Transformation from Sulcal Correspondences

In Sec. 11.1.2, we gave the solution of a system to estimate the affine transformation for each subject based on the correspondence mappings from each sulci to the mean representation. We show here how to obtain this solution (actually, we show it only for the linear component of the affine transformation, the translation was discussed in Sec. 11.1.2). This system to solve is:

$$C(A) = \sum_{i=1}^{N} \int_0^1 \|Ay_i(s) - z_i(s)\|^2 ds,$$

where A is the linear component of the affine transformation we want to estimate, y_i is the sulcus i and z_i the mean sulcus. By doing a Taylor expansion, we get:

$$\begin{aligned}
C(A+dA) &= \sum_{i=1}^{N} \int_0^1 \langle (A+dA)y_i(s) - z_i(s) | (A+dA)y_i(s) - z_i(s) \rangle \, ds \\
&= \sum_{i=1}^{N} \left[\int_0^1 \langle Ay_i(s) - z_i(s) | Ay_i(s) - z_i(s) \rangle \, ds + \right. \\
&\quad \left. + 2 \int_0^1 \langle Ay_i(s) - z_i(s) | dAy_i(s) \rangle \, ds + O(\|dA\|^2) \right] \\
C(A+dA) &= C(A) + 2 \sum_{i=1}^{N} \int_0^1 \langle Ay_i(s) - z_i(s) | dAy_i(s) \rangle \, ds + O(\|dA\|^2)
\end{aligned}$$

By analogy with the Taylor expansion of C: $C(A+dA) = C(A) + < \nabla C(A)|dA > + O(\|dA\|^2)$, we obtain:

$$\begin{aligned}
\langle \nabla C(A)|dA \rangle &= 2\sum_{i=1}^{N} \int_0^1 \langle Ay_i(s) - z_i(s)|dAy_i(s)\rangle \, ds \\
&= 2\sum_{i=1}^{N} \int_0^1 \text{Trace}\left((Ay_i(s) - z_i(s)) \, (dAy_i(s))^\top \right) ds \\
&= 2\sum_{i=1}^{N} \int_0^1 \text{Trace}\left(Ay_i(s)y_i(s)^\top dA^\top - z_i(s)y_i(s)^\top dA^\top \right) ds \\
&= 2\sum_{i=1}^{N} \int_0^1 \left\langle Ay_i(s)y_i(s)^\top - z_i(s)y_i(s)^\top | dA \right\rangle ds
\end{aligned}$$

By identification, we get $\nabla C(A)$:

$$\nabla C(A) = 2\sum_{i=1}^{N} \int_0^1 \left(Ay_i(s)y_i(s)^\top - z_i(s)y_i(s)^\top \right) ds.$$

Finally, solving $\nabla C(A) = 0$ gives:

$$A = \left(\sum_{i=1}^{N} \int_0^1 z_i(s)y_i(s)^\top ds \right) \left(\sum_{i=1}^{N} \int_0^1 y_i(s)y_i(s)^\top ds \right)^{-1}.$$

Appendix C

Practical Implementation of the Matrix Exponential Directional Derivative

In Sec. 6.2, the directional derivative of the exponential $\partial_{\mathbf{G}} \exp(\mathbf{L})$ is used. For general matrices, one has to compute the series [Arsigny 2006c]:

$$\partial_{\mathbf{G}} \exp(\mathbf{L}) = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \mathbf{L}^i \mathbf{G} \mathbf{L}^{k-i-1}. \tag{C.1}$$

In general, Eq. [C.1] cannot be simplified. However, in the case of symmetric matrices, the differential can take a much simpler form. We know that $\mathbf{G}$ and $\mathbf{L}$ are both symmetric matrices: In our application, $\mathbf{G}$ is the tensor product of a vector $\mathbf{g}$ with itself: $\mathbf{G} = \mathbf{g}\mathbf{g}^\top$, which is by construction symmetric, and $\mathbf{L}$ is the matrix logarithm of a diffusion tensor, therefore symmetric.

Let $\mathbf{L} = \mathbf{R}^\top \mathbf{S} \mathbf{R}$ be an eigen decomposition of $\mathbf{L}$. $\mathbf{S}$ is diagonal: $\mathbf{S} = \mathrm{diag}(s_1, s_2, s_3)$. From Eq. [C.1], we have:

$$\begin{aligned}
\partial_{\mathbf{G}} \exp(\mathbf{L}) &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \left(\mathbf{R}^\top \mathbf{S}^i \mathbf{R}\right) \mathbf{G} \left(\mathbf{R}^\top \mathbf{S}^{k-i-1} \mathbf{R}\right) \\
&= \mathbf{R}^\top \left(\sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \mathbf{S}^i \left(\mathbf{R}\mathbf{G}\mathbf{R}^\top\right) \mathbf{S}^{k-i-1} \right) \mathbf{R} && \text{(C.2)} \\
&= \mathbf{R}^\top \left(\partial_{\mathbf{R}\mathbf{G}\mathbf{R}^\top} \exp\left(\mathbf{S}\right)\right) \mathbf{R} && \text{(C.3)}
\end{aligned}$$

We have $\partial_{\mathbf{G}} \exp(\mathbf{L}) = \mathbf{R}^\top \partial_{\mathbf{R}\mathbf{G}\mathbf{R}^\top} \exp(\mathbf{S}) \mathbf{R}$. Let $\mathbf{M}$ be $\partial_{\mathbf{R}\mathbf{G}\mathbf{R}^\top} \exp\left(\mathbf{S}\right)$. We denote by $[\mathbf{M}]_{(l,m)}$ the coefficient (l, m) of matrix $\mathbf{M}$ (i.e. the matrix coefficient stored at row l and column m). $(l, m) \in \{1, 2, 3\}$ for diffusion tensors. From Eq. [C.2] we can express the coefficient (l, m) of $\mathbf{M}$ as:

$$[\mathbf{M}]_{(l,m)} = \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} \left[\mathbf{S}^i \mathbf{R}\mathbf{G}\mathbf{R}^\top \mathbf{S}^{k-i-1}\right]_{(l,m)}.$$

As $\mathbf{S}$ is diagonal, we can further write:

$$\begin{aligned}[\mathbf{M}]_{(l,m)} &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s_l^i \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} s_m^{k-i-1} \\ &= \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s_l^i s_m^{k-i-1}.\end{aligned}$$

Now, we only need to express the series $\sum_{k=1}^{+\infty} 1/k! \sum_{i=0}^{k-1} s_l^i s_m^{k-i-1}$. If $s_l \neq s_m$, we have:

$$\begin{aligned}\sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s_l^i s_m^{k-i-1} &= \sum_{k=1}^{+\infty} \frac{1}{k!} s_m^{k-1} \sum_{i=0}^{k-1} \left(\frac{s_l}{s_m}\right)^i \\ &= \sum_{k=1}^{+\infty} \frac{1}{k!} s_m^{k-1} \frac{1-(s_l/s_m)^k}{1-s_l/s_m} \\ &= \frac{1}{s_m - s_l} \sum_{k=1}^{+\infty} \frac{1}{k!} s_m^k \frac{s_m^k - s_l^k}{s_m^k} \\ &= \frac{\exp(s_m) - \exp(s_l)}{s_m - s_l}.\end{aligned}$$

If $sl = s_m = s$, then the series can be even more simplified:

$$\begin{aligned}\sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s^i s^{k-i-1} &= \sum_{k=1}^{+\infty} \frac{1}{k!} \sum_{i=0}^{k-1} s^{k-1} = \sum_{k=1}^{+\infty} \frac{1}{k!} \left(k s^{k-1}\right) \\ &= \sum_{k=1}^{+\infty} \frac{s^{k-1}}{(k-1)!} = \sum_{k=0}^{+\infty} \frac{s^k}{k!} \\ &= \exp(s).\end{aligned}$$

Finally, one can access the coefficient (l, m) of $\mathbf{M}$ as:

$$\begin{aligned}[M]_{(l,m)} &= \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} \frac{e^{s_m} - e^{s_l}}{s_m - s_l} \text{ if } s_l \neq s_m, \\ &= \left[\mathbf{R}\mathbf{G}\mathbf{R}^\top\right]_{(l,m)} e^{s_l} \text{ if } s_l = s_m. \end{aligned} \tag{C.4}$$

From Eq. [C.4], and Eq. [C.2], we deduce the directional derivative of the matrix exponential $\partial_{\mathbf{G}} \exp(\mathbf{L})$.

Appendix D

Practical Implementation of $\alpha(x)$

In Sec. 6.2.3, the coefficient $\alpha(x) = I_0'/I_0(x)$ is needed to be able to compute the gradients of the ML and MAP estimators. α takes its values between 0 and 1. A practical implementation of $\alpha(x)$ is given in the following.

The 0^{th} order modified Bessel function of the first kind $I_0(x)$ is expressed as follows: $I_0(x) = \sum_{k=0}^{+\infty} \frac{(x^2/4)^k}{k!^2}$. Consequently, $\alpha(x)$ can be written as:

$$\alpha(x) = \frac{I_0'(x)}{I_0(x)} = \frac{x}{2} \frac{\sum_{k=0}^{+\infty} \frac{\left(\frac{1}{4}x^2\right)^k}{k!(k-1)!}}{\sum_{k=0}^{+\infty} \frac{\left(\frac{1}{4}x^2\right)^k}{k!^2}}. \tag{D.1}$$

Unfortunately, Eq. [D.1] cannot be simplified. However, for small values of x (typically $x \leq 10$) the series converges very quickly, and in practice 10 terms of the series are sufficient for an accuracy of 1.10^{-12}, which is by far enough. For higher values of x, one may experience numerical stability issues. To prevent this, one may use the following Taylor expansion of $\alpha(x)$ valid for $x > 10$:

$$\begin{aligned}\alpha(x) \quad &\simeq 1.0 - 0.5x^{-1} - 0.125x^{-2} - 0.125x^{-3} - 0.1953125x^{-4} \\ &-0.40625x^{-5} - 1.0478515625x^{-6} - 3.21875x^{-7} \\ &-11.46646118164x^{-8} - 46.478515625x^{-9} \\ &-211.276149749755x^{-10} - 1064.67822265625x^{-11}\end{aligned}$$

Bibliography

[Aja-Fernández 2007] S. Aja-Fernández, C. Alberola-López et C.-F. Westin. *Signal LMMSE estimation from multiple samples in MRI and DT-MRI.* In Proc. of MICCAI'07, volume 4792 of *LNCS*, pages 368–375, 2007. 153

[Alexander 2001a] A. L. Alexander, K. M. Hasan, M. Lazar, J. S. Tsuruda et D. L. Parker. *Analysis of partial volume effects in diffusion-tensor MRI.* Magnetic Resonance in Medicine, vol. 45, no. 5, pages 770–780, 2001. 7, 25

[Alexander 2001b] D.C. Alexander, C. Pierpaoli, P.J. Basser et J.C. Gee. *Spatial Transformations of Diffusion Tensor Magnetic Resonance Images.* IEEE Transactions on Medical Imaging, vol. 20, no. 11, pages 1131–1139, Nov 2001. 133

[Anderson 2001] Adam W. Anderson. *Theoretical analysis of the effects of noise on diffusion tensor imaging.* Magnetic Resonance in Medicine, vol. 46, no. 6, pages 1174–1188, December 2001. 114

[Arsigny 2005a] Vincent Arsigny, Pierre Fillard, Xavier Pennec et Nicholas Ayache. *Fast and Simple Calculus on Tensors in the Log-Euclidean Framework.* In J. Duncan et G. Gerig, editors, Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part I, volume 3749 of *LNCS*, pages 115–122, Palm Springs, CA, USA, October 26-29, 2005. Springer Verlag. PMID: 16685836. 15, 34

[Arsigny 2005b] Vincent Arsigny, Pierre Fillard, Xavier Pennec et Nicholas Ayache. *Fast and Simple Computations on Tensors with Log-Euclidean Metrics.* Research Report RR-5584, INRIA, Sophia-Antipolis, France, May 2005. 15, 34, 55, 57

[Arsigny 2006a] Vincent Arsigny, Olivier Commowick, Xavier Pennec et Nicholas Ayache. *A Log-Euclidean Framework for Statistics on Diffeomorphisms.* In Proc. of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Part I, number 4190 of LNCS, pages 924–931, 2-4 October 2006. PMID: 17354979. 11, 30

[Arsigny 2006b] Vincent Arsigny, Pierre Fillard, Xavier Pennec et Nicholas Ayache. *Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices.* SIAM Journal on Matrix Analysis and Applications, vol. 29, no. 1, pages 328–347, 2006. 15, 34

[Arsigny 2006c] Vincent Arsigny, Pierre Fillard, Xavier Pennec et Nicholas Ayache. *Log-Euclidean Metrics for Fast and Simple Calculus on Diffusion Tensors.* Magnetic Resonance in Medicine, vol. 56, no. 2, pages 411–421, August 2006. PMID: 16788917. 15, 34, 120, 217

[Ashburner 2005] J. Ashburner et K. J. Friston. *Unified segmentation.* NeuroImage, vol. 26, no. 3, pages 839–851, 2005. 157

[Assaf 2005] Y. Assaf et P. J. Basser. *Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain.* Magnetic Resonance in Medicine, vol. 27, no. 1, pages 48–58, 2005. 7, 25, 97

[Bakircioglu 1998] M. Bakircioglu, U. Grenander, N. Khaneja et M.I. Miller. *Curve Matching on Brain Surfaces Using Induced Frenet Distance Metrics.* In Human Brain Mapping, volume 6, pages 329–331, 1998. 165

[Bammer 2000] R. Bammer, F. Fazekas, M. Augustin, J. Simbrunner, S. Strasser-Fuchs, T. Seifert, R. Stollberger et H.-P. Hartung. *Diffusion-Weighted MR Imaging of the Spinal Cord.* American Journal of Neuroradiology, vol. 21, pages 587–591, 2000. 138

[Barmpoutis 2007a] A. Barmpoutis, B. Jian, B. C. Vemuri et T. M. Shepherd. *Symmetric Positive 4^{th} Order Tensors & Their Estimation from Diffusion Weighted MRI.* In Proc. of IPMI, volume 4584 of *LNCS*, pages 308–319, 2007. 97

[Barmpoutis 2007b] A. Barmpoutis et B.C. Vemuri. *Exponential Tensors: A Framework for Efficient Higher-Order DT-MRI Computations.* In Proc. of ISBI'07, pages 792–795, 2007. 7, 25

[Basser 1994a] P. Basser, J. Mattiello et D. Le Bihan. *MR diffusion tensor spectroscopy and imaging.* Biophysical Journal, vol. 66, pages 259–267, 1994. 4, 7, 23, 26, 43, 115, 138

[Basser 1994b] P.J. Basser, J. Mattiello et D. Le Bihan. *Estimation of the effective self-diffusion tensor from the NMR spin echo.* Journal of Magnetic Resonance, vol. 103, pages 247–254, 1994. 99, 100

[Basser 1996] P. Basser et C. Pierpaoli. *Microstructural and physiological features of tissues elucidated by quantitative diffusion tensor MRI.* Journal of Magnetic Resonance, vol. 111, pages 209–219, 1996. 138

[Basser 2000] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda et A. Aldroubi. *In vivo fiber tractography using DT-MRI data.* Magnetic Resonance in Medicine, vol. 44, pages 625–632, 2000. 140

[Basu 2006] S. Basu, T. Fletcher et R. Whitaker. *Rician Noise Removal in Diffusion Tensor MRI.* In Proc. of MICCAI'06, volume 4190 of *LNCS*, pages 117–125. Springer, October 2006. 101

[Baumberg 1994] Adam Baumberg et David Hogg. *Learning Flexible Models from Image Sequences.* In ECCV (1), pages 299–308, 1994. 165

[Bhatia 2003] R. Bhatia. *On the exponential metric increasing property.* Linear Algebra and its Applications, vol. 375, pages 211–220, 2003. 48

[Bigün 1991] J. Bigün, G. H. Granlund et J. Wiklund. *Multidimensional orientation estimation with applications to texture analysis and optical flow.* IEEE Transactions on pattern Analysis and Machine Intelligence, vol. 13, no. 8, pages 775–790, August 1991. 91

[Bookstein 1991] F. Bookstein. Morphometric tools for landmark data: Geometry and biology. Cambridge University Press, 1991. 164

[Bourbaki 1989] Nicolas Bourbaki. Elements of mathematics: Lie groups and lie algebra, chapters 1-3. Springer-Verlag, 2nd printing édition, 1989. 54

[Brodmann 1905] K. Brodmann. *Beiträge zur histologischen Lokalisation der Grosshirnrinde: dritte Mitteilung: Die Rindenfelder der niederen Affen.* Journal für Psychologie und Neurologie, vol. 4, pages 177–226, 1905. 160

[Cachia 2003] A. Cachia, J.-F. Mangin, D. Rivière, D. Papadopoulos-Orfanos, F. Kherif, I. Bloch et J. Régis. *A generic framework for parcellation of the cortical surface into gyri using geodesic Voronoï diagrams.* Medical Image Analysis, vol. 7, no. 4, pages 403–416, 2003. 11, 30

[Cannon 2005] T.D. Cannon, W. Hennah, T.G.M. van Erp, P.M. Thompson, J. Lonnqvist, M. Huttunen, T. Gasperoni, A. Tuulio-Henriksson, T. Pirkola, A. W. Toga, J. Kaprio, J. Mazziotta et L. Peltonen. *Association of DISC1/TRAX Haplotypes With Schizophrenia, Reduced Prefrontal Gray Matter, and Impaired Short- and Long-term Memory.* Arch Gen Psychiatry, vol. 62, pages 1205–1213, 2005. 157

[Cassol 2004] E. Cassol, J.-P. Ranjeva, D. Ibarrola, C. Mékies, C. Manelfe, M. Clanet et I. Berry. *Diffusion tensor imaging in multiple sclerosis: a tool for monitoring changes in normal-appearing white matter.* Multiple Sclerosis, vol. 10, pages 188–196, 2004. 147

[Castaño-Moraga 2006] C. A. Castaño-Moraga, C. Lenglet, R. Deriche et J. Ruiz-Alzola. *A fast and rigorous anisotropic smoothing method for DT-MRI.* In Proc. of ISBI, pages 93–96, 2006. 48

[Castaño-Moraga 2007] C. A. Castaño-Moraga, C. Lenglet, R. Deriche et J. Ruiz-Alzola. *A Riemannian Approach to Anistropic Filtering of Tensor Fields.* Signal Processing, special issue on tensor signal processing, vol. 87, no. 2, pages 263–76, 2007. 48

[Cazals 2001] F. Cazals et J.-D. Boissonnat. *Natural Coordinates of Points on a Surface.* Comp. Geometry Theory and Applications, vol. 19, pages 155–173, 2001. 67

[Cercignani 2003] M. Cercignani, M. A. Horsfield, F. Agosta et M. Filippi. *Sensitivity-Encoded Diffusion Tensor MR Imaging of the Cervical Cord.* American Journal of Neuroradiology, vol. 24, pages 1254–1256, June 2003. 138

[Chang 2005] Lin-Ching Chang, Derek K. Jones et Carlo Pierpaoli. *RESTORE: Robust Estimation of Tensors by Outlier Rejection.* Magnetic Resonance in Medicine, vol. 53, no. 5, pages 1088–1095, 2005. 100, 107

[Chefd'hotel 2002] C. Chefd'hotel, C. Tschumperlé, R. Deriche et O. Faugeras. *Constrained flows of matrix-valued functions: Application to diffusion tensor regularization.* In Proc. of ECCV'02, volume 2350 of *LNCS*, pages 251–265. Springer, 2002. 109

[Chefd'hotel 2004] C. Chefd'hotel, D. Tschumperlé, R. Deriche et O. Faugeras. *Regularizing Flows for Constrained Matrix-Valued Images.* Journal of Mathematical Imaging and Vision, vol. 20, no. 1-2, pages 147–162, 2004. 102

[Ciccarelli 2003] O. Ciccarelli, D. J. Werring, G. J. Barker, C. M. Griffin, C. A. M. Wheeler-Kingshott, D. H. Miller et A. J. Thompson. *A study of the mechanisms of normal-appearing white matter damage in multiple sclerosis using diffusion tensor imaging Evidence of Wallerian degeneration.* Journal of Neurology, vol. 250, no. 3, March 2003. 147

[Clark 2000] C. A. Clark, D. J. Werring et D. H. Miller. *Diffusion imaging of the spinal cord in vivo: Estimation of the principal diffusivities and application to multiple sclerosis.* Magnetic Resonance in Medicine, vol. 43, no. 1, pages 133–138, January 2000. 138

[Clatz 2005] Olivier Clatz, Maxime Sermesant, Pierre-Yves Bondiau, Hervé Delingette, Simon K. Warfield, Grégoire Malandain et Nicholas Ayache. *Realistic Simulation of the 3D Growth of Brain Tumors in MR Images Coupling Diffusion with Mass Effect.* IEEE Transactions on Medical Imaging, vol. 24, no. 10, pages 1334–1346, October 2005. 154

[Cointepas 2001] Y. Cointepas, J.-F. Mangin, L. Garnero, J.-B. Poline et H. Benali. *BrainVISA: Software platform for visualization and analysis of multimodality brain data.* NeuroImage, vol. 13, no. 6, 2001. 125

[Collins 1995] D. Collins, C. Holmes, T. Peters et A. Evans. *Automatic 3D model-based neuroanatomical segmentation.* Human Brain Mapping, vol. 3, no. 3, pages 190–208, 1995. 162, 188

[Cook 2006] P. A. Cook, Y. Bai, S. Nedjati-Gilani, K. K. Seunarine, M. G. Hall, G. J. Parker et D. C. Alexander. *Camino: Open-Source Diffusion-MRI Reconstruction and Processing.* In Proc. of the 14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine, page 2759, Seattle, WA, USA, May 2006. 125

[Cootes 1995] T. F. Cootes, C. J. Taylor, D. H. Cooper et J. Graham. *Active shape models-their training and application.* Comput. Vis. Image Underst., vol. 61, no. 1, pages 38–59, 1995. 164

[Corouge 2003] I. Corouge, P. Hellier, B. Gibaud et C. Barillot. *Interindividual functional mapping: a nonlinear local approach.* NeuroImage, vol. 19, no. 4, pages 1337–1348, August 2003. 160

[Corouge 2006] I. Corouge, P.T. Fletcher, S. Joshi, S. Gouttard et G. Gerig. *Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis.* Medical Image Analysis, vol. 10, no. 5, pages 786–798, 2006. 11, 30, 154

[Coulon 2004] O. Coulon, D. Alexander et S. Arridge. *Diffusion tensor magnetic resonance image regularization.* Medical Image Analysis, vol. 8, no. 1, pages 47–67, 2004. 43, 102

[Culver 1966] W. J. Culver. *On the existence and uniqueness of the real logarithm of a matrix.* Proc. of the American Mathematical Society, vol. 17, no. 5, pages 1146–1151, October 1966. 54

[de Luis-Garcìa 2005] R. de Luis-Garcìa, R. Deriche, M. Rousson et C. Alberola-López. *Tensor Processing for Texture and Colour Segmentation.* In Proc. of SCIA, pages 1117–1127, 2005. 91

[de Luis-Garcìa 2008] R. de Luis-Garcìa, R. Deriche et C. Alberola-López. *Texture and color segmentation based on the combined use of the structure tensor and the image components.* Signal Processing, vol. 88, no. 4, pages 776–795, 2008. 91

[Demir 2003] A. Demir, M. Ries, C. T. W. Moonen, J.-M. Vital, J. Dehais, P. Arne, J.-M. Caillé et V. Dousset. *Diffusion-weighted MR Imaging with Apparent Diffusion Coefficient and Apparent Diffusion Tensor Maps in Cervical Spondylotic Myelopathy.* Radiology, vol. 229, page 37, 2003. 138

[Deriche 2004] R. Deriche, D. Tschumperlé et C. Lenglet. *DT-MRI Estimation, Regularization and Fiber Tractography.* In Proc. of ISBI'04, pages 9–12, 2004. 100

[Descoteaux 2006] M. Descoteaux, E. Angelino, S. Fitzgibbons et R. Deriche. *Apparent Diffusion Coefficients from High Angular Resolution Diffusion Images: Estimation and Applications.* MRM, vol. 56, no. 2, pages 395–410, 2006. 97

[Descoteaux 2007a] M. Descoteaux, E. Angelino, S. Fitzgibbons et R. Deriche. *Regularized, Fast and Robust Analytical Q-Ball Imaging.* MRM, vol. 58, no. 3, pages 497–510, 2007. 97

[Descoteaux 2007b] M. Descoteaux, R. Deriche et A. Anwander. *Deterministic and Probabilistic Q-Ball Tractography: from Diffusion to Sharp Fiber Distributions.* Technical report, INRIA, August 2007. 7, 25

[Ducreux 2005] D. Ducreux, G. Nasser, C. Lacroix, D. Adams et P. Lasjaunias. *MR Diffusion Tensor Imaging, Fiber Tracking, and Single-Voxel Spectroscopy Findings in an Unusual MELAS Case.* American Journal of Neuroradiology, vol. 26, pages 1840–1844, August 2005. 149

[Ducreux 2007] D. Ducreux, P. Fillard, D. Facon, A. Ozanne, J.-F. Lepeintre, J. Renoux, M. Tadié et P. Lasjaunias. *Diffusion Tensor Magnetic Resonance Imaging and Fiber Tracking in Spinal Cord Lesions: Current and Future Indications.* Neuroimaging Clinics of North America, vol. 17, no. 1, pages 137–147, February 2007. 15, 35, 137

[Durrlemann 2007] S. Durrlemann, X. Pennec, A. Trouvé et N. Ayache. *Measuring Brain Variability via Sulcal Lines Registration: a Diffeomorphic Approach.* In Proc. of MICCAI'07., volume 4791 of *LNCS*, pages 675–682, 2007. 198

[Einstein 1956] A. Einstein. Investigations on the theory of the brownian movement. New York: Dover, 1956. 5, 23

[Facon 2005] D. Facon, A. Ozanne, P. Fillard, J.-F. Lepeintre, C. Tournoux-Facon et D. Ducreux. *MR Diffusion Tensor Imaging and Fiber Tracking in Spinal Cord Compression.* American Journal of Neuroradiology, vol. 26, pages 1587–1594, June-July 2005. 115, 138, 145

[Fillard 2003] P. Fillard, J. Gilmore, J. Piven, W. Lin et G. Gerig. *Quantitative analysis of white matter fiber properties along geodesic paths.* In Proc. of MICCAI'03, volume 2879 of *LNCS*, pages 16–23. Springer, 2003. 119

[Fillard 2005a] Pierre Fillard, Vincent Arsigny, Nicholas Ayache et Xavier Pennec. *A Riemannian Framework for the Processing of Tensor-Valued Images.* In Ole Fogh Olsen, Luc Florak et Arjan Kuijper, editors, Deep Structure, Singularities, and Computer Vision (DSSCV), number 3753 of LNCS, pages 112–123. Springer Verlag, June 2005. 15, 34

[Fillard 2005b] Pierre Fillard, Vincent Arsigny, Xavier Pennec et Nicholas Ayache. *Joint Estimation and Smoothing of Clinical DT-MRI with a Log-Euclidean Metric.* Research Report RR-5607, INRIA, Sophia-Antipolis, France, June 2005. 15, 35

[Fillard 2005c] Pierre Fillard, Vincent Arsigny, Xavier Pennec, Paul M. Thompson et Nicholas Ayache. *Extrapolation of Sparse Tensor Fields: Application to the Modeling of Brain Variability.* In Gary Christensen et Milan Sonka, editors, Proc. of Information Processing in Medical Imaging 2005 (IPMI'05), volume 3565 of *LNCS*, pages 27–38, Glenwood springs, Colorado, USA, July 2005. Springer. PMID: 17354682. 16, 35

[Fillard 2006a] Pierre Fillard, Vincent Arsigny, Xavier Pennec et Nicholas Ayache. *Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean*

metrics. In Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2006), pages 786–789, Crystal Gateway Marriott, Arlington, Virginia, USA, April 2006. 15, 35

[Fillard 2006b] Pierre Fillard, Vincent Arsigny, Xavier Pennec, Kiralee M. Hayashi, Paul M. Thompson et Nicholas Ayache. *Measuring Brain Variability by Extrapolating Sparse Tensor Fields Measured on Sulcal Lines*. Research Report 5887, INRIA, April 2006. 16, 35

[Fillard 2007a] P. Fillard, X. Pennec, P.M. Thompson et N. Ayache. *Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines*. In Proc. of MICCAI'07 Workshop on Statistical Registration: Pairwise and Group-wise Alignment and Atlas Formation, Brisbane, Australia, 2007. 16, 35

[Fillard 2007b] Pierre Fillard, Vincent Arsigny, Xavier Pennec et Nicholas Ayache. *Clinical DT-MRI Estimation, Smoothing and Fiber Tracking with Log-Euclidean Metrics*. IEEE Transactions on Medical Imaging, vol. 26, no. 11, pages 1472–1482, November 2007. PMID: 18041263. 15, 35

[Fillard 2007c] Pierre Fillard, Vincent Arsigny, Xavier Pennec, Kiralee M. Hayashi, Paul M. Thompson et Nicholas Ayache. *Measuring Brain Variability by Extrapolating Sparse Tensor Fields Measured on Sulcal Lines*. Neuroimage, vol. 34, no. 2, pages 639–650, January 2007. Also as INRIA Research Report 5887, April 2006. PMID: 17113311. 16, 35

[Fillard 2007d] Pierre Fillard, Xavier Pennec, Paul M. Thompson et Nicholas Ayache. *Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines*. Research Report 6241, INRIA, 07 2007. 16, 35

[Fischl 2002] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen et A. M. Dale. *Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain*. Neuron, vol. 33, pages 341–355, 2002. 157

[Fletcher 2004] P.T. Fletcher et 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. 48

[Förstner 1987] W. Förstner et E. Glüch. *A fast operator for detection and precise location of distinct points, corners and centers of circular features*. In Proc. of ISPRS'87, pages 281–305. Interlaken, June 1987. 91

[Förstner 1999] W. Förstner et B. Moonen. *A Metric for Covariance Matrices*. In F. Krumm et V. S. Schwarze, editors, Quo vadis geodesia...? Festschrift for

Erik W. Grafarend on the occasion of his 60th birthday, number 1999.6 of Tech. Report of the Dept of Geodesy and Geoinformatics, pages 113–128. Stuttgart University, 1999. 48, 50

[Fujikoshi 1979] Y. Fujikoshi et L.G. Veitch. *Estimation of dimensionality in canonical correlation analysis.* Biometrika, vol. 66, no. 2, pages 345–351., 1979. 188

[Gallot 1993] S. Gallot, D. Hulin et J. Lafontaine. Riemannian geometry. Springer, 1993. 55, 76

[Gamkrelidze 1991] R.V. Gamkrelidze, editor. Geometry I, volume 28 of *Encyclopaedia of Mathematical Sciences.* Springer, 1991. 48, 51

[Gee 1998] J.C. Gee et R.K.Bajcsy. Brain warping, chapter Elastic matching: continuum mechanical and probabilistic analysis. Academic Press, 1998. 157

[Gerig 1992] G. Gerig, R. Kikinis, O. Kübler et F.A. Jolesz. *Nonlinear Anisotropic Filtering of MRI Data.* IEEE Transactions on Medical Imaging, vol. 11, no. 2, pages 221–232, June 1992. 78, 102

[Geschwind 2002] D. H. Geschwind, B. L. Miller, C. DeCarli et D. Carmelli. *Heritability of lobar brain volumes in twins supports genetic models of cerebral laterality and handedness.* PNAS, vol. 99, pages 3176–3181, 2002. 157

[Glaunès 2005] Joan Alexis Glaunès. *Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique.* PhD thesis, Université Paris 13, 2005. 11, 30

[Goodlett 2006] C. Goodlett, B. Davis, R. Jean, J. Gilmore et G. Gerig. *Improved Correspondence for DTI Population Studies via Unbiased Atlas Building.* In Proc. of MICCAI'06, volume 4191 of *LNCS*, pages 260–267, 2006. 154

[Gray 1958] Henry Gray. Gray's anatomy: Descriptive and surgical. 1958. 2, 6, 20, 24

[Guéziec 1994] A. Guéziec et N. Ayache. *Smoothing and Matching of 3-D Space Curves.* The International Journal of Computer Vision, vol. 12, no. 1, pages 79–104, January 1994. 165

[Guimond 2000] A. Guimond, J. Meunier et J.-P. Thirion. *Average Brain Models: A Convergence Study.* Computer Vision and Image Understanding, vol. 77, no. 2, pages 192–210, 2000. 166

[Hagmann 2000] P. Hagmann, J. Thirian, P. Vandergheynst, S. Clarke et R. Meuli. *Statistical Fiber Tracking on DT-RI data as a Potential Tool for Morphological Brain Studies.* In Proc. of ISMRM Workshop on Diffusion MRI: Biophysical Issues, 2000. 140

[Haselgrove 1996] J. C. Haselgrove et J. R. Moore. *Correction for distortion of echo-planar images used to calculate the apparent diffusion coefficient.* Magnetic Resonance in Medicine, vol. 36, no. 6, pages 960–964, December 1996. 140

[Helgason 1978] S. Helgason. Differential geometry, Lie groups, and symmetric spaces. Academic Press, 1978. 48, 51

[Herve 2005] P.Y. Herve, B. Mazoyer, F. Crivello, G. Perchey et N. Tzourio-Mazoyer. *Finger tapping, handedness and grey matter amount in the Rolando's genu area.* NeuroImage, vol. 25, no. 4, pages 1133–45, 2005. 190

[Holder 2000] C. A. Holder, R. Muthupillai, S. Mukundan, J. D. Eastwood et P. A. Hudgins. *Diffusion-weighted MR Imaging of the Normal Human Spinal Cord in Vivo.* American Journal of Neuroradiology, vol. 21, pages 1799–1806, March 2000. 138

[Ibanez 2003] L. Ibanez, W. Schroeder, L. Ng et J. Cates. *The ITK Software Guide.* Kitware, Inc. ISBN 1-930934-10-6, http://www.itk.org/ItkSoftwareGuide.pdf, first édition, 2003. 126, 134

[Kendall 1990] W.S. Kendall. *Probability, convexity, and harmonic maps with small image I: uniqueness and fine existence.* Proc. London Math. Soc., vol. 61, no. 2, pages 371–406, 1990. 52, 60

[Kindlmann 2004] G. Kindlmann. *Superquadric Tensor Glyphs.* In Proc. of IEEE TCVG Symposium on Visualization, pages 147–154, May 2004. 132

[Kindlmann 2007] G. Kindlmann, R.S.J. Estépar, M. Niethammer, S. Haker et C.-F. Westin. *Geodesic-Loxodromes for Diffusion Tensor Interpolation and Difference Measurement.* In Proc. of MICCAI'07, volume 4791 of *LNCS*, pages 1–9, 2007. 88

[Kobayashi 1969] S. Kobayashi et K. Nomizu. Foundations of differential geometry, vol. II. Number 15 of Interscience tracts in pure and applied mathematics. John Whiley & Sons, 1969. 48, 51

[Kochunov 2002] P. Kochunov, J. Lancaster, P.M. Thompson, A.W. Toga, P. Brewer, J. Hardies et P. Fox. *An optimized individual brain target in the Talairach coordinate system.* NeuroImage, vol. 17, no. 2, pages 922–927, October 2002. 166

[Konukoglu 2007] E. Konukoglu, M. Sermesant, O. Clatz, J.-M. Peyrat, H. Delingette et N. Ayache. *A Recursive Anisotropic Fast Marching Approach to Reaction Diffusion Equation: Application to Tumor Growth Modeling.* In Proceedings of the 20th International Conference on Information Processing in Medical Imaging (IPMI'07), volume 4584 of *LNCS*, pages 686–699, 2-6 July 2007. 154

[Le Bihan 2001] D. Le Bihan, J.-F. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko et H. Chabriat. *Diffusion tensor imaging: Concepts and applications.* Journal of Magnetic Resonance Imaging, vol. 13, no. 4, pages 534–546, 2001. 43

[Le Goualher 1997] G. Le Goualher, C. Barillot et Y. Bizais. *Three-dimensional segmentation and representation of cortical sulci using active ribbons.* Int. J. of Pattern Recognition and Artificial Intelligence, vol. 11, no. 8, pages 1295–1315, 1997. 11, 30

[Le Goualher 1999] G. Le Goualher, E. Procyk, L. Collins, R. Venegopal, C. Barillot et A. Evans. *Automated extraction and variability analysis of sulcal neuroanatomy.* IEEE Transactions on Medical Imaging, vol. 18, no. 3, pages 206–217, 1999. 11, 30

[LeBihan 1991] D. LeBihan. *Molecular diffusion nuclear magnetic resonance imaging.* Magm. Reson. Q., vol. 7, pages 1–30, 1991. 138

[Lenglet 2004] C. Lenglet, R. Deriche et O. Faugeras. *Inferring white matter geometry from diffusion tensor MRI: Application to connectivity mapping.* In Proc. of the 8th European Conference on Computer Vision, 2004. 140

[Lenglet 2006a] C. Lenglet. *Geometric and Variational Methods for DTI Processing.* PhD thesis, University of Nice-Sophia Antipolis, 2006. 63

[Lenglet 2006b] C. Lenglet, M. Rousson et R. Deriche. *Statistics on the Manifold of Multivariate Normal Distributions: Theory and Application to Diffusion Tensor MRI Processing.* Journal of Mathematical Imaging and Vision, vol. 25, no. 3, pages 423–444, October 2006. 48

[MacDonald 1998] D. MacDonald, D. Avis et A.C. Evans. *Proximity Constraints in Deformable Models for Cortical Surface Identification.* In Proc. of MICCAI'98, pages 650–659, 1998. 160

[Maddah 2007] M. Maddah, W. M. Wells III, S. K. Warfield, C.-F. Westin et W. E. L. Grimson. *Probabilitic Clustering and Quantitative Analysis of White Matter Fiber Tracts.* In Proc. of IPMI'07, volume 4584 of *LNCS*, pages 372–383, 2007. 154

[Mangin 2002] J.-F. Mangin, C. Poupon, Y. Cointepas, D. Rivière, D. Papadopoulos-Orfanos, C. A. Clark, J. Régis et D. Le Bihan. *A Framework based on Spin Glass Models for the Inference of Anatomical Connectivity from Diffusion-Weighted MR Data - A Technical Review.* NRM in Biomedicine, vol. 15, pages 481–492, 2002. 140

[Mangin 2004a] J.-F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, D.L. Collins, A.C. Evans et J. Regis. *Object-based*

morphometry of the cerebral cortex. IEEE Transactions on Medical Imaging, vol. 23, no. 8, pages 968–982, 2004. 157

[Mangin 2004b] J.-F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, P. Scifo, T. Ochiai, F. Brunelle et J. Regis. *A framework to study the cortical folding patterns.* NeuroImage, vol. 23, no. Supplement 1, pages S129–S138, 2004. 11, 30, 160

[Mechelli 2005] A. Mechelli, K. J. Friston, R. S. Frackowiak et C. J. Price. *Structural Covariance in the Human Cortex.* Journal of Neurosience, vol. 25, pages 8303–8310, September 2005. 185

[Meijering 2002] E. Meijering. *A Chronology of Interpolation: From Ancient Astronomy to Modern Signal and Image Processing.* Proc. of the IEEE, vol. 90, no. 3, pages 319–342, March 2002. 64

[Moakher 2002] Maher Moakher. *Means and Averaging in the Group of Rotations.* SIAM J. Matrix Anal. Appl., vol. 24, no. 1, pages 1–16, 2002. 55

[Nomizu 1954] K. Nomizu. *Invariant affine connections on homogeneous spaces.* American J. of Math., vol. 76, pages 33–65, 1954. 48

[O'Donnel 2007] L.J. O'Donnel, C.-F. Estin et A.J. Golby. *Tract-Based Morphometry.* In Proc. of MICCAI'07, volume 4792 of *LNCS*, pages 161–168, 2007. 154

[O'Donnell 2002] L. O'Donnell, S. Haker et C.-F. Westin. *New Approaches to Estimation of White Matter Connectivity in Diffusion Tensor MRI: Elliptic PDEs and Geodesics in a Tensor-Warped Space.* In Proc. of MICCAI'02, pages 459–466, 2002. 140

[O'Donnell 2006] L. O'Donnell et C.-F. Westin. *High-Dimensional White Mapper Atlas Generation and Group Analysis.* In Proc. of MICCAI'06, volume 4191 of *LNCS*, pages 243–251, 2006. 154

[Ozanne 2007] A. Ozanne, T. Krings, D. Facon, P. Fillard, J.L. Dumas, H. Alvarez, D. Ducreux et P. Lasjaunias. *MR Diffusion Tensor Imaging and Fiber Tracking in Spinal Cord Arteriovenous Malformations: A Preliminary Study.* American Journal of Neuroradiology, vol. 28, pages 1271–1279, August 2007. 148

[Paulsen 2003] R. R. Paulsen et K. B. Hilger. *Shape Modelling Using Markov Random Field Restoration of Point Correspondences.* In Information Processing in Medical Imaging, IPMI, 2003. 164

[Pennec 1996] Xavier Pennec. *L'incertitude dans les problèmes de reconnaissance et de recalage – Applications en imagerie médicale et biologie moléculaire.* Thèse de sciences (phd thesis), Ecole Polytechnique, Palaiseau (France), December 1996. 47, 53, 55

[Pennec 1997] Xavier Pennec et Jean-Philippe Thirion. *A Framework for Uncertainty and Validation of 3D Registration Methods based on Points and Frames.* Int. Journal of Computer Vision, vol. 25, no. 3, pages 203–229, December 1997. 53

[Pennec 1998] Xavier Pennec et Nicholas Ayache. *Uniform distribution, distance and expectation problems for geometric features processing.* Journal of Mathematical Imaging and Vision, vol. 9, no. 1, pages 49–67, July 1998. A preliminary version appeared as INRIA Research Report 2820, March 1996. 50, 53

[Pennec 1999] Xavier Pennec. *Probabilities and Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements.* In A.E. Cetin, L. Akarun, A. Ertuzun, M.N. Gurcan et Y. Yardimci, editors, Proc. of Nonlinear Signal and Image Processing (NSIP'99), volume 1, pages 194–198, June 20-23, Antalya, Turkey, 1999. IEEE-EURASIP. 44, 47

[Pennec 2004a] Xavier Pennec. *Probabilities and Statistics on Riemannian Manifolds: A Geometric approach.* Research Report 5093, INRIA, January 2004. An extended version will appear in the Int. Journal of Mathematical Imaging and Vision. 44, 47, 52, 55, 61, 63, 84

[Pennec 2004b] Xavier Pennec, Pierre Fillard et Nicholas Ayache. *A Riemannian Framework for Tensor Computing.* Research Report 5255, INRIA, July 2004. Published in Int. Journal of Computer Vision 65(1), October 2005. 15, 34

[Pennec 2006] Xavier Pennec, Pierre Fillard et Nicholas Ayache. *A Riemannian Framework for Tensor Computing.* International Journal of Computer Vision, vol. 66, no. 1, pages 41–66, January 2006. A preliminary version appeared as INRIA Research Report 5255, July 2004. 15, 34, 48

[Perona 1990] P. Perona et J. Malik. *Scale-space and edge detection using anisotropic diffusion.* IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol. 12, no. 7, pages 629–639, 1990. 78, 102

[Perrin 2005] M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivière, A. Constantinesco, D. Le Bihan et J.-F. Mangin. *Fiber Tracking in q-Ball Fields using Regularized Particle Trajectories.* In Gary E. Christensen et Milan Sonka, editors, Proc. of IPMI'05, volume 1 of *LNCS 3565*, pages 52–63. Springer, July 2005. 153

[Pieper 2004] S.D. Pieper, M. Halle et R. Kikinis. *3D Slicer.* In Proc. of ISBI'04, 2004. 125

[Pitiot 2003] A. Pitiot, H. Delingette, A. Toga et P. M. Thompson. *Learning Object Correspondences with the Observed Transport Shape Measure.* In Chris Taylor et J. A. Noble, editors, Information Processing in Medical Imaging IPMI'03, volume 2732 of *LNCS*, pages 25–37. Springer Verlag, July 2003. 159, 165

[Pitiot 2004] A. Pitiot, H. Delingette, P.M. Thompson et N. Ayache. *Expert Knowledge Guided Segmentation System for Brain MRI.* NeuroImage, vol. 23, no. supplement 1, pages S85–S96, 2004. Special Issue: Mathematics in Brain Imaging. 157

[Poupon 1999] Cyril Poupon. *Détection des faisceaux de fibres de la substance blanche pour l'étude de la connectivité anatomique cérébrale.* Thèse de sciences, Ecole Nationale Supérieure de Télécommunications, 1999. 8, 26

[Poupon 2000] C. Poupon, C. A. Clark, V. Frouin, J. Régis, I. Bloch, D. Le Bihan et J.-F. Mangin. *Regularization of Diffusion-Based Direction Maps for the Tracking of Brain White Matter Fascicles.* NeuroImage, vol. 12, no. 2, pages 184–195, August 2000. 140

[Rao 1991] A. R. Rao et B. G. Schunck. *Computing oriented texture fields.* CVGIP: Graphical Models and Image Processing, vol. 53, pages 157–185, 1991. 91

[Rao 2006] A Rao, K Babalola et Daniel Rueckert. *Canonical correlation analysis of sub-cortical brain structures using non-rigid registration.* In Proc. of WBIR, July 2006. 187

[Renoux 2006] J. Renoux, D. Facon, P. Fillard, I. Huynh, P. Lasjaunias et D. Ducreux. *MR Diffusion Tensor Imaging and Fiber Tracking in Inflammatory Diseases of the Spinal Cord.* American Journal of Neuroradiology, vol. 27, pages 1947–1951, October 2006. 138, 146

[Rey 2002] David Rey, Gérard Subsol, Hervé Delingette et Nicholas Ayache. *Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis.* Medical Image Analysis, vol. 6, no. 2, pages 163–179, June 2002. 73

[Ries 2000] M. Ries, R. A. Jones, V. Dousset et C.T.W. Moonen. *Diffusion tensor MRI of the spinal cord.* Magnetic Resonance in Medicine, vol. 44, no. 6, pages 884–892, December 2000. 138

[Rivière 2002] D. Rivière, J.-F. Mangin, D. Papadopoulos-Orfanos, J.-M. Martinez, V. Frouin et J. Régis. *Automatic recognition of cortical sulci of the Human Brain using a congregation of neural networks.* Medical Image Analysis, vol. 6, no. 2, pages 77–92, 2002. 11, 30

[Roche 2000] A. Roche, G. Malandain et N. Ayache. *Unifying Maximum Likelihood Approaches in Medical Image Registration.* International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging, vol. 11, no. 1, pages 71–80, 2000. 158

[Salavador 2005] R. Salavador, A. Peña, D.K. Menon, T.A. Carpenter, J.D. Pickard et E.T. Bullmore. *Formal Characterization and Extension of the Linearized*

Diffusion Tensor Model. Human Brain Mapping, vol. 24, pages 144–155, 2005. 100

[Savadjiev 2006] P. Savadjiev, J.S.W. Campbell, G.B. Pike et K. Siddiqi. *3D Curve Inference for Diffusion MRI Regularization and Fibre Tractography.* Medical Image Analysis, vol. 10, no. 5, pages 799–813, 2006. 153

[Sibson 1981] R. Sibson. *A brief description of natural neighbour interpolation.* In Vic Barnet, editor, Interpreting Multivariate Data, pages 21–36. John Wiley & Sons, Chichester, 1981. 65, 67

[Sijbers 1998] J. Sijbers, A.J. den Dekker, P. Scheunders et D. Van Dyck. *Maximum Likelihood Estimation of Rician Distribution Parameters.* IEEE Transactions on Medical Imaging, vol. 17, no. 3, June 1998. 101, 108

[Skare 2000] S. Skare, T. Li, B. Nordell et Ingvar M. *Noise considerations in the determination of diffusion tensor anisotropy.* MRI, vol. 18, no. 6, pages 659–669, July 2000. 114

[Skovgaard 1984] L.T. Skovgaard. *A Riemannian geometry of the multivariate normal model.* Scand. J. Statistics, vol. 11, pages 211–223, 1984. 48, 52, 60

[Sowell 2004] E.R. Sowell, P.M. Thompson et A.W. Toga. *Mapping Changes in the Human Cortex throughout the Span of Life.* The Neuroscientist, vol. 10, no. 4, pages 372–92, August 2004. 160

[Stefanescu 2005] R. Stefanescu. *Parallel nonlinear registration of medical images with a priori information on anatomy and pathology.* PhD thesis, Nice-Sophia Antipolis University, 2005. 91

[Stein 1992] B. M. Stein et P. C. McCormick. *Intramedullary neoplasms and vascular malformations.* Clin. Neurosurg., vol. 39, pages 361–387, 1992. 148

[Stejskal 1965] E.O. Stejskal et J.E. Tanner. *Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient.* Journal of Chemical Physics, vol. 42, pages 288–292, 1965. 5, 24

[Studholme 2007] C. Studholme. *Incorporating DTI Data as a Constraint in Deformation Tensor Morphometry between T1 MR Images.* In Proc. of MICCAI'07, volume 4584 of *LNCS*, pages 223–232, 2007. 154

[Talairach 1988] J. Talairach et P. Tournoux. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system - an approach to cerebral imaging. Thieme Medical Publishers, 1988. 11, 30

[Thévenaz 2000] P. Thévenaz, T. Blu et M. Unser. *Interpolation Revisited.* IEEE Transactions on Medical Imaging, vol. 19, no. 7, pages 739–758, July 2000. 64

[Thirion 1998] J.-P. Thirion. *Image matching as a diffusion process: an analogy with Maxwell's demons.* Medical Image Analysis, vol. 2, no. 3, pages 243–260, 1998. 134

[Thompson 2000] P.M. Thompson, M.S. Mega, K.L. Narr, E.R. Sowell, R.E. Blanton et A.W. Toga. *Brain Image Analysis and Atlas Construction.* In M. Fitzpatrick et M. Sonka, editors, Handbook of Medical Image Proc. and Analysis, chapter 17. SPIE, 2000. 159, 172, 175

[Thompson 2001a] P.M. Thompson, T.D. Cannon, K.L. Narr, T.G.M. van Erp, V.P. Poutanen, M. Huttunen, J. Lönnqvist, C.G. Standertskjï¿½d-Nordenstam, J. Kaprio, M. Khaledy, R. Dail, C.I. Zoumalan et A.W. Toga. *Genetic Influences on Brain Structure.* Nature Neuroscience, vol. 4, no. 12, pages 1253–1258, December 2001. 157

[Thompson 2001b] P.M. Thompson, M.S. Mega, R.P. Woods, C.I. Zoumalan, C.J. Lindshield, R.E. Blanton, J. Moussai, C.J. Holmes, J.L. Cummings et A.W. Toga. *Cortical Change in Alzheimer's Disease Detected with a Disease-Specific Population-Based Brain Atlas.* Cerebral Cortex, vol. 11, no. 1, pages 1–16, January 2001. 159

[Thompson 2004] P.M. Thompson, K.M. Hayashi, E.R. Sowell, N. Gogtay, J.N. Giedd, J.L. Rapoport, G.I. de Zubicaray, A.L. Janke, S.E. Rose, J. Semple, D.M. Doddrell, Y.L. Wang, T.G.M. van Erp, T.D. Cannon et A.W. Toga. *Mapping Cortical Change in Alzheimer's Disease, Brain Development, and Schizophrenia.* NeuroImage Special Issue on Mathematics in Brain Imaging, September 2004. 160

[Thompson 2005] P.M. Thompson, A.D. Lee, R.A. Dutton, J.A. Geaga, K.M. Hayashi, M.A. Eckert, U. Bellugi, A.M. Galaburda, J.R. Korenberg, D.L. Mills, A.W. Toga et A.L. Reiss. *Abnormal Cortical Complexity and Thickness Profiles Mapped in Williams Syndrome.* Journal of Neuroscience, vol. 25, no. 16, pages 4146–4158, April 2005. 183

[Toga 2003] A.W. Toga et P.M. Thompson. *Mapping Brain Asymmetry.* Nature Reviews Neuroscience, vol. 4, no. 1, pages 37–48, January 2003. 157, 172, 180

[Tosun 2005] D. Tosun et J. L. Prince. *Cortical Surface Alignment using Geometry Driven Multispectral Optical Flow.* In Proc. of IPMI'05, volume 3565 of *LNCS*, pages 480–492, 2005. 159

[Tournier 2004a] J.-D. Tournier, F. Calamante, D. G. Gadian et A. Connelly. *Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution.* NeuroImage, vol. 23, no. 3, pages 1176–1185, 2004. 7, 25

[Tournier 2004b] J.-D. Tournier, F. Calamante, D.G. Gadian et A. Connelly. *Direct estimation of the fiber orientation density function from diffusion-weightedMRI data using spherical deconvolution.* NeuroImage, vol. 23, pages 1176–1185, 2004. 97

[Toussaint 2007a] N. Toussaint, M. Sermesant et P. Fillard. *vtkINRIA3D: A VTK Extension for Spatiotemporal Data Synchronization, Visualization and Management.* In Proc. of Workshop on Open Source and Open Data for MICCAI, Brisbane, Australia, October 2007. 126

[Toussaint 2007b] N. Toussaint, J.C. Souplet et P. Fillard. *MedINRIA: Medical Image Navigation and Research Tool by INRIA.* In Proc. of MICCAI'07 Workshop on Interaction in medical image analysis and visualization, Brisbane, Australia, 2007. 15, 35

[Trouve 2000] A. Trouve et L. Younes. *Diffeomorphic Matching Problems in One Dimension: Designing and Minimizing Matching Functionals.* In ECCV00, pages I: 573–587, 2000. 164

[Tschumperlé 2002] D. Tschumperlé et R. Deriche. *Orthonormal vector sets regularization with PDE's and applications.* Int. J. of Computer Vision (IJCV), vol. 50, no. 3, pages 237–252, 2002. 43, 65

[Tschumperlé 2003] D. Tschumperlé et R. Deriche. *Variational Frameworks for DT-MRI Estimation, Regularization and Visualization.* In Proc. of ICCV, pages 116–121. IEEE Computer Society, 2003. 100

[Tuch 2002] D. S. Tuch. *High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity.* Magn. Reson. Med., vol. 48, pages 577–582, 2002. 7, 25

[Tuch 2004] D.S. Tuch. *Q-Ball Imaging.* Magnetic Resonance in Medicine, vol. 52, pages 1358–1372, 2004. 7, 25, 97

[Vaillant 2005] M. Vaillant et J. Glaunès. *Surface matching via currents.* In Proc. of IPMI'05, pages 381–392, 2005. 11, 30, 159

[Vaillant 2006] M. Vaillant, A. Qiu, J. Glaunès et M. I. Miller. *Diffeomorphic metric surface mapping in subregion of the superior temporal gyrus.* NeuroImage, vol. 34, no. 3, pages 1149–1159, 2006. 159

[Vercauteren 2007] Tom Vercauteren, Xavier Pennec, Aymeric Perchant et Nicholas Ayache. *Non-parametric Diffeomorphic Image Registration with the Demons Algorithm.* In Nicholas Ayache, Sébastien Ourselin et Anthony J. Maeder, editors, Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI'07), volume 4792 of *Lecture Notes in Computer Science*, pages 319–326, Brisbane, Australia, October 2007. Springer-Verlag. PMID: 18044584. 134

[VTK 2007] *The Visualization ToolKit: http://www.vtk.org/*, 2007. 126

[Wang 2004a] Z. Wang, B.C. Vemuri, Y. Chen et T.H. Mareci. *A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI.* IEEE Transactions on Medical Imaging, vol. 23, no. 8, August 2004. 44, 101, 102, 103, 110

[Wang 2004b] Zhizhou Wang et Baba C. Vemuri. *An Affine Invariant Tensor Dissimilarity Measure and Its Applications to Tensor-Valued Image Segmentation.* In Proc. of CVPR'04, pages 228–233, 2004. 44

[Wang 2005] Y.L. Wang, M.C. Chiang et P.M. Thompson. *Automated Surface Matching using Mutual Information Applied to Riemann Surface Structures.* In Proc. of MICCAI'05, number 2 of lncs, pages 666–674, Palm Springs, CA, October 2005. 159

[Weickert 1999] J. Weickert. *Coherence-enhancing diffusion filtering.* IJCV, vol. 31, no. 2/3, pages 111–127, April 1999. 91

[Weinstein 1999] David M. Weinstein, Gordon L. Kindlmann et Eric C. Lundberg. *Tensorlines: Advection-Diffusion based Propagation through Diffusion Tensor Fields.* vis, vol. 00, page 40, 1999. 140

[Westin 2002] C.-F. Westin, S. Maier, H. Mamata, A. Nabavi, F. Jolesz et R. Kikinis. *Processing and visualization for diffusion tensor MRI.* Medical Image Analysis, vol. 6, no. 2, pages 93–108, 2002. 43, 100, 107, 140

[Wheeler-Kingshott 2002] C. Wheeler-Kingshott, S. Hickman et G. Parker. *Investigating cervical spinal cord structure using diffusion tensor imaging.* NeuroImage, vol. 16, pages 93–102, 2002. 138

[wxWidgets 2007] wxWidgets. *http://www.wxwidgets.org/*, 2007. 126

[Xu 2002] D. Xu, S. Mori, M. Solaiyappan, P. C. M. van Zijl et C. Davatzikos. *A Framework for Callosal Fiber Distribution Analysis.* NeuroImage, vol. 17, no. 3, pages 1131–1143, November 2002. 140, 142

[Zhang 2007a] F. Zhang, C. Goodlett, E. Hancock et G. Gerig. *Probabilistic Fiber Tracking Using Particle Filtering.* In Proc. of MICCAI'07, volume 4792 of *LNCS*, pages 144–152, 2007. 140

[Zhang 2007b] H. Zhang, P. A. Yushkevich, D. Rueckert et J. C. Gee. *Unbiased White Matter Atlas Construction using Diffusion Tensor Images.* In Proc. of MICCAI'07, volume 4792 of *LNCS*, pages 211–218, 2007. 154

[Ziyan 2007] U. Ziyan, M. R. Sabuncu, L. J. O'Donnell et C.-F. Westin. *Nonlinear Registration of Diffusion MR Images Based on Fiber Bundles.* In Proc. of MICCAI'07, volume 4791 of *LNCS*, pages 351–358, 2007. 154