UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE MSII no 269

Laboratoire ICube

THÈSE présentée par :

Theodosios GKAMAS

soutenue le : **29 septembre 2015**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**
Discipline/ Spécialité : **Traitement du signal et des images**

MODÉLISATION STATISTIQUE DE TENSEURS D'ORDRE SUPÉRIEUR EN IMAGERIE PAR RÉSONANCE MAGNÉTIQUE DE DIFFUSION

~

STATISTICAL MODELLING OF HIGH ORDER TENSORS IN DIFFUSION WEIGHTED MAGNETIC RESONANCE IMAGING

THÈSE dirigée par :

M. Christian HEINRICH	Professeur, université de Strasbourg
M. Stéphane KREMER	Professeur, université de Strasbourg

RAPPORTEURS :

Mme Florence FORBES	Directrice de recherche, INRIA Grenoble Rhône-Alpes
M. Jean-Philippe RANJEVA	Professeur, université d'Aix-Marseille

AUTRES MEMBRES DU JURY :

M. Jean-Michel DISCHLER	Professeur, université de Strasbourg
M. Christophoros NIKOU	Professeur Associé, université de Ioannina
M. Félix RENARD	Docteur, AGIM Lab, université de Grenoble

Acknowledgements

Many people may claim that following the path of research is a tough and durable task to do, like a marathon, independently the scientific field, especially in the first steps as researcher when the experience is limited. Moreover, the path from the starting point to the achievement of the goals is not always a short straight line, thus a researcher should be equipped with patience, tolerance, motivation, curiosity and thirst for knowledge.

My first early steps as a researcher were taken in the last year of my diploma in 2008, and thankfully continued in my master degree for the next two years in 2008-2010, under the supervision of the same person, Christophoros Nikou, Associate Professor at the Department of Computer Science and Engineering at the University of Ioannina in Greece, who sparked my interest in research. In addition, his assistance to find a Ph.D. thesis in France was very important and I am very grateful to him.

Supplementary to the previously mentioned virtues that a researcher should possess, it is mandatory that the working environment is intimate and friendly, consisting of people who are willing to collaborate and offer to young researchers the opportunity to exchange their opinion with, and moreover to enlighten them with their experience and knowledge. Thinking in this direction, I would like to thank my supervisors Christian Heinrich and Stéphane Kremer, who offered me the opportunity to fulfil one of my dreams, to enrich my knowledge in computer science, through a financial doctoral contract related to a proposed Ph.D. thesis in topics directly applicable to medical science. Furthermore, their flawless scientific knowledge and professional expertise assisted crucially to understand deeply the new research topics and guided me skillfully to find the answers to our scientific problems. Likewise, I owe a great thanks to Félix Renard, currently post-doctoral researcher in AGIM laboratory (formerly known as Gipsa lab) at the University of Grenoble, for his honor, availability and experience in DW-MRI modelling, processing and analysis, providing me with invaluable information in these topics.

Additionally, I would like to express my gratitude to every member of the jury, and particularly to Florence Forbes and Jean-Philippe Ranjeva who accepted the invitation to review this thesis.

To continue, the accomplishment of my thesis would not be possible without the establishment of a collaboration with the ICube-IMIS research team in order to gain access to the medical database. For that reason, I specially thank Jean-Paul Armspach, leader of ICube-IMIS team, who welcomed me in his team, and additionally to Frédérique Ostré and Julien Lamy for their technical support.

Many thanks should also be delivered to my colleagues in ICube-MIV team. Being a member of such a large well-organised research group working on similar projects it is a bless, since it can always result into fruitful conversations, opening new horizons for research. During these years, I met remarkable researchers and I made friends that I will not forget.

Taking the decision to live for at least three years in a new place, thousands of kilometers away from your homeland, it is similarly like starting a new life almost from scratch. Starting a new life in France without speaking not even a single word in French was a drawback in daily life. Although it was not an obstacle in my working environment, I thought it would be an important missing part of the French culture that I could not realise it if I was not interested in learning and assimilating this delegate language into my way of living. Thus, I voluntarily started following French courses, in Spiral and CRL language centers at the University of Strasbourg, in parallel with my Ph.D. studies, all these years, leading me in claiming with proud that I can speak and understand French in an acceptable level.

Furthermore, living in Strasbourg for almost four years up to now, new people got into my life, and therefore I cannot forget to thank my friends Angeliki, Emilie, Marios, Taka, Ting, KoHsin, Mickaël, Anne, Kathi, Hélène, Ralph and Florian for their kindness and rare personality, but also for the unlimited, joyful and memorable time that we spent together in Strasbourg. Of course, I will not forget my friends and classmates in Greece, Petros, Zaher, Vasilis K., Vasilis G., Michalis V., Michalis G., Aris, Vangelis and Kostas that honor our friendship all these years, even though the geographical distance was separating us for a long time.

Last but not least, I would like to thank every member of my family from the bottom of my heart, for their endless, selfless and unconditional feelings and support that they generously offer to me and characterize their presence from the first moments of my life.

Abstract

Diffusion weighted magnetic resonance imaging (DW-MRI) is a non-invasive modality, able to measure the diffusion of water molecules in living tissues. Fitting tensors models on DW-MRI data allows to represent the diffusion in $3D$ space at every voxel in that tissue. Second (T2) and fourth (T4) order tensors are studied extensively. This thesis focuses on the problems of statistical populations comparison and individual against a healthy group comparison, in DW-MRI.

Examining the diffusivity allows us to study the structure of complex organs. For the purposes of this study, the human brain is selected. A variety of brain pathologies alters the structure of the neural fibers in the brain, either globally (*e.g.* multiple sclerosis (MS), Alzheimer's disease (AD), neuromyelitis optica (NMO)), or in specific regions (*e.g.* locked-in syndrome (LIS)). Therefore, DW-MRI analysis is suitable to extract knowledge related to a particular disease *e.g.* through biomarker detection, disease staging and patient follow-up.

Given a healthy population (as a reference) and a group of patients (related to the same disease), capturing the variability of each group, biomarkers detection, disease staging (through different acquisitions in time) and patient follow-up can be performed *via* statistical populations comparisons. On the other hand, lacking of enough pathological data, patient follow-up can be achieved through individual statistical comparisons against the reference population. In this thesis, two methods are proposed (one for each problem, populations or individual comparisons), capable to produce fruitful maps of statistics (*i.e.* statistical atlases).

Undoubtedly, comparing different subjects presupposes that all data are normalized in a common space. In the case of orientated data (*e.g.* DW-MRI, tensor images), a single registration step will produce incoherent fiber orientations. Thus the registration should be always followed by a reorientation step. In this dissertation, reorientation methods for T4s are studied.

To continue, one of the fundamental points of the proposed statistical methods is the estimation of the reduced dense space of the tensor models using Isomap, a nonlinear dimensionality reduction technique. Once the reduced space is estimated, a flexible Gaussian mixture model is fitted to each group (or only to the reference group) and statistics are calculated robustly. Moreover, p-values are estimated with the aid of permutation testing or Monte Carlo simulations. Furthermore, in contrast to many statistical approaches found in the literature which halt their calculations in a single p-value estimation per voxel, we propose to further analyse our approximations by finding a Highest Probability Density interval for each p-value.

Applications of the proposed methods to synthetic and real cases were accomplished. The effectiveness of the proposed methods is compared favorable or even better than many state of the art approaches. In the case of real data, the NMO disease and the LIS syndrome were selected to be analysed. The obtained results are coherent with medical knowledge.

Résumé long en français

1. Introduction

Le cerveau humain est un organe multi-fonctionnel, et l'un des organes les plus importants du corps humain. Étudier la structure et la fonctionnalité du cerveau a toujours fasciné les médecins. De nos jours, une grande partie des capacités du cerveau reste inconnue.

Par ailleurs, l'étude de la spécificité de nombreuses pathologies du cerveau et l'extraction des formes et structures qui les caractérisent sont des sujets auxquels la communauté scientifique s'intéresse. Par exemple, les maladies inflammatoires peuvent se développer initialement dans des régions spécifiques du cerveau puis diffuser progressivement, tandis que d'un autre côté peuvent exister des pathologies strictement localisées dans certaines zones. En conséquence, des outils d'analyse spécifiques et efficaces qui peuvent aider au diagnostic précoce et au pronostic, mais aussi à l'extraction de connaissances liées à une maladie donnée et au suivi des patients, sont nécessaires. Ces examens médicaux devraient permettre d'extraire des caractéristiques multidimensionnelles qui ne sont pas extractibles avec les examens classiques.

L'imagerie par résonance magnétique de diffusion

Durant les dernières décennies, la physique, l'informatique et la médecine ont joint leurs forces pour numériser, modéliser et étudier le cerveau, sous le nom de Neuroscience. Des examens ont été élaborés et adaptés pour mesurer les différentes propriétés des tissus cérébraux d'une manière non-invasive, *in vivo*, afin de caractériser correctement de nombreuses pathologies. Par exemple, l'Imagerie par Résonance Magnétique de diffusion (IRMd) est une variation de l'IRM classique et permet de suivre le mouvement des molécules d'eau en $3D$. L'IRMd renseigne ainsi sur l'anatomie structurelle des connexions neuronales dans le cerveau.

De plus, la neuroinformatique, une science dont certains objectifs sont de traiter des images médicales, de modéliser les données acquises et de les analyser, a sa propre place dans ce contexte. Il est courant de modéliser des données IRMd en utilisant des tenseurs (pour plus d'informations, le lecteur est renvoyé au chapitre 2), afin de visualiser et d'analyser des données brutes d'IRMd. Les tenseurs d'ordre deux (T2s) sont largement utilisés et bien connus, mais leurs capacités sont limitées à la description de fibres simples. Pour cette étude, nous avons choisi de travailler avec les tenseurs d'ordre quatre (T4s), car ils peuvent représenter beaucoup plus en détail la structure sous-jacente de la fibre que les T2s, en particulier dans le cas du croisement de fibres.

Applications médicales de cette thèse

L'analyse d'une pathologie donnée, avec pour objectif ultime l'extraction de biomarqueurs, exige une très grande base de données de sujets traités, afin de disposer de populations représentatives (une pour le groupe témoin et l'autre pour le groupe pathologique), capables de capturer la variabilité de la maladie. Alternativement, si notre objectif principal est de suivre l'état du patient avec de multiples analyses dans le temps, il n'est pas obligatoire de construire deux populations, puisque chaque sujet pathologique sera examiné individuellement au regard de la population saine.

Dans cette thèse, les comparaisons de populations ainsi que les comparaisons d'une personne considérée individuellement avec une population normale sont traitées par des tests statistiques spécifiques originaux, dans un but de diagnostic précoce, de pronostic, de détection de biomarqueurs, d'évaluation du stade de la maladie et de suivi des patients.

Lorsqu'un patient arrive en vue d'un diagnostic, on utilise les biomarqueurs extraits via la comparaison de populations pour analyser les données du patient ou l'on réalise une analyse statistique en comparant ce patient au groupe de référence. Le diagnostic est réalisé en fonction des résultats obtenus.

Enfin, les applications des méthodes proposées à la neuromyélite optique aiguë (*neuromyelitis optica*, NMO), ou maladie de Devic et à la maladie *locked-in syndrome* (LIS) sont présentées.

Contributions de la thèse

Le traitement d'un grand nombre de sujets pour finalement construire un atlas statistique lié à une maladie, par exemple par comparaison de populations, exige une étape de prétraitement cruciale, appelée normalisation des données, de sorte que les sujets seront correctement recalés spatialement les uns avec les autres, par exemple en utilisant un modèle de référence commun. Malheureusement, dans le cas de données IRMd (ou d'images de tenseurs) une simple étape de recalage ne suffit pas, car les données contiennent des informations orientées. En conséquence, une étape de réorientation doit compléter la tâche de normalisation. Pendant la première année de cette thèse, nous avons traité le problème de la réorientation des tenseurs d'ordre quatre. Notre étude sur la réorientation de T4 est présentée dans le chapitre 3.

Afin de comparer les sujets dans l'espace des tenseurs, des métriques efficaces qui considèrent l'ensemble des informations comprises dans le modèle de tenseur doivent être choisies. Malheureusement, la majorité des métriques de tenseur de la littérature sont définies uniquement pour les modèles T2s, alors que nous nous intéressons essentiellement aux tenseurs d'ordre supérieur. Suite à l'étude des métriques de la littérature, nous avons proposé une nouvelle métrique de tenseur, adaptée à tout modèle de tenseur (un état de l'art des différentes métriques de tenseur, ainsi que la métrique proposée, sont présentés dans le chapitre 2).

Au cours des deux années suivantes de préparation de la thèse, nos efforts se sont focalisés sur le développement d'approches statistiques avancées pour résoudre deux problèmes, d'une part les comparaisons de populations (le groupe normal *versus* le groupe pathologique/anormal) et d'autre part les comparaisons d'une personne anormale *versus* la population normale, avec pour objectif ultime la détection de biomarqueurs, le suivi des patients et le diagnostic précoce. Les méthodes statistiques proposées ainsi que leur application à des cas réels et synthétiques sont présentés dans les chapitres 5 à 7.

Organisation de la thèse

L'organisation de cette thèse est la suivante. Le chapitre 1 est l'introduction de cette thèse.

Le chapitre 2 est un chapitre d'introduction aux données IRMd et aux moyens de les modéliser (par exemple, à l'aide de tenseurs). En outre, un ensemble de métriques de tenseur dépendant de l'ordre du modèle est présenté, ainsi que la métrique proposée. Le chapitre se termine par une mention brève de descriptions de niveau supérieur de données IRMd, à savoir les fibres et les connectomes.

Le chapitre 3 détaille les étapes de prétraitement des données IRMd, soulignant la réorientation des T4s. Notre étude sur ce sujet avec l'évaluation expérimentale est incluse.

Le chapitre 4 présente le problème de la construction d'atlas statistiques permettant l'extraction de biomarqueurs. En outre, les méthodes de l'état de l'art sont présentées.

Les chapitres 5 et 6 contiennent la première approche statistique proposée, pour le problème de la comparaison de populations. L'application à la maladie NMO, ainsi que l'application aux tests synthétiques, sont également présentées.

Dans le chapitre 7, la deuxième approche statistique proposée, pour le problème de la comparaison d'un individu avec une population normale, est décrite. Cette approche est applicable dans le cas de populations pathologiques clairsemées. L'application à la maladie LIS est présentée.

Enfin, le chapitre 8 contient la conclusion de cette thèse ainsi que les perspectives proposées.

2. Modèles de tenseur pour les données IRMd

Un ensemble de modèles de tenseurs permettant de décrire les données IRMd existe dans la littérature. Dans le cadre de cette thèse, les tenseurs d'ordre deux et quatre sont utilisés.

Tenseurs d'ordre deux

L'imagerie du tenseur de diffusion (*diffusion tensor imaging*, DTI) a été la première tentative pour représenter des données IRMd à l'aide de T2s. Un tenseur T2 peut être défini comme une matrice symétrique 3×3 comme suit :

$$\mathbf{D}_2 = \begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{bmatrix} . \tag{1}$$

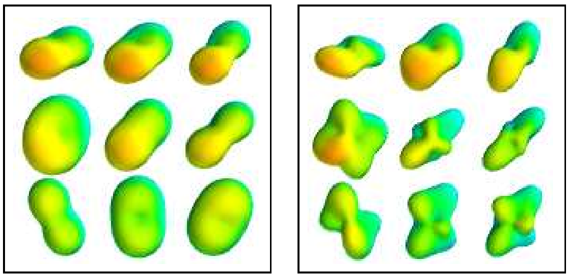


Figure 1: Modèles T2s (à gauche) et modèles T4s (à droite) dans une certaine région du cerveau. Comme prévu, le modèle T4 est plus fin que le modèle T2, et capture avec plus de détails les croisements de fibres.

La fonction de diffusivité $d(\mathbf{g})$ (pour la modélisation de tenseurs de diffusion) ou la fonction de distribution d'orientation de fibres (*fiber orientation distribution function*, fODF) $f(\mathbf{g})$ (pour la modélisation de tenseurs de fODF) (pour plus d'informations, voir le chapitre 2) liée à un T2 constitue la grandeur d'intérêt (la diffusion ou l'orientation de fibres). Selon une direction $\mathbf{g} = (g_1, g_2, g_3)$ de gradient donnée, la grandeur d'intérêt, $d(\mathbf{g})$ ou $f(\mathbf{g})$, est définie par la relation :

$$\sum_{i=1}^{3}\sum_{j=1}^{3} D_{ij}\ g_i\ g_j\ . \tag{2}$$

Tenseurs d'ordre quatre

Les modèles d'ordre supérieur à l'ordre deux existent dans la littérature. Dans ce mémoire, nous allons nous concentrer sur les modèles T4s qui sont en mesure de représenter jusqu'à trois faisceaux de fibres distincts en un seul voxel et peuvent être décrits par la matrice symétrique 6 × 6 suivante :

$$\mathbf{D}_4 = \begin{bmatrix} D_{1111} & D_{1122} & D_{1133} & D_{1112} & D_{1123} & D_{1113} \\ D_{2211} & D_{2222} & D_{2233} & D_{2212} & D_{2223} & D_{2213} \\ D_{3311} & D_{3322} & D_{3333} & D_{3312} & D_{3323} & D_{3313} \\ D_{1211} & D_{1222} & D_{1233} & D_{1212} & D_{1223} & D_{1213} \\ D_{2311} & D_{2322} & D_{2333} & D_{2312} & D_{2323} & D_{2313} \\ D_{1311} & D_{1322} & D_{1333} & D_{1312} & D_{1323} & D_{1313} \end{bmatrix} . \tag{3}$$

Par similarité au cas T2, la fonction de diffusivité (ou la fonction de fODF) d'un T4 s'écrit :

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} D_{ijkl}\ g_i\ g_j\ g_k\ g_l\ . \tag{4}$$

Un exemple visuel de la supériorité des T4s par rapport aux T2s pour produire des descriptions plus représentatives de la structure des fibres est proposé figure 1. Il est à noter que le modèle T4 (appartenant à $\mathbb{R}^{15}$, décrit par 15 coefficients uniques) est plus performant que le modèle T2 (appartenant à $\mathbb{R}^{6}$), en particulier dans le cas du croisement de fibres. Ceci provient du fait que le modèle T2 est un cas particulier de modèle T4.

Métrique de tenseur proposée

En raison de l'absence de métriques définies pour des tenseurs d'ordre supérieur et bénéficiant des propriétés souhaitées du logarithme, nous avons proposé une nouvelle metrique.

D'après le travail de Tarantola [153] sur les distances entre fonctions positives, nous proposons ci-dessous la distance entre deux profils d_1, d_2 de diffusivité (ou deux profils de fODF) qui peut être utilisée pour tous les modèles de tenseurs :

$$dist(d_1, d_2) = \iint \left| \log \frac{d_1(\theta, \phi)}{d_2(\theta, \phi)} \right| \sin\theta \ \mathtt{d}\theta \ \mathtt{d}\phi \ \ , \tag{5}$$

où $\phi \in [0, \pi]$ est l'angle polaire et $\theta \in [0, 2\pi]$ est l'angle d'azimut qui paramètrent la sphère unité en $3D$.

3. Étapes de prétraitement des données IRMd et importance de la réorientation des T4s

Étapes de prétraitement

Avant d'appliquer un test statistique approprié, un ensemble d'étapes de prétraitement est obligatoire. La première étape est la correction de courants de Foucault (*eddy current correction*) où les données sont débarrassées de tout mouvement de l'objet et de tout mouvement dû à des pulsations sanguines. Dans la deuxième étape, nous devons extraire le volume du cerveau en éliminant certaines zones, par exemple le crâne ou les yeux. Troisièmement, nous devons normaliser les données (recalage spatial et réorientation de tenseurs ou de données IRMd) dans un espace de référence commun, en calculant une transformation linéaire ou non-linéaire à l'aide de la carte d'anisotropie fractionnelle (image FA, *fractional anisotropy*) et enfin, nous devons réduire l'erreur de recalage soit par lissage, soit par l'approche par patch que nous proposons (voir la section 5.1.2 pour plus de détails).

Étude sur la réorientation des T4s

Dans la première année de cette thèse, le problème de la normalisation de tenseur a été étudié, et en particulier celui des réorientations de T4. L'importance de l'étape de réorientation après le recalage spatial des données est mise en évidence dans la figure 2, où il est évident que les modèles recalés (simplement recalés et non réorientés) ne sont pas adaptés à la nouvelle orientation de la fibre.

Les procédures de réorientation de T4 étudiées (voir le chapitre 3) sont basées sur les décompositions en T2s suivies par les réorientations des T2s. Les décompositions spectrales (*spectral decomposition*, SD) de T4s produisant six T2s, ainsi que les décompositions basées sur le théorème de Hilbert (*Hilbert decomposition*, HD) produisant trois T2s ont été testées, avec les réorientations de T2s de type *finite strain* (FS) et la préservation des directions

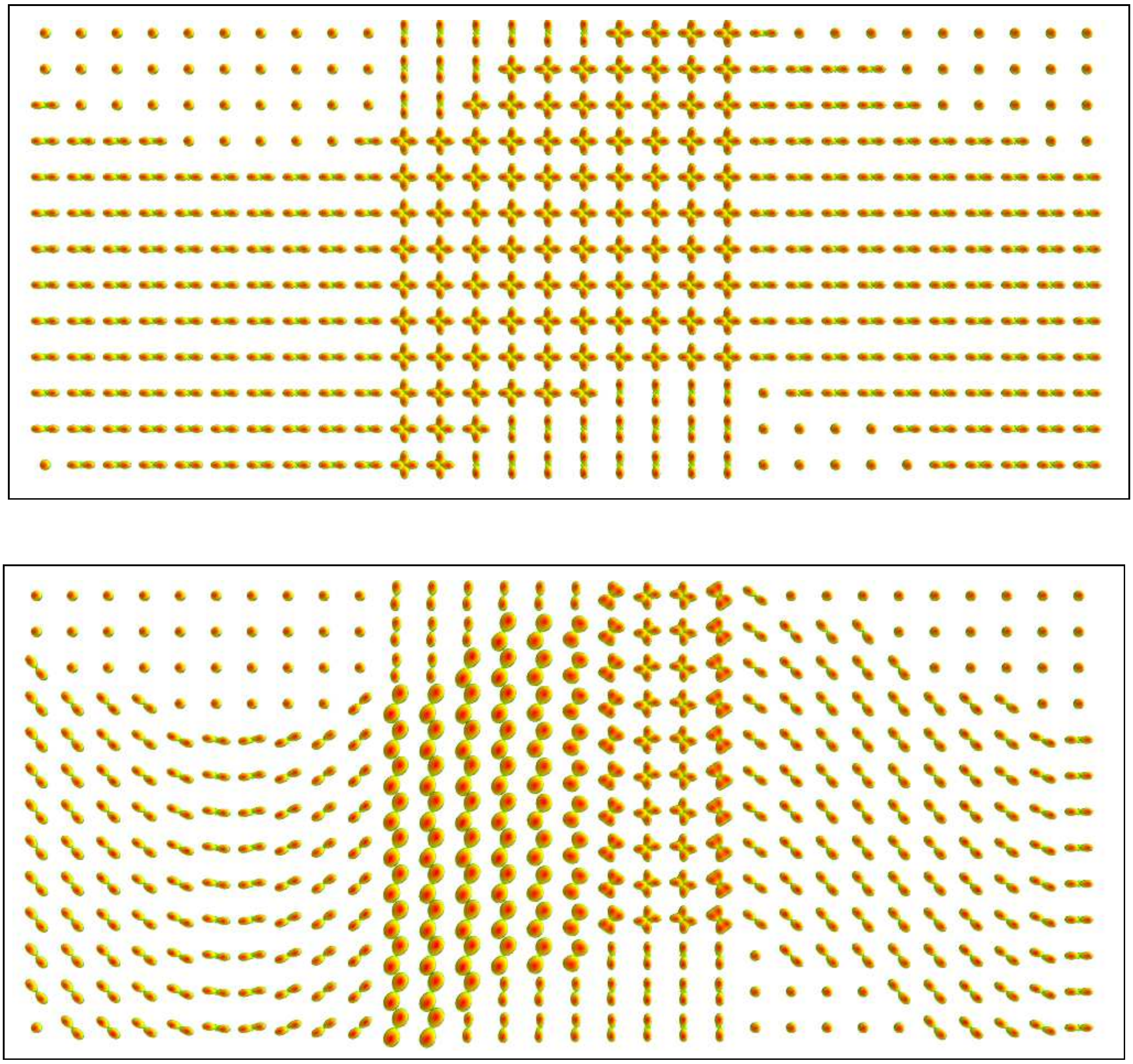


FIGURE 2: Champs synthétiques de T4 après normalisation spatiale ($[x', y']^T = [x, y + \sin(x)]^T$). Figure du haut : après recalage, sans réorientation. Figure du bas : après recalage et réorientation. On constate que la figure du haut ne rend pas compte de la structure des fibres sous-jacentes.

principales (*preservation of principal directions*, PPD). La procédure de réorientation de T4 proposée par Renard pendant sa thèse [134] est construite sur la combinaison HD avec PPD. Des résultats expérimentaux à la fois sur des données synthétiques et sur des données réelles, complétés des procédures d'évaluation pour chaque cas, figurent dans le chapitre 3.

Bien entendu, au lieu de normaliser les images de tenseurs, il serait possible de normaliser les données IRMd brutes. Cela semble être plus sûr, car une variation des coefficients du tenseur aura un plus grand impact, dans de nombreuses directions de diffusion, qu'une perturbation dans les données IRMd.

4. Introduction aux atlas statistiques pour les données IRMd

Un atlas statistique est une description de la variabilité d'une ou de plusieurs grandeurs d'intérêt au sein d'une ou de plusieurs populations. Les atlas statistiques permettront de détecter des biomarqueurs liés à une pathologie. Il existe trois moyens pour effectuer une analyse statistique : a) l'analyse selon une région d'intérêt (ROI - *region of interest*), où les régions d'intérêt sont prédéfinies b) l'analyse fondée sur les voxels (VB - *voxel-based*) et enfin c) l'analyse basée sur les fibres prédéfinies (*tract-based*). Dans cette thèse, nous nous concentrerons sur l'analyse statistique VB.

Application d'un test statistique approprié

Le choix d'un test statistique approprié pour la comparaison de populations est une étape importante. Il devrait être sensible et suffisamment flexible pour extraire les zones comportant des lésions. Dans cette thèse, trois tests statistiques particuliers ont été choisis, qui apparaissent dans la littérature, ou que nous avons synthétisé en combinant différents éléments de la littérature. Plus précisément, le premier test a été proposé pour la comparaison de populations et sera utilisé tel quel. Nous avons synthétisé un second test en combinant deux composants trouvés dans la littérature. Le troisième test provient de la théorie des forêts aléatoires et sera adapté à notre problème. Ces tests correspondent à l'ensemble des approches de l'état de l'art et seront comparés à notre propre approche pour le problème de la comparaison de populations.

La première approche choisie a été proposée par Verma *et al.* en 2007 [165] et consiste en une analyse statistique VB pour les modèles T2s. Les auteurs ont remarqué que l'application d'un modèle statistique standard pour les T2s n'était pas fiable, car les T2s ne suivent pas des lois gaussiennes multivariées dans leur espace initial (*i.e.* $\mathbb{R}^6$). Par conséquent, la tentative des auteurs pour estimer la sous-variété non-linéaire de l'espace des T2s, utilisant l'Isomap [154] (une technique de réduction de dimension non-linéaire, qui combine la méthode *multidimensional scaling* (MDS) [96] avec la théorie des graphes) était fondée et totalement innovante. De cette manière, les données sont transformées d'un espace de grande dimension où la métrique est riemannienne vers un espace de faible dimension où la distance est euclidienne. Une fois l'espace réduit déterminé (c'est-à-dire $\mathbb{R}^2$ selon leurs travaux), Verma *et al.* ont proposé de comparer les populations en utilisant le test T^2 de Hotelling, qui compare la moyenne des deux populations en supposant qu'elles ont la même matrice de covariance. Bien sûr, leur test statistique est également applicable au cas des T4s, en estimant l'espace réduit correspondant aux T4s. Bien que leur méthode comporte beaucoup d'idées très prometteuses et intéressantes, le test T^2 de Hotelling n'est pas très puissant et son inconvénient sera mis en évidence dans la section expérimentale sur des données synthétiques (section 6.3.2).

Nous avons synthétisé le deuxième test statistique en combinant un test sur deux populations [23], capable d'analyser des données de grande dimension même lorsque la dimension des données est très supérieure au nombre des observations, avec le test de permutation appliqué sur la matrice de distance inter-point proposé en [133]. Le test proposé par Biswas et Ghosh en 2014 [23] est non paramétrique et défini pour tous les types de matrices de distance. Il nous offrira la statistique d'intérêt. Ce composant donne une grande performance au test statistique, alors que la majorité des tests statistiques paramétriques ou non paramétriques est incapable de traiter ces problèmes mal posés. En outre, nous avons choisi d'effectuer des tests de permutation basés sur la redistribution des étiquettes des observations (c'est-à-dire de sujets normaux/anormaux), spécialement construits pour des matrices de distance entre points par Reiss *et al.* en 2010 [133], afin de simuler la distribution de la statistique d'intérêt (c'est-à-dire la statistique de Biswas et Ghosh). Cette distribution nous permettra de tester si la statistique associée au vrai étiquetage est une valeur extrême. Par ailleurs, un intervalle de crédibilité (*highest probability density (HPD) interval*) est calculé pour chaque p-valeur. Pour plus d'informations, le lecteur est renvoyé à la section 4.3.2.

La troisième méthode bénéficie de la théorie des forêts aléatoires (*random forests*, RF), proposée par Breiman en 2001 [30]. Les RFs sont un outil polyvalent et compétitif, y compris pour l'analyse statistique. Ses applications sont nombreuses, par exemple pour des problèmes de classification/régression, de détection d'anomalies (*via* l'estimation de densité), d'apprentissage de variétés (*manifold learning*), etc. [41, 42]. Dans le cas de la comparaison de populations, les classifieurs RF peuvent être utilisés, tandis que l'erreur de généralisation (GE) mesurée pour chaque donnée inconnue sera la statistique d'intérêt. Si les deux populations sont similaires, la GE sera très élevée, ce qui signifie qu'il est difficile de discriminer les deux groupes, tandis que d'autre part la GE est faible lorsque les populations se distinguent nettement. En outre, puisque le RF est un ensemble d'arbres de décision formés aléatoirement, où chaque arbre utilise un sous-ensemble aléatoire pour l'apprentissage, nous pouvons calculer une p-valeur avec son intervalle de crédibilité en divisant le nombre de mauvaises classifications par le nombre total d'échantillons inconnus utilisés dans la validation de chaque arbre de décision.

5. Modèles statistiques proposés

Dans les deux approches statistiques proposées, les modèles T4s sont sélectionnés en raison de leur grande capacité à représenter des structures complexes de fibres (les T2s fournissent des représentations moins justes). Ces T4s ont été estimés sur les données IRMd qui ont été initialement normalisées à l'aide de la méthode de Duarte-Carvajalino *et al.* [51] proposée en 2013. Cette méthode est une extension, pour les transformations non-linéaires, de la méthode de Tao et Miller [152] proposée pour les transformations linéaires en 2006.

Un autre point commun aux deux méthodes proposées est que nous avons besoin de construire une matrice de distance inter-point, pour toutes les paires possibles de données (normales et anormales) où chaque distance est calculée conformément à la métrique de tenseur proposée (eq. 5). En outre, les distances entre voxels sont imprécises en raison de l'erreur de recalage résiduelle. Pour contourner cet obstacle, au lieu de lisser les données, nous proposons d'introduire des informations de voisinage, sous la forme de deux patchs $3 \times 3 \times 3$ (un patch par sujet et par voxel d'intérêt). Chaque patch est extrait de deux voisinages plus grands (par exemple, des voisinages $5 \times 5 \times 5$). Les patchs sélectionnés sont ceux, parmi ceux possibles, qui minimisent la somme des distances inter-voxel.

Ensuite, d'après le travail de Verma *et al.* [165] pour estimer la sous-variété de T2 utilisant l'Isomap [154], nous avons choisi de calculer de façon similaire la sous-variété de T4. Plusieurs méthodes de réduction de dimension ont été testées, par exemple le *maximum variance unfolding* (MVU) [171], le *locality preserving projection* (LPP) [77, 78] et l'Isomap, sans remarquer aucune différence particulière d'un point de vue discriminatoire. Par conséquent, l'Isomap a également été sélectionné dans le cas des T4s. La matrice de distance inter-point contient toutes les informations nécessaires pour effectuer l'Isomap. La représentation de l'erreur de reconstruction en fonction de la dimension réduite d ($1 \leqslant d \leqslant 15$) a permis de conclure que le travail en $2D$ est adéquat pour le cas des T4s (cette dimension avait été retenue pour les T2s, par Verma *et al.*).

Enfin, les deux méthodes sont appliquées à l'analyse statistique voxel par voxel (VB). En outre, nous avons pour objectif de calculer des p-valeurs, c'est-à-dire la probabilité d'obtenir une statistique (par exemple une dissimilarité entre les populations) plus extrême que la valeur courante étant donnée la distribution de notre statistique sous l'hypothèse d'indiscernabilité. De plus, un intervalle de crédibilité sera estimé pour chaque p-valeur.

Première méthode proposée : comparaison statistique de populations - Application à la pathologie NMO

La première méthode proposée se réfère aux travaux présentés dans le chapitre 5, avec son évaluation dans le chapitre 6. Dans ce cas, les modèles de fODF T4 ont été sélectionnés [172] pour décrire les données IRMd, car notre objectif ultime était d'appliquer le test statistique proposé à la neuromyélite optique (NMO), pathologie qui en général provoque des lésions qui modifient l'orientation de la diffusion.

Pour chaque voxel dans le cerveau, une fois que toutes les données normales et anormales sont transformées dans l'espace réduit, nous proposons de décrire chaque population à l'aide de modèles de mélange de lois gaussiennes (*Gaussian mixture model*, GMM), en considérant un noyau gaussien associé à chaque point. Nous travaillons ainsi dans le cadre *kernel density estimation* (KDE). En outre, nous définissons comme statistique d'intérêt la distance (c'est-à-dire l'écart) entre les PDFs des deux GMMs. L'idée initiale était d'utiliser la version symétrique de la divergence de Kullback-Leibler (sKL), mais malheureusement il n'y a pas de formulation exacte pour calculer le sKL pour les GMMs et le calcul numérique est très long. En conséquence, nous avons trouvé dans la littérature une distance, notée $\mathcal{P}$, proposée par Sfikas *et al.* en 2005 [143], directement applicable aux GMMs.

Afin de calculer la p-valeur $\nu^\star$ liée à ce problème, la distribution $p(\mathcal{P})$ de la divergence $\mathcal{P}$ est nécessaire. Puisque cette distribution ne peut pas être déterminée analytiquement, nous proposons de déterminer la p-valeur par méthode de Monte Carlo, en redistribuant les étiquettes des données (test de permutation), ce qui nous permettra de produire de façon aléatoire un grand nombre d'échantillons $\{\mathcal{P}_1, \ldots, \mathcal{P}_N\}$ de cette distribution. La p-valeur $\nu^\star$

est définie comme $P(\mathcal{P}_n \geqslant \mathcal{P}_0)$ sous l'hypothèse nulle que les populations sont indiscernables, où $\mathcal{P}_0$ est la distance se référant à l'étiquetage réel des points. Par ailleurs, il est possible de calculer la distribution *a posteriori* de la p-valeur ν, c'est-à-dire $p(\nu|\mathcal{P}_1, \ldots, \mathcal{P}_N)$, ce qui nous permet d'extraire l'intervalle de crédibilité de la p-valeur, comprenant par exemple 99% de la masse *a posteriori* de $p(\nu|\mathcal{P}_1, \ldots, \mathcal{P}_N)$.

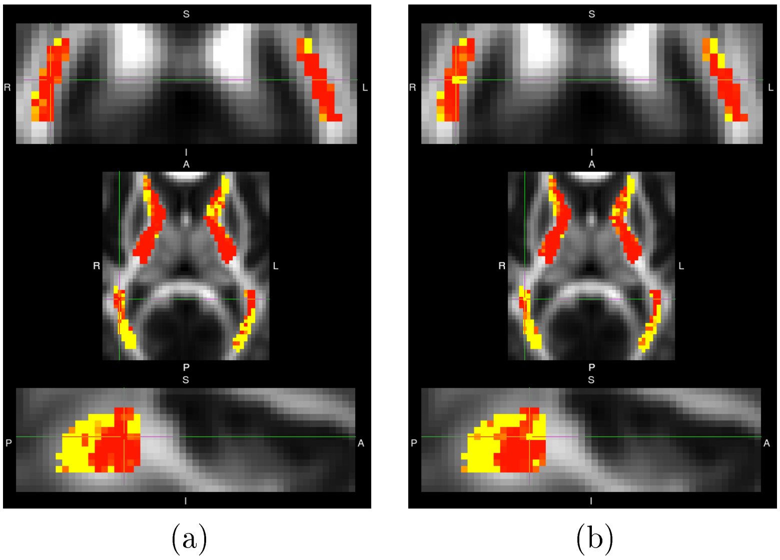


FIGURE 3: Visualisation des biomarqueurs obtenus (correspondant aux voxels pour lesquels la limite supérieure d'intervalle de crédibilité de la p-valeur est inférieure à 0.05 mise en évidence par la couleur rouge) d'une région particulière, représentée sur une image FA, pour (a) T4 fODF et (b) T2 fODF. On peut constater que le cas T4 fODF produit plus de biomarqueurs que le cas T2 fODF.

Le premier test statistique proposé a été évalué avec des données synthétiques et des données réelles. On met en évidence que les performances des modèles T4s sont meilleures que celles des modèles T2s, car plusieurs de biomarqueurs sont extraits dans le cas de T4 que le cas de T2 et les résidus de T2 contiennent de l'informations. En outre, l'approche statistique proposée est plus sensible que le test T^2 de Hotelling. De plus, dans le cas de données réelles où les régions liées à la maladie NMO ont été identifiées, l'approche statistique proposée est en cohérence avec les tests statistiques construits sur classifieurs RF et les tests de permutation de la matrice de distance inter-point. La figure 3 représente la limite supérieure des intervalles de crédibilité des p-valeurs induites (les biomarqueurs extraits avec les p-valeurs inférieures à 0.05 sont représentés en rouge).

Deuxième méthode proposée : comparaison statistique d'un sujet anormal *versus* la population normale - Application à la pathologie LIS

La seconde approche statistique est proposée au chapitre 7 pour le problème de la comparaison d'un individu avec la population normale, dans le cas de données anormales dispersées qui ne peuvent pas capturer toute la variabilité de la population anormale.

De la même manière que pour la première méthode, les données IRMd sont normalisées dans un espace commun. Initialement, nous avons pensé travailler avec les fODF T4s, mais nos expérimentations ont montré que la maladie LIS ne produisait pas beaucoup de lésions dans l'orientation de la diffusion. Nous avons donc choisi de considérer les profils de diffusion, modélisés comme des T4s [11].

Une autre différence se situe dans la détermination de l'espace réduit *via* Isomap. Contrairement à la méthode précédente, nous déterminons maintenant l'espace réduit en introduisant dans Isomap seulement les distances inter-point relatives aux distances entre les couples de données normales. De cette manière, le nuage normal ne sera pas influencé par les données anormales. Une fois l'espace réduit déterminé, les points anormaux sont placés dans l'espace réduit sans modifier la position des points normaux.

Concernant le test statistique révisé, un seul GMM est nécessaire, construit comme précédemment, afin de décrire la population normale. La similitude de chaque point anormal au nuage normal est mesurée en calculant sa densité de probabilité avec le KDE/GMM du groupe normal. Cette densité est la statistique d'intérêt.

Ensuite, comme pour le test de permutation, on va approximer la p-valeur à l'aide de simulations de Monte Carlo. On génère des échantillons selon le KDE/GMM, ce qui délivre des densités p_i. La p-valeur mise en jeu est égale à la probabilité $P(p_i \leqslant p_0)$ sous l'hypothèse nulle que le sujet à évaluer appartient à la population normale. En plus, un intervalle HPD peut être extrait pour chaque p-valeur, de façon similaire à l'approche précédente.

L'application aux données réelles a été effectuée, en particulier pour la pathologie LIS. Des zones spécifiques, comprenant le système moteur, ont été sélectionnées pour être étudiées. Dans cette étude, notre objectif était de calculer le pourcentage de lésions par zone (c'est-à-dire le pourcentage de voxels pour lesquels la limite supérieure de l'intervalle de crédibilité de la p-valeur est inférieure à 0.05). Ces zones peuvent être séparées en deux grandes régions. La première région (région 1) contient des zones proches de la moelle épinière (située dans la partie inférieure du cerveau, par exemple *pontine crossing tract*, *corticospinal tract* gauche et droit, *medial lemniscus* gauche et droit), tandis que la seconde région (région 2) regroupe les zones situées dans les parties moyenne et supérieure du volume du cerveau relié à la moelle épinière (comme *posterior limb of internal capsule* gauche et droit, *superior corona radiata* gauche et droit).

L'analyse statistique a conclu que le pourcentage de lésions dans la région 1 était plus élevé que dans la région 2. En outre, la quantité de lésions dans la région 2 dépend du patient. Par ailleurs, l'analyse statistique est cohérente avec l'avis médical.

Les pourcentages de lésions détectées par les méthodes proposées ont suscité notre intérêt pour explorer les performances de différentes variations de la métrique de tenseur, construites à partir de la métrique initiale, proposée dans l'équation 5. Ces variations peuvent nous fournir un outil plus sensible pour discriminer plusieurs différences. Nos expérimentations (voir la section 7.2.4) ont montré qu'en effet il y a une variation particulière de la métrique de tenseur qui peut fournir de meilleures performances que la métrique utilisée pour le moment.

Enfin, les comparaisons entre les modèles T4 et T2 de diffusion sont présentées. De plus, des comparaisons avec les approches classiques fondées sur les statistiques du z-score relatives aux mesures de FA et MD ont été effectuées. Les analyses des T2s de diffusion et d'images FA/MD produisent des pourcentages de lésions plus élevés que l'approche proposée sur le modèle de T4. Malheureusement, en raison de l'absence de vérité terrain, il est difficile de tirer des conclusions sûres. Peut-être une comparaison sur des données synthétiques pourrait-elle nous éclairer davantage pour tirer des conclusions. D'autre part, nous pouvons affirmer qu'un modèle de tenseur d'ordre supérieur, comme le T4, a la potentialité de mieux capturer la variabilité de la maladie qu'un autre modèle moins adapté, comme le T2, ou les mesures scalaires simples (par exemple, les images FA et MD).

6. Conclusion et perspectives

Pour conclure ce résumé, nous avons choisi de souligner quelques points caractéristiques et de donner des orientations pour les travaux futurs.

Le premier point auquel prêter attention est la normalisation des données. Puisque les données de tenseur ou d'IRMd contiennent des informations d'orientation, la normalisation de données se compose de deux étapes, le recalage spatial et la réorientation de données. À ce stade, nous devrions indiquer que la normalisation d'IRMd est moins risquée que la normalisation de tenseurs.

Deuxièmement, la prise en compte des T4s au lieu des T2s a conduit à une analyse statistique plus efficace et robuste, en particulier dans le cas des croisements de fibres. Il faut noter qu'un modèle plus juste présente un meilleur potentiel pour un diagnostic précoce.

Troisièmement, le calcul de la sous-variété des modèles T4 a l'un des rôles les plus incontournables dans nos approches. Les distances euclidiennes en grande dimension peuvent lisser les différences. Par opposition, les distances géodésiques utilisées pour déterminer la matrice de distances inter-points ensuite utilisée par l'Isomap, permettent de mieux mettre en évidence les dissimilarités. En outre, la prise en compte de l'erreur de recalage résiduelle à l'aide de patchs les mieux adaptés et en introduisant des informations de voisinage dans l'estimation de la matrice de distance inter-point a montré de meilleurs résultats que le lissage des données.

Par ailleurs, avoir les deux approches statistiques proposées nous donne la flexibilité nécessaire pour analyser les données pathologiques indépendamment de leur nombre.

En ce qui concerne les orientations pour les travaux futurs, nous proposons d'examiner la combinaison du recalage d'IRMd (on recale les données brutes sans les orienter) suivi par l'estimation de tenseurs sur les données obtenues complétées par la réorientation de ces tenseurs. Par exemple, dans le cas de la réorientation de T4 (ou même T2), les méthodes présentées dans le chapitre 3 peuvent être utilisées.

Deuxièmement, l'analyse statistique basée sur la nouvelle variation de la métrique proposée sur le tenseur (mentionnée dans la section 7.2.4) devrait être également évaluée. Probablement, les pourcentages de lésions détectées peuvent être plus élevés que ceux détecte actuellement.

Une autre perspective est liée à la construction de chaque GMM par population. Il sera très intéressant de connecter plusieurs sujets dans l'espace réduit au même noyau gaussien, au lieu de considérer un noyau par sujet. De cette manière, tous les problèmes de surajustement occasionnel peuvent être évités.

Pour continuer, dans cette étude nous avons concentré nos efforts dans l'étude de ce qui se passe dans certaines régions du cerveau, déjà connues comme atteintes par la maladie. Dans le cas des maladies inflammatoires, il peut être fascinant d'étudier et de détecter d'éventuelles lésions potentielles dans d'autres régions du cerveau, peut-être totalement nouvelles pour la communauté médicale.

Ensuite, une étude approfondie des propriétés de diffusivité liées à chaque voxel extrait comme biomarqueur, par exemple en analysant le niveau significatif des variations de la diffusion dans chaque direction, pourrait permettre d'évaluer quelles directions de diffusion sont responsables de la caractérisation du voxel comme biomarqueur.

De plus, les capacités des approches statistiques proposées pour réaliser un diagnostic précoce restent à examiner, sous l'expertise de neurologues.

Pour conclure, les applications des approches proposées pour analyser des descripteurs de plus haut niveau (par example les faisceaux de fibres et les connectogrammes) peuvent être testées.

Contents

List of Figures

List of Tables

Abbreviations

AD	**A**lzheimer's **D**isease
ADC	**A**pparent **D**iffusion **C**oefficient
AE	**A**ngular **E**rror
CDF	**C**umulative **D**istribution **F**unction
dODF	**d**iffusion **O**rientation **D**istribution **F**unction
DSI	**D**iffusion **S**pectrum **I**maging
DT	**D**ecision **T**ree
DTI	**D**iffusion **T**ensor **I**maging
DT4	**4**th order **D**iffusion **T**ensor
DW-MRI	**D**iffusion **W**eighted **M**agnetic **R**esonance **I**maging
EAP	**E**nsemble **A**verage **P**ropagator
EEG	**E**lectro**E**ncephalo**G**raphy
FA	**F**ractional **A**nisotropy
fMRI	**f**unctional **M**agnetic **R**esonance **I**maging
fODF	**f**iber **O**rientation **D**istribution **F**unction
FS	**F**inite **S**train algorithm
GA	**G**eneralized **A**nisotropy
GFA	**G**eneralized **F**ractional **A**nisotropy
GMM	**G**aussian **M**ixture **M**odel
HARDI	**H**igh **A**ngular **R**esolution **D**iffusion **I**maging
HD	**H**ilbert **D**ecomposition
HOT	**H**igher **O**rder **T**ensor
HPD	**H**ighest **P**robability **D**ensity
i.i.d.	**i**ndependent and **i**dentically **d**istributed

IVIM	**I**ntra-**V**oxel **I**ncoherent **M**otion
KDE	**K**ernel **D**ensity **E**stimation
LIS	**L**ocked-**I**n **S**yndrome
LPP	**L**ocality **P**reserving **P**rojection
MD	**M**ean **D**iffusivity
MDS	**M**ulti**D**imensional **S**caling
MEG	**M**agneto**E**ncephalo**G**raphy
MRI	**M**agnetic **R**esonance **I**maging
MS	**M**ultiple **S**clerosis
MVU	**M**aximum **V**ariance **U**nfolding
NMO	**N**euro**M**yelitis **O**ptica disease
NMR	**N**uclear **M**agnetic **R**esonance
ODF	**O**rientation **D**istribution **F**unction
PCA	**P**rincipal **C**omponent **A**nalysis
PDF	**P**robability **D**ensity **F**unction
PPD	**P**reservation of **P**rincipal **D**irections algorithm
QBI	**Q**-**b**all **I**maging
RA	**R**elative **A**nisotropy
RF	**R**andom **F**orest
ROI	**R**egion **O**f **I**nterest
SD	**S**pectral **D**ecomposition
SH	**S**pherical **H**armonics
sKL	**s**ymmetrized **K**ullback **L**eibler divergence
SVD	**S**ingular **V**alue **D**ecomposition
SVM	**S**upport **V**ector **M**achine
T2	**2**nd order **T**ensor
T4	**4**th order **T**ensor
TB	**T**ract-**B**ased analysis
VB	**V**oxel-**B**ased analysis
WM	**W**hite **M**atter of brain

Chapter 1

Introduction

This introductory chapter presents the scientific problem, the contribution of the thesis, and the organization of the dissertation.

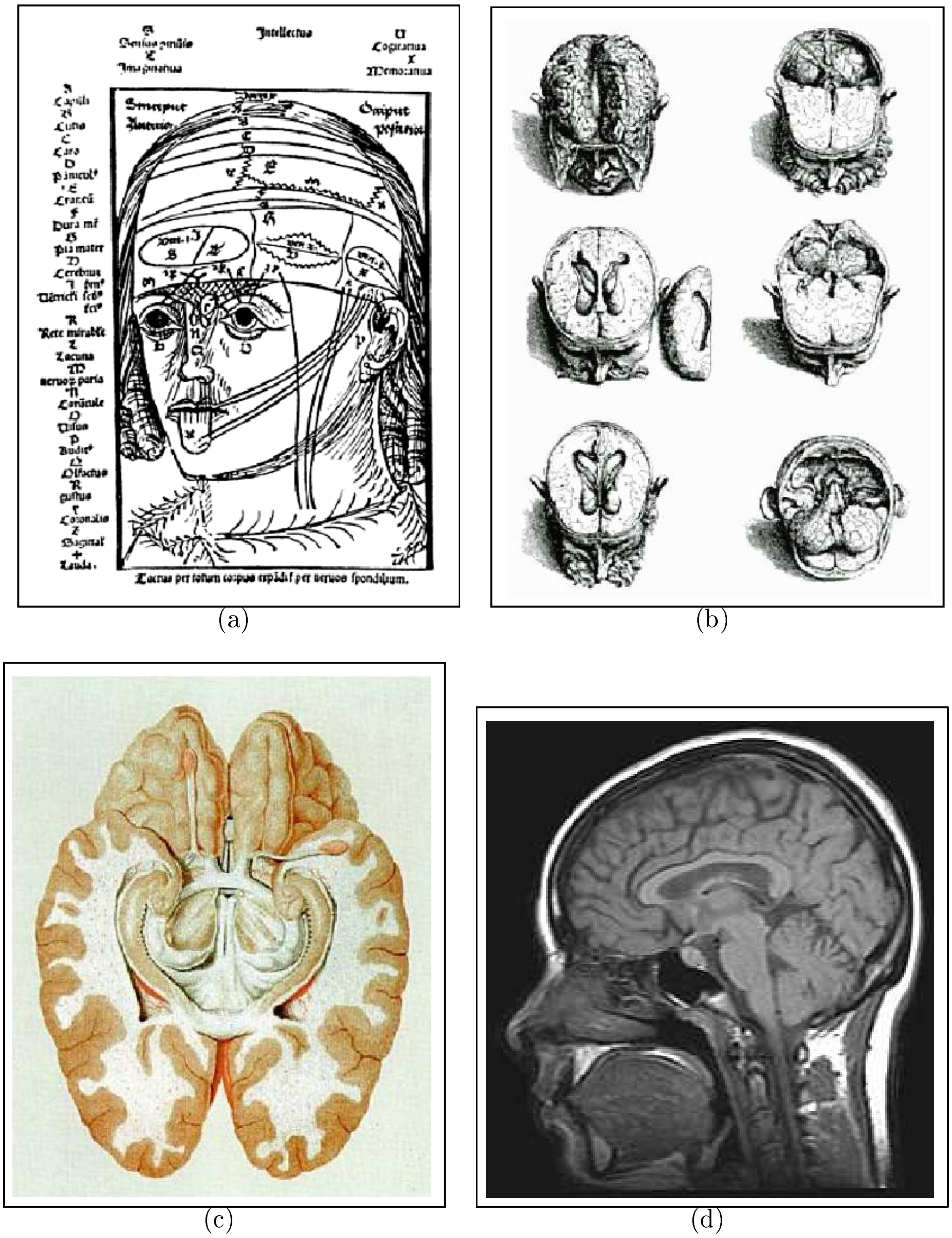
1.1 Definition of the Scientific Problem

Understanding human brain's structure and functionality has always fascinated the human kind, since the brain is one of the most fundamental organs, but also the most complex and multi-task organ in our body. Figure 1.1 illustrates a few characteristic steps through the evolution of brain studies in the past centuries. Nowadays, a considerable percentage of the brain's capabilities still remains unidentified and open to research.

Another reason that triggered the interest of the scientific community to extensively comprehend the brain is the lack of deep knowledge concerning the specificity of various brain pathologies. For instance, certain disorders are restricted to specific areas of the brain (*e.g.* locked-in syndrome (LIS), Parkinson's disease, dyslexia etc.), while, on the contrary, several inflammatory or neurodegenerative diseases can potentially affect the entire human brain (neuromyelitis optica (NMO), multiple sclerosis (MS), Alzheimer's disease etc.) [104]. The development of efficient and comprehensive automated diagnostic tools can help us understand the aspects of the disease, and eventually allows us to monitor the patient's condition. Additionally, treating the disease as early as possible by systematically guiding the doctor's decisions concerning the adjustment of patient's treatment is always desirable.

For many decades, computer and medical sciences have been collaborating in this direction with the aid of physics, under the names of Neuroscience and Neuroinformatics. In order to gain knowledge related to a disease, handy examinations were constructed that are suitable to extract measurements useful to characterize particular diseases. Highly informative data acquisition techniques, such as Magnetic Resonance Imaging (MRI), Diffusion Weighted

Magnetic Resonance Imaging (DW-MRI) and functional Magnetic Resonance Imaging (f-MRI) were developed to observe the brain and gather various interesting measurements concerning the structure and the function of the brain.



(a) (b)

(c) (d)

Figure 1.1: The evolution of brain studies. (a) medieval view of brain's structure and function (by Magnus Hundt, in 1501 - probably first printed anatomical figure of the head), (b) drawing appeared in Book VII of Andreas Vesalius's Fabrica (1543) depicting horizontal segments of the human head, (c) image from Félix Vicq d'Azyr's atlas of the human brain (1786 - the most accurate before the development of neurohistology) and (d) a modern MRI scan of the brain (1971 - nowadays).

For example, DW-MRI is the first technique to capture information related to the structural anatomy of white matter (WM) or even grey matter of the brain, *in vivo* (although other competitive tools, more suitable for grey matter studies than DW-MRI, also exist). In fact, DW-MRI measures the diffusion of water molecules across several directions in the $3D$

space, revealing the structure of the white matter (WM) of the brain. Moreover, DW-MRI is a totally non-invasive routine, due to the fact that the patient is not injected with any radioactive tracer. Neither does any exposure to ionizing radiation take place through the whole examination, reducing the appearance of complications to a minimum level. Generally speaking, DW-MRI data will monopolize our interest in this study.

Such an innovative imaging modality lead us to seek advanced tools for image processing, modelling and analysis. For instance, DW-MRI data are commonly modelled using tensors (the reader is referred to chapter 2 and [112] for more information). These tensor models concentrate solid clues about the structural representation of the diffusion and offer an easier way to visualize the diffusion properties at each voxel of the brain than raw DW-MRI data. Due to the fact that the human brain contains complex structures representing several bundles of fibers in a high percentage of voxels (almost 50%), powerful and competitive tensor models such as High Order Tensor (HOT) models [14, 122, 172] are needed. HOT tensors can assist us to define representative descriptions that can capture as much information as possible that is embedded in the DW-MRI data.

Finally, studying a particular disease, for example with a view to biomarker extraction, requires a significantly large repository of data, so as to define the control (normal) population with healthy subjects, and to characterize the variability of the disease by considering an extensive pathological (abnormal) population. On the other hand, if the desirable task is to follow up patients' state, each patient can be alternatively tested individually, for a given set of different in time scans, against the normal population.

In this thesis, populations comparisons, along with individual *versus* normal population comparisons, are addressed *via* specific statistical tests that we propose, potentially aiming at early diagnosis, biomarker extraction and patient follow-up. Noticing the lack of mapping techniques and statistical analysis tools for HOTs in the literature, and considering their powerful abilities to describe DW-MRI data, we chose to work with fourth order tensors (T4s) [11, 172], which is a particular case of HOTs.

1.2 Thesis' Contributions

Performing population comparisons requires an initial important task, known as data normalization (*e.g.* for DW-MRI or tensor data), in order to align brains of different subjects together. Due to the fact that DW-MRI data and tensor models contain orientated information, a single spatial registration is not enough. A reorientation step is crucial and mandatory, too. During the first year of this thesis, we focused on the problem called tensor reorientation [2, 5, 135]. In this case, the tensor models should be correctly reorientated in order to match with the new underlying fiber orientation in the new common space. The results of this work on T4 models are presented in chapter 3.

To continue, in order to compare different subjects in the tensor space, metrics taking into consideration the properties of the diffusivity profiles should be defined (see chapter 2 where a proposed tensor metric is presented, inspired by the work of Tarantola [153] in physics, along with other metrics that can be found in the literature). These tensors metrics will allow us to construct inter-point distance matrices that will be introduced into dimensionality reduction techniques in order to perform statistical analysis, robustly, in a reduced space.

The need for sensitive models and statistical tests with a view to early diagnosis and prognosis, disease staging, patient follow-up etc. started to grow rapidly. Comparisons between healthy and pathological data in order to detect patterns of lesions for a given disease, or to extract biomarkers *via* population modelling and comparisons, drew our attention during the next two years of this work. The majority of this thesis is focused on developing advanced statistical tests for high order tensor models in order to solve the problem of populations comparison (*i.e.* healthy *versus* pathological groups, see chapter 5). In this case, given a normal population corresponding to a set of healthy individuals, and an abnormal population containing pathological datasets, we can highlight voxels, or group of voxels in the brain with structural abnormalities resulting from the disease (*i.e.* biomarkers), with the aid of a suitable statistical test, that will compare the two populations.

Furthermore, individual pathological datum *versus* the healthy population comparisons were studied and proposed in this thesis with view to patient follow-up (see chapter 7). In this case, methods perform statistical comparisons individually for each patient against the normal population in time series of scans. The last application is considered to be very helpful and sometimes the only solution in cases where the variability of the diseased population cannot be fully captured (*e.g.* diseases with no specific drawn patterns, either much variability *e.g.* often in traumatic brains, or a minor number of patients available related to the same disease).

To sum up, the main contributions of this thesis are contained in the following points:

- Proposed tensor metric (chapter 2, section 2.3.4).
- Study and evaluation of T4 reorientation schemes (chapter 3).
- Statistical models using tensor information for the following problems:
  - Population *VS* Population: application to NMO disease (chapters 5, 6).
  - Individual *VS* Normal Population: application to LIS syndrome (chapter 7).

1.3 Organization of the Dissertation

The dissertation is organized as follows:

Chapter 2 presents the DW-MRI modality and the possible ways to model the acquired data, for example using tensor models. In addition, several tensor metrics are presented, along with an original one. Furthermore, high level descriptions of DW-MRI data are discussed, such as fiber tracts and connectomes.

Chapter 3 addresses the pre-processing steps for DW-MRI data. The problem of tensor reorientation is explained and a study on reorientation schemes for T4 models is presented, based on T4 decompositions into T2s followed by T2 reorientations. Two T4 decompositions are described along with two T2 reorientation methods. Experimental results on synthetic and real data are included.

In chapter 4, the general steps to devise statistical atlases are highlighted and the problem of biomarkers detection is explained. Moreover, state-of-art techniques are presented and discussed.

The construction of a statistical atlas, for the problem of population *versus* population comparison, is addressed in chapter 5. The proposed approach fits T4 models on the DW-MRI data and performs voxelwise statistical analysis in a reduced space. Applications to synthetic cases were achieved along with application to NMO disease which allowed us to evaluate the proposed method in comparison with several other methods. Experimental results are included in chapter 6.

For the case of sparse (pathological) populations, we propose in chapter 7 a solution to the problem of individual *versus* normal population. Application to LIS pathology is selected and presented.

Finally, the conclusion of this thesis along with suggestions for future work are contained in chapter 8.

Chapter 2

In Vivo Probing and Modelling the Diffusion of Water Molecules in the Human Brain

In 1965, Stejskal's and Tanner's method for probing the diffusion of water molecules by using Nuclear Magnetic Resonance (NMR) resulted into a non-invasive technique called Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) [149]. DW-MRI's ability to capture information related to the underlying white matter (WM) structure of the brain was revolutionary. In this chapter, a brief introduction about the properties of the diffusion of water molecules in the human brain is presented, along with the description of several suitable models to represent DW-MRI data.

2.1 Brownian Motion of Water Molecules

Around $77-78\%$ of the human brain consists of water [105]. DW-MRI measures the signal of the proton (^1H) in water molecules (H_20), which corresponds to the movement of water molecules, by applying a set of magnetic gradient directions to the subject that we examine. This permits us to measure the motion of the molecules across these directions. This random movement is known as intra-voxel incoherent motion (IVIM), random motion, or **Brownian motion**. Examples showing different kind of Brownian motions are presented in figure 2.1.

Unfortunately, this is not the only motion which can be observed in the measured signal. Another type of detected motion is known as bulk motion. It is the result of the subject's movements during the scanning procedure. Furthermore, brain pulsation is also considered as bulk motion. The magnitude of this motion is usually larger than Brownian motion and can be easily removed or corrected as an artifact.

Concerning the aspects of the applied magnetic field, it is always parallel to the z axis, but it varies in space. It is modulated according to the current $3D$ position (x, y, z):

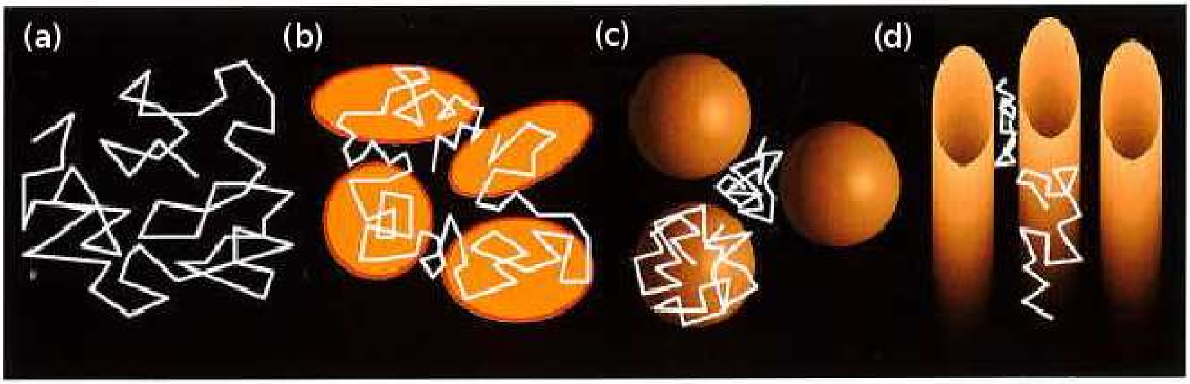


FIGURE 2.1: Different examples of diffusion. (a) random isotropic diffusion of water with unconstrained shape, (b) and (d) anisotropic shape constrained diffusion and (c) isotropic shape constrained diffusion. Image reproduced from [112].

$B(x, y, z) = B_0\,[0, 0, (ax + by + cz + 1)]^T$, given the main magnetic field B_0 (see figure 2.2). In this way, the set of magnetic gradient directions consists of unit vectors of the form $g = [a, b, c]^T \,/\, \| \, [a, b, c]^T \, \|$. Applying specific gradients in many directions allows us to measure the Brownian motion by sampling the $3D$ unit sphere or hemisphere, useful to understand the neural network (*i.e.* structure) of the brain. For more details about the acquisition process, the reader is referred to [112].

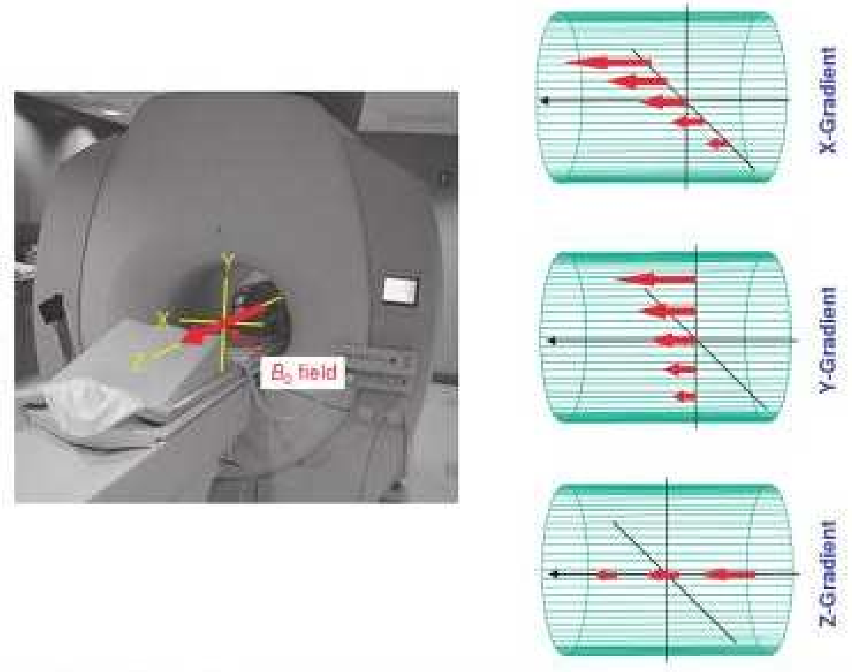


FIGURE 2.2: The strength of the magnetic field B (red arrow) is linearly modulated along each of the three axes. Figure appeared in [112].

The measured diffusion can be **isotropic** (meaning the same signal in each gradient direction that can correspond to trapped water without any particular information, *e.g.* fig. 2.1 a) and c)), or **anisotropic** that reveals, except from the magnitude of the diffusion, also the orientation of the underlying fiber structure and has the potential to indicate several neural connections, structures etc. (*e.g.* fig. 2.1 b) and d)).

The Propagator Description of Water Molecules

Random diffusion of water molecules flowing from point $\boldsymbol{r}$ to $\boldsymbol{r}'$ in time t can be described by a probability density function (PDF) $p(\boldsymbol{r}', t|\boldsymbol{r})$ [92] which follows Fick's law:

$$\frac{\partial p(\boldsymbol{r}', t \mid \boldsymbol{r})}{\partial t} = \mathbf{D}\nabla^2 p(\boldsymbol{r}', t \mid \boldsymbol{r}), \tag{2.1}$$

assuming that the diffusion is homogeneous in the medium, where ∇^2 is the Laplacian operator and $\mathbf{D}$ is the corresponding diffusion tensor (see section 2.3.1.1).

Equation 2.1 describes the propagator of a single water molecule, but several water molecules can exist in a voxel, participating in the same motion. For that reason, another useful quantity is the probability of molecules to reach point $\boldsymbol{r}'$ in time t, also known as *ensemble average propagator* (EAP):

$$p(\boldsymbol{r}', t) = \int p(\boldsymbol{r})\ p(\boldsymbol{r}', t \mid \boldsymbol{r})\ d\boldsymbol{r}, \tag{2.2}$$

where $p(\boldsymbol{r})$ corresponds to the density of water molecules in the initial position $\boldsymbol{r}$.

Given the initial state where particles start at point $\boldsymbol{r}$, $p(\boldsymbol{r}', 0 \mid \boldsymbol{r}) = \delta(\boldsymbol{r}' - \boldsymbol{r})$, the Dirac function, the solution of eq. 2.1 provides us with the following Gaussian description of the propagator [92, 134]

$$p(\boldsymbol{r}' - \boldsymbol{r}, t) = \left((4\pi t)^3 \, |\mathbf{D}|\right)^{-1/2} \exp\left(-\frac{(\boldsymbol{r}' - \boldsymbol{r})^T \mathbf{D}^{-1} (\boldsymbol{r}' - \boldsymbol{r})}{4t}\right). \tag{2.3}$$

2.2 DW-MRI Data

Several books address the DW-MRI acquisition process (see *e.g.* [91, 92, 112]). As a result, the purpose of this section is not to focus on this procedure, but to mention the most important properties of DW-MRI data and to remind the equations that describe them.

First of all, the characteristic term of "*diffusion-weighted*" is given due to the utilization of a set of magnetic gradient directions along which the diffusion is measured [112]. In other words, grey levels in DW images represent the diffusion (along each gradient direction). The most common equation that defines the acquired signal intensity S in a gradient direction $\mathbf{g}$ ($3D$ unit vector) is the following:

$$S(\mathbf{g}) = S_0 \mathrm{e}^{-b\ d(\mathbf{g})}, \tag{2.4}$$

where S_0 is the signal intensity with zero gradient (*i.e.* $\mathbf{g} = (0,0,0)$), $S(\mathbf{g})$ is the measured signal along the gradient direction $\mathbf{g}$, b is known as b-value and is the acquisition parameter, and finally $d(\mathbf{g})$ is the (positive) diffusion value depending on $\mathbf{g}$. For example, given diffusion tensor $\mathbf{D}$ (*i.e.* second order tensor, section 2.3.1.1), equation 2.4 is written as:

$$S(\mathbf{g}) = S_0 \mathrm{e}^{-b\ \mathbf{g}^T\ \mathbf{D}\ \mathbf{g}}. \tag{2.5}$$

For more information about how equations 2.4 and 2.5 were derived the reader is referred to [92, 112]. An example of DW-MRI signal is shown in figure 2.3.

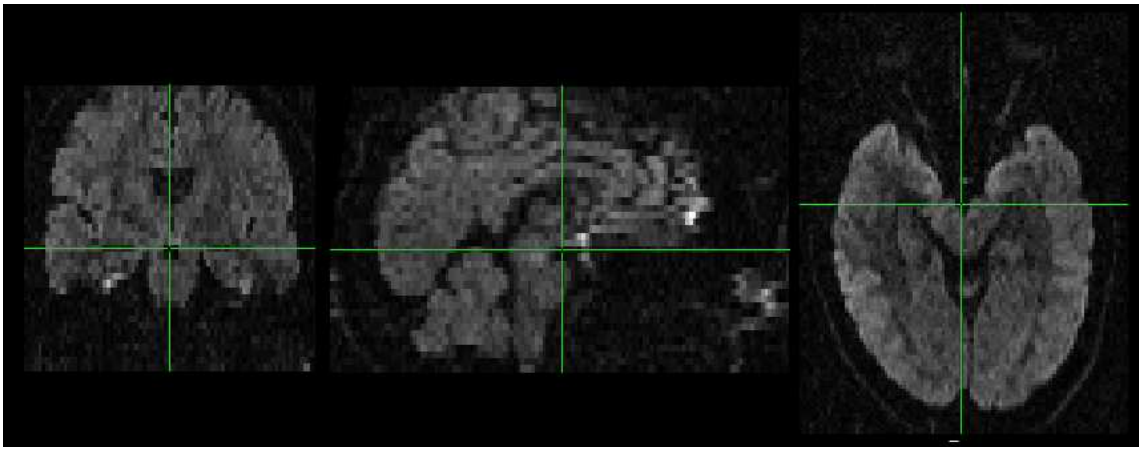


FIGURE 2.3: Three slices of DW-MRI data in a given gradient direction. (left) coronal view, (middle) sagittal view and (right) axial view of the human head.

Due to the fact that each DW-MRI image, that stores the signal connected to the diffusion in a single direction, can be altered by the presence of noise, we need to gather as many images as possible, for a set of different gradient directions in a limited scan-time, in order to increase the accuracy of the measurements. Acquisition techniques with a lot of gradients directions (*e.g.* 30, 40 etc.) are known as **HARDI** (High Angular Resolution Diffusion Imaging) methods [160]. A more extensive discussion about HARDI techniques will take place in the following sections.

2.3 Models for DW-MRI Data

2.3.1 Tensor Models

Diffusion models are tools that allow us to represent the diffusion of water molecules that captures the structure of the WM of the brain. The first attempt took place by the introduction of Diffusion Tensor Imaging (DTI). DTI uses very few gradients (*e.g.* 6) and can only describe a single direction of diffusion in the underlying fiber architecture (see subsection 2.3.1.1). As neuroimaging gained knowledge, it was proved that more complex structures containing bundles of fibers can appear in the human brain (in almost 50% of voxels) [91, 92, 160, 161], meaning that more flexible models that can capture, in detail, the underlying shape of the fibers are required [12, 14, 159, 172, 173]. In this direction, Higher Order Tensor (HOT) models became popular (see subsection 2.3.1.2). Tensor models can describe diffusion (*i.e.* apparent diffusion coefficient - ADC profiles) and fibers orientation (fiber Orientation Distribution Functions - fODF profiles) (subsection 2.3.2). Subsection 2.3.4 defines some important tensor metrics, that will be needed in the rest of the dissertation.

2.3.1.1 Diffusion Tensor Imaging and Second Order Tensors

DTI tensors, often called second order tensors or simply T2s, were proposed by Basser *et al.* in 1994 [18]. Basser *et al.* modelled the propagator (eq. 2.3) in the form of ellipsoids. Although it is named as diffusion ellipsoid, in fact it represents the iso-probability density of the propagator (eq. 2.3) and it should not be confused with the diffusion profile (which is commonly visualized by a "peanut" shaped representation, as we will show later).

As it is known, an ellipsoid is defined by six parameters (fig. 2.4). Theoretically, at least six DW-MRI measurements across six independent gradient directions are need to determine an ellipsoid.

The information enclosed in the six previously mentioned parameters of an ellipsoid can be mathematically gathered in a tensor matrix $\mathbf{D}_2$, such that the diagonalization of $\mathbf{D}_2$ produces the three eigenvectors v_1, v_2, v_3 and their corresponding eigenvalues $\lambda_1, \lambda_2, \lambda_3$. A T2 tensor can be represented by the following 3x3 symmetrical matrix:

$$\mathbf{D}_2 = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{bmatrix}. \tag{2.6}$$

Given matrix $\mathbf{D}_2$ and a variable $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$ in the $3D$ space, Basser *et al.*'s ellipsoid is defined as the set of $\boldsymbol{\mu}$'s such that $\boldsymbol{\mu}^T\ \mathbf{D}_2^{-1}\ \boldsymbol{\mu} = c$, where c is a constant (notice the connection with the numerator in the exponential of the Gaussian propagator in eq. 2.3).

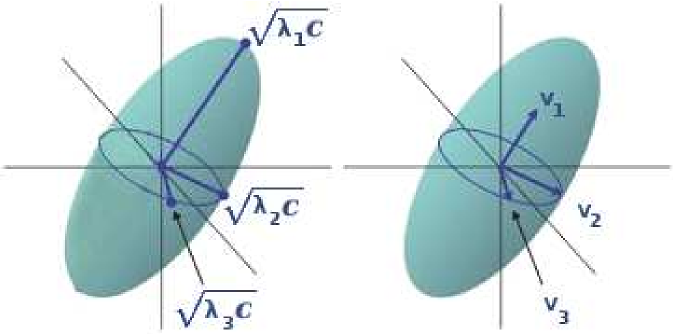


FIGURE 2.4: Six parameters (three eigenvalues $\lambda_1, \lambda_2, \lambda_3$, two parameters to define the first eigenvector v_1, one for the second v_2, and zero for the third one v_3) are needed to define a $3D$ ellipsoid (image reproduced from [112]).

The symmetry property of the tensor matrix $\mathbf{D}_2$ ($D_{ij} = D_{ji}$) results into having six unique coefficients (same in number as the parameters of the ellipsoid). These coefficients allow us to define another representation of the T2 as a 6×1 vector using Voigt's notation, $[D_{xx}, D_{yy}, D_{zz}, D_{yz}, D_{xz}, D_{xy}]^T$.

In addition, a useful function that can be defined with the use of a tensor matrix (eq. 2.6), is known as "*diffusivity function*". It is a symmetrical positive real-valued function which returns the diffusion along a given gradient direction $\mathbf{g} = (g_1, g_2, g_3)^T$, where $\|\mathbf{g}\| = 1$:

$$d(\mathbf{g}) = \sum_{i=1}^{3} \sum_{j=1}^{3} D_{ij}\ g_i\ g_j = \mathbf{g}^T\ \mathbf{D}_2\ \mathbf{g}. \tag{2.7}$$

The diffusivity function $d(\mathbf{g})$ has the shape of a peanut in contrast to the ellipsoid of the propagator (see figure 2.5). In this dissertation, we will work with the peanut shaped glyphs.

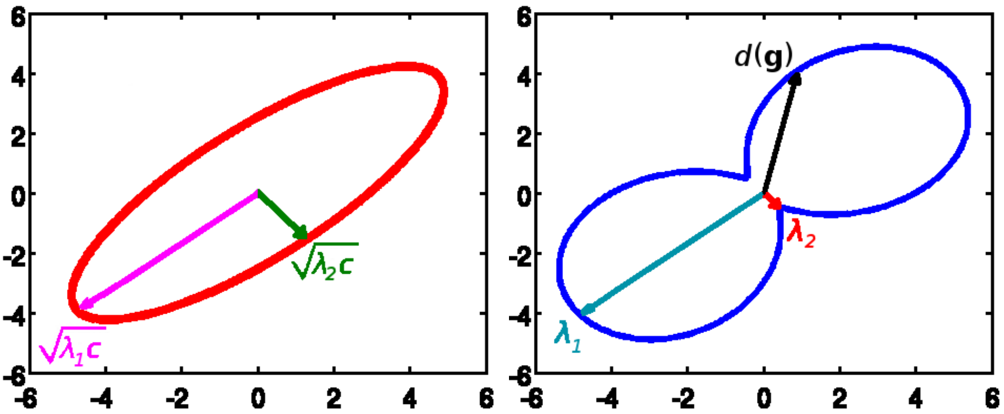


FIGURE 2.5: Plotting $2D$ slices of the $3D$ diffusion ellipsoid for $c = \lambda_1$ (on the left) and the diffusivity function $d(\mathbf{g})$ as "peanut shaped" representation (on the right).

Anisotropy indices for DTI

Two popular scalar maps proposed by Basser and Pierpaoli in 1996 [21], the fractional anisotropy **(FA)** and the relative anisotropy **(RA)**, have been widely used in DTI:

$$\mathbf{FA} = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{2.8}$$

$$\mathbf{RA} = \sqrt{\frac{1}{3}} \frac{\sqrt{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}}{\langle\lambda\rangle}, \tag{2.9}$$

where $\langle\lambda\rangle$ is the mean diffusivity **(MD)** and is defined as:

$$\langle\lambda\rangle = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}, \tag{2.10}$$

and $\mathbf{FA} \in [0, 1]$, while $\mathbf{FA} = 0$ corresponds to the isotropic diffusion.

T2 Model limitations and efficient ways to circumvent them

As can be imagined, the simplicity of the T2 model comes with a price. It appears that the human brain contains complex structures with more than one bundle of fibers crossing each other, in almost 50% of voxels [91, 92, 160, 161]. Figure 2.6 describes some characteristic cases of fiber bundles that appear frequently in real DW-MRI data, and unfortunately T2 fails to represent correctly the underlying fiber structures. Fiber Orientation Distribution Functions (fODF), shown in the fourth column, give better results than T2s. A fODF is a function on the sphere, describing the orientation of the fibers (see subsection 2.3.2).

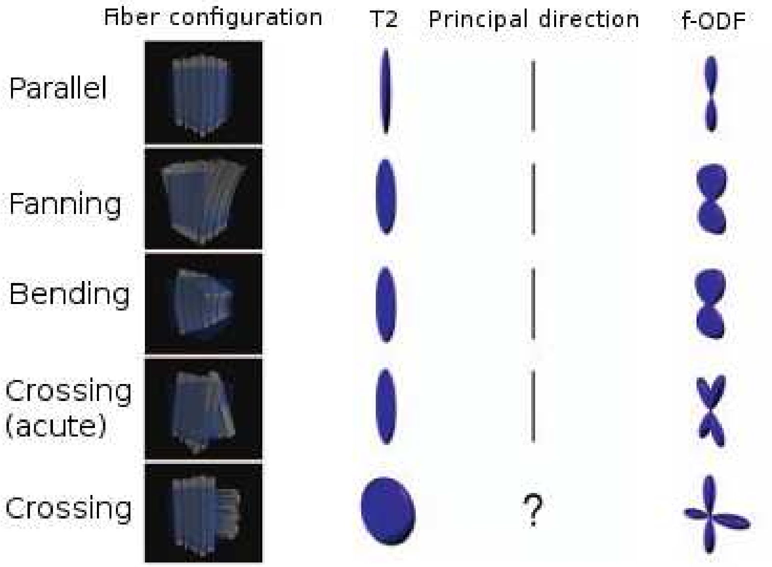


FIGURE 2.6: Several examples of bundles of fibers that occur frequently in real data (first column). The second column contains the T2 models, while the third column contains the principal direction of the T2 models and finally the fourth column shows the fODF results, which are better than T2s. (original image appeared in [91]).

In order to bypass the T2 limitations, researchers started thinking about the points where information is lost. Firstly, it is undoubtable that fitting a model able to describe more complex data is the first key point. A solution to this problem is given by considering more complex tensor models (further discussion is proposed in section 2.3.1.2). But before addressing this point, there are two more points to stand.

The second limitation takes place during the data acquisition process. Usually, hardware's ability to measure the signal is limited around $2-3\ mm$ per dimension. In order to achieve higher resolution of DW-MRI images, novel hardware should be constructed, which could produce larger magnetic fields (corresponding to b-value higher than 1000 s/mm^2 currently used in DTI). Moreover, increasing the magnitude of the magnetic field also increases the risk taken by the patient and the researcher should further study that risk. At the time when T2 models appeared, 1.5 Tesla fields were commonly used. But nowadays the situation is better since for example, in experimental level, human studies have been performed at up to 9.4 Tesla [162] and animal studies have been performed at up to 21.1 Tesla [132].

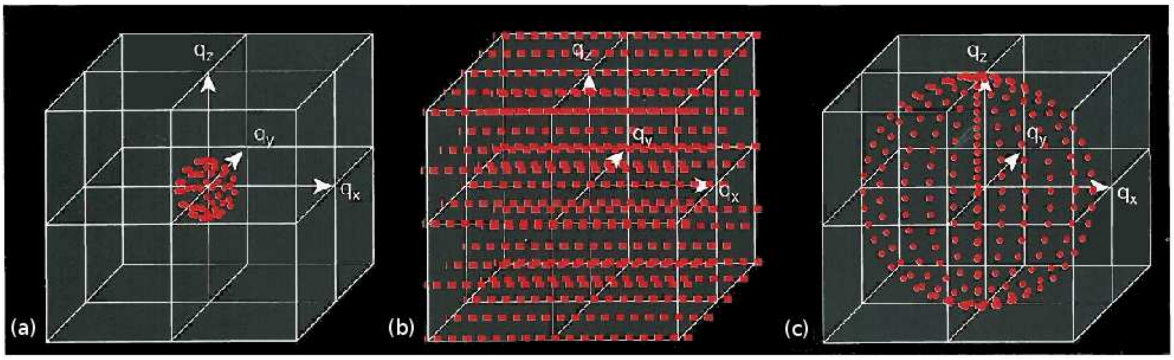


FIGURE 2.7: Three different sampling schemes for diffusion MRI. (a) DTI , (b) DSI and (c) QBI (image reproduced from [92]).

The third cause of information loss is related to the low number of gradients used to acquire the signal, which may not be sufficient to capture in detail the underlying fiber structure. For this reason, methods belonging to the **HARDI** family [121, 160], which increase the angular resolution by using a large number of gradient directions, started to be developed. Multi-model fitting algorithms [84, 170], or model-independent methods [159], benefited from specific HARDI acquisitions (such as diffusion spectrum imaging (DSI), Q-ball imaging (QBI)), which estimate the *diffusion Orientation Distribution Function* - dODF (see subsection 2.3.2), have been proposed. Moreover, methods based on spherical deconvolution *e.g.* [4, 44, 90, 157, 158], which estimate the fODF and require significantly less samples than DSI, became popular in the following years after T2 models. Figure 2.7 shows different sampling schemes and highlights the large number of gradient directions needed in DSI.

Furthermore, figure 2.8 illustrates a dODF patch stemming from Q-ball imaging (a model-free approach) on real data that contains crossing fibers.

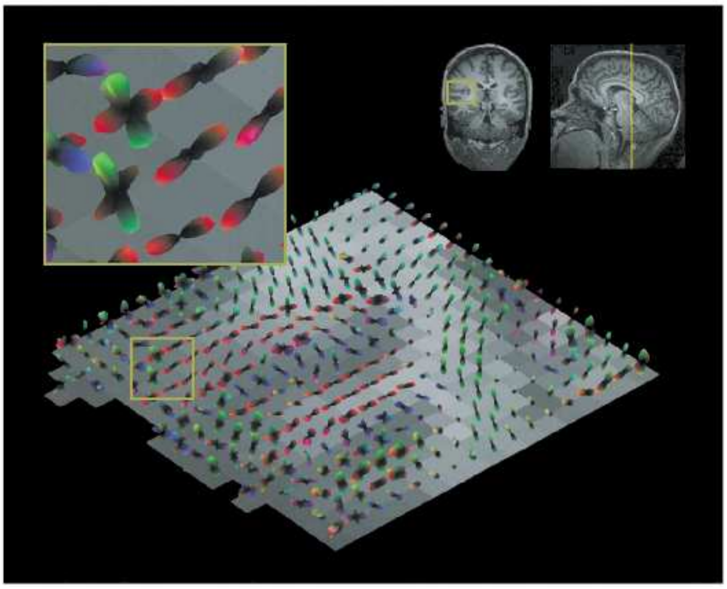


FIGURE 2.8: A patch with dODF profiles resulted from Q-Ball imaging using 492 gradient directions in an area containing crossing fibers (image appeared in [161]).

2.3.1.2 Higher Order Tensors

As previously said, it is possible to increase the number of gradient directions during DW-MRI acquisition, so that more accurate representations of the underlying fODF (or other profiles) can be obtained. Due to the fact that HARDI approaches are not based on structured models (such as tensors), numerous samples are required, resulting into long time scans.

Alternatively, higher order tensors increase the accuracy of the description. Defining a more flexible model than the limiting T2 model, by increasing the order of the tensor, looks as the ideal case. The order N of a suitable tensor must be an even number, since the diffusion $d(\mathbf{g})$ is a symmetrical function. A T2 tensor ($N = 2$) is represented by $K = 6$ unique coefficients (as was previously shown). A fourth order tensor (T4) is described by $N = 4$ and $K = 15$ (as we will justify later), and so on, $N = 6$, $K = 28$ and $N = 8$, $K = 45$ etc. (see section 2.3.3 for the computation of K). As a consequence, it is important to select wisely which model is suitable to our data, and avoid increasing the tensor's order with no control, since it will increase the number of the unknowns and probably ineffective and pointless coefficients. Examples of methods modelling diffusion or fODF profiles with HOT estimation can be found in [14, 122, 172], while for using T4 with positive definite constraints estimations the reader is referred to [11, 12, 64, 176].

According to [121], the diffusivity function of a Nth order tensor $\mathbf{D}_N$ with elements $D_{i_1 i_2 \dots i_N}$, given a $3D$ gradient direction $\mathbf{g} = (g_1, g_2, g_3)^T$ can be written as:

$$d(\mathbf{g}) = \sum_{i_1=1}^{3} \sum_{i_2=1}^{3} \dots \sum_{i_N=1}^{3} D_{i_1 i_2 \dots i_N}\ g_{i_1}\ g_{i_2}\ \dots\ g_{i_N}, \tag{2.11}$$

where $g_{i_1}, \dots, g_{i_N}$ take values from the coefficients $\{g_1, g_2, g_3\}$ of $\mathbf{g}$.

For the purposes of this dissertation, we will focus on the fourth order tensor model (T4). T4 models can represent up to three clearly separated bundles of fibers. A T4 tensor can be described by the following 6×6 symmetric matrix:

$$\mathbf{D}_4 = \begin{bmatrix} D_{1111} & D_{1122} & D_{1133} & D_{1112} & D_{1123} & D_{1113} \\ D_{2211} & D_{2222} & D_{2233} & D_{2212} & D_{2223} & D_{2213} \\ D_{3311} & D_{3322} & D_{3333} & D_{3312} & D_{3323} & D_{3313} \\ D_{1211} & D_{1222} & D_{1233} & D_{1212} & D_{1223} & D_{1213} \\ D_{2311} & D_{2322} & D_{2333} & D_{2312} & D_{2323} & D_{2313} \\ D_{1311} & D_{1322} & D_{1333} & D_{1312} & D_{1323} & D_{1313} \end{bmatrix}, \tag{2.12}$$

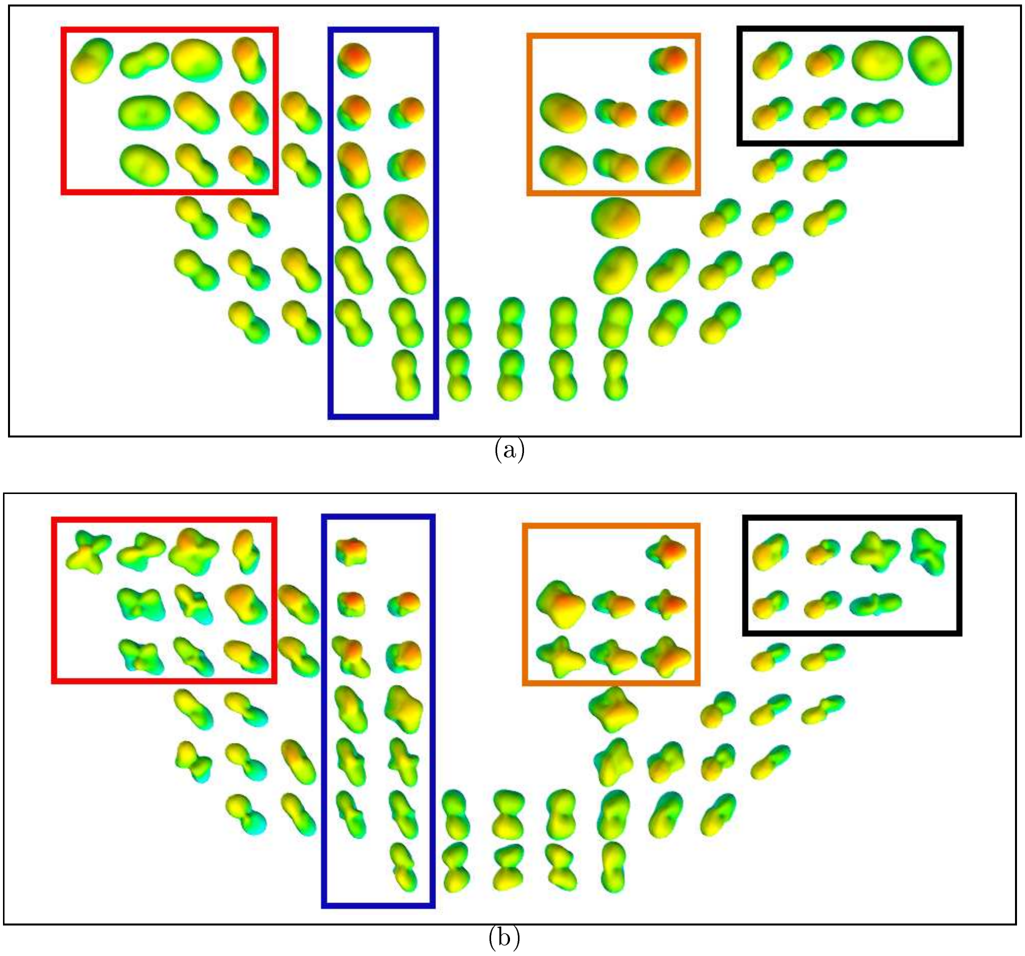
which can be compressed in a vector with 15 unique coefficients $[D_{1111}, D_{2222}, D_{3333}, D_{1122}, D_{2233}, D_{1133}, D_{1233}, D_{1123}, D_{1333}, D_{1223}, D_{1113}, D_{1112}, D_{2333}, D_{2223}, D_{1222}]^T$, since for example $D_{1122} = D_{2211} = D_{1212}$, $D_{1333} = D_{3313}$ etc. Moreover, the estimation of a T4 requires at least 15 DW-MRI measurements instead of 6 in the T2 case.

Following equation 2.11, the diffusivity function of a T4 writes:

$$d(\mathbf{g}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} D_{ijkl}\; g_i\; g_j\; g_k\; g_l. \tag{2.13}$$

It should be mentioned that the T2 model D is a particular case of the T4 model T, justified by equalizing the diffusivity functions in equations 2.7, 2.13 and considering $g_1^2+g_2^2+g_3^2=1$:

$$\left(\sum_{i=1}^{3}\sum_{j=1}^{3} D_{ij}\; g_i\; g_j\right)(g_1^2+g_2^2+g_3^2) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} T_{ijkl}\; g_i\; g_j\; g_k\; g_l. \tag{2.14}$$


(a)

(b)

Figure 2.9: Comparison between T2 and T4 models for a ROI that contains crossing fibers in the human brain. (a) T2 and (b) T4 resulting tensor fields. T4s represent the fiber structure more accurately than T2s.

Figure 2.9 shows that fourth order models are flexible and allow to capture in much more detail the underlying **fODF** in comparison to second order models. It can be noticed that when the voxels contain crossing fibers, the T2 approach is unable to capture in detail the shape of the fibers, given a solution more close to isotropic diffusion, while the T4 models better their shape. Moreover, it is mandatory to mention that a more detailed model can be much more useful to early diagnosis than less accurate techniques.

Anisotropy indices for HOTs

Interesting indices for high order tensors have been proposed in [122]. Among them, one really useful index is the Generalized Anisotropy **(GA)** which is a contrast function of the variance of the diffusion and is defined as:

$$\mathbf{GA} = 1 - \frac{1}{1 + (250V)^{\varepsilon(V)}}, \tag{2.15}$$

where $V \in [0, 1]$ is the variance of the normalized diffusivity (see [122] for the details of that index) and the exponent $\varepsilon(V)$ writes:

$$\varepsilon(V) = 1 + \frac{1}{1 + 5000V}. \tag{2.16}$$

GA's objective is to highlight areas in WM with anisotropic properties (*e.g.* fibers), similarly as **FA**, **RA** in the T2 case. High values of **GA** characterize the WM. Moreover, visualizing the **GA** as greyscale $2D/3D$ images, low contrast can be noticed between anisotropic voxels or between very low anisotropy, such as isotropic water. As a result, it is useful to separate the white matter from the grey matter. Figure 2.10 plots both **GA** and $\varepsilon(V)$ functions.

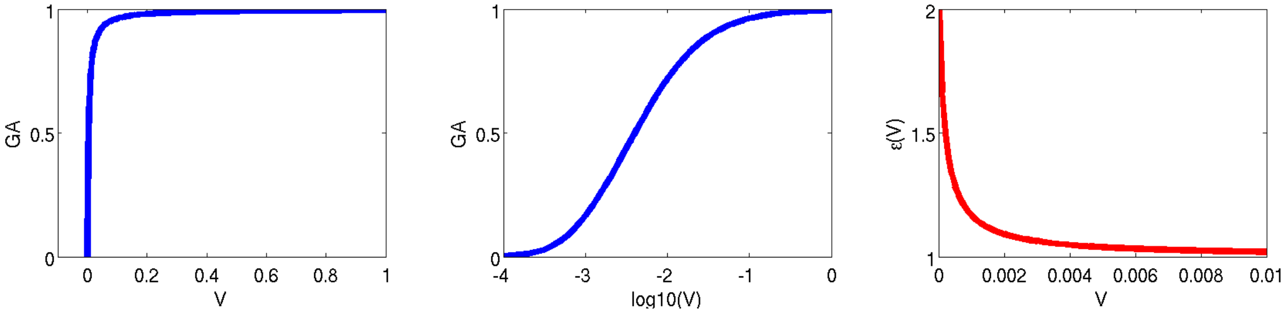


FIGURE 2.10: Plotting **GA** (on the left and the center) and $\varepsilon(V)$ (on the right).

2.3.2 ADC, dODF and fODF Profiles

In 1977, Tanner proposed to relate the measured NMR signal to a single scalar, called *apparent diffusion coefficient* (ADC), since only one gradient direction ($\mathbf{g} = \mathbf{g}^{(1)}$) was used. Replacing $d(\mathbf{g})$ in equation 2.4 with ADC, yields equation 2.17:

$$\text{ADC} = -\frac{1}{b} \ln \left(\frac{S\left(\mathbf{g}^{(1)}\right)}{S_0} \right). \tag{2.17}$$

Using more than one gradient $\mathbf{g}$ in the acquisition results into estimating whole diffusion (ADC) profiles similarly to eq. 2.17: $\text{ADC}\big(\mathbf{g}^{(i)}\big) = -(1/b) \ln \big(S\left(\mathbf{g}^{(i)}\right)/S_0\big)$.

Unfortunately, the diffusion (ADC) profile does not match with the underlying fiber orientation (see fig. 2.12). This mismatch was efficiently explained in [71] with the next example.

Given two crossing fibers with different scales f_1, f_2, modelled by two tensors $\mathbf{D^{(1)}}$, $\mathbf{D^{(2)}}$ the acquired signal can be described by the following equation:

$$\begin{aligned} S(\mathbf{g},b) &= f_1 S_1(\mathbf{g},b) + f_2 S_2(\mathbf{g},b) \\ &= \left[f_1 \exp\left\{-b\mathbf{g}^T\mathbf{D^{(1)}}\mathbf{g}\right\} + f_2 \exp\left\{-b\mathbf{g}^T\mathbf{D^{(2)}}\mathbf{g}\right\}\right] S_0 = \left[\exp\left\{-b\mathrm{ADC}(\mathbf{g})\right\}\right] S_0. \end{aligned} \tag{2.18}$$

In this case, the diffusion (ADC) profile will be given by:

$$\mathrm{ADC}(\mathbf{g}) = -\frac{1}{b} \ln\left[f_1 \exp\left\{-b\mathbf{g}^T\mathbf{D^{(1)}}\mathbf{g}\right\} + f_2 \exp\left\{-b\mathbf{g}^T\mathbf{D^{(2)}}\mathbf{g}\right\}\right]. \tag{2.19}$$

As can be noticed, the two fiber descriptions are combined non linearly in eq. 2.19, while a linear combination of the fibers would match the correct orientation (see fig. 2.12).

To bypass this misalignment, *Orientation Distribution Functions* (ODF) were proposed [47, 158, 159, 170, 172].

To be more precise, *diffusion Orientation Distribution Functions* (dODF) choose to describe the EAP propagator (eq. 2.2) as follows:

$$\mathrm{dODF}(\mathbf{g}) = \int_0^\infty p(r\mathbf{g},t)\, dr. \tag{2.20}$$

In 2004, Tuch [159] initially proposed the "model-free" concept of Q-Ball imaging, in order to approximate the dODF using acquisitions on the spherical q space (presented in figure 2.7(c)). One year later, Hess *et al.* [79] used Spherical Harmonics to define dODFs. In 2007, Descoteaux *et al.* [47] proposed a more robust Q-Ball imaging.

Recalling the previous example, the corresponding EAP propagator will be defined by the following linear equation, given the two individual propagators one for each fiber $p_1(r,t)$, $p_2(r,t)$:

$$p(r,t) = f_1 p_1(r,t) + f_2 p_2(r,t). \tag{2.21}$$

In the same direction, *fiber Orientation Distribution Functions* (fODF) (*e.g.* [4, 158, 172]) describe the acquired signal S *via* spherical convolution of the fODF profile with a response function R (modelling a single fiber) over the unit sphere (see fig. 2.11):

$$S(\theta,\phi) = fODF(\theta,\phi) \otimes R(\theta), \tag{2.22}$$

where θ and ϕ are the spherical coordinates.

Moreover, the spherical convolution is visually expressed in figure 2.11. The synthetic fODF function presented in fig. 2.11 is defined as the linear combination of two Dirac delta functions. In practice, fODF functions can appear differently than linear combinations of Dirac delta functions due to the presence of noise.

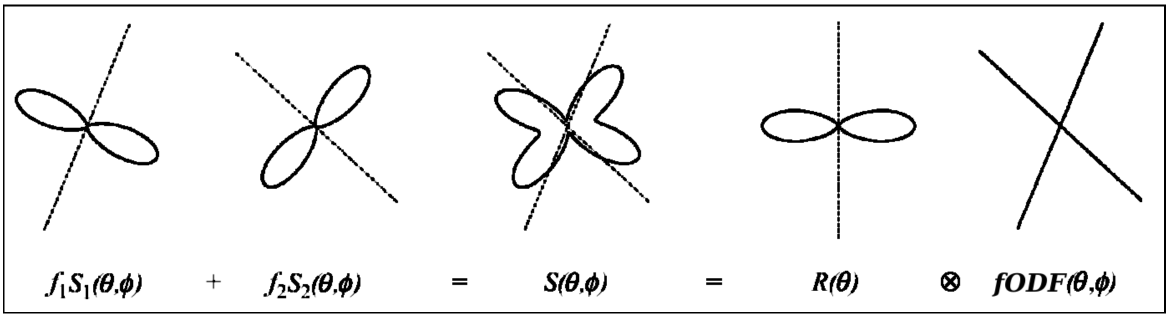


FIGURE 2.11: An explanation in $2D$ of the spherical convolution procedure used in fODF estimations (image reproduced from [158]).

In order to better understand the differences between ADC and fODF profiles, we consider the fiber structures (drawn as blue crossing lines) illustrated in figure 2.12. It is clear that the orientation of the ADC profiles (second row) has nothing in common with the underlying fiber orientation (especially when the two bundles are well separated). This happens due to the non linear combination of the fibers in the ADC profiles (eq. 2.19). On the contrary, fODF profiles (third row) describe significantly better the correct orientation by combing linearly the fibers.

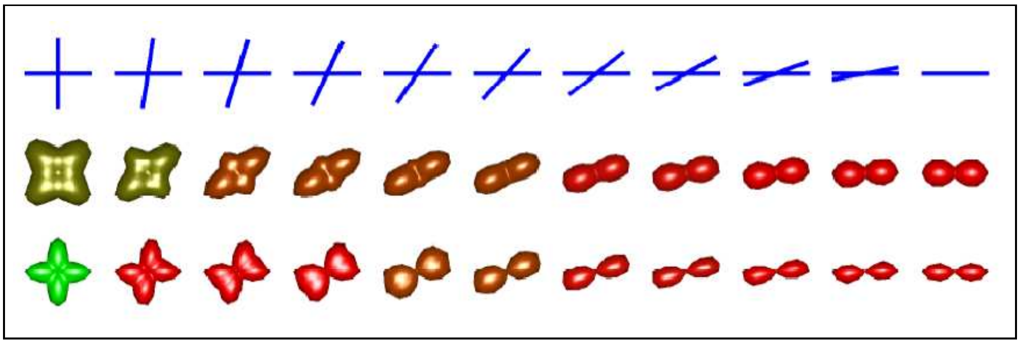


FIGURE 2.12: Comparisons of ADC and fODF profiles. The first row corresponds to the underlying fiber structure, the second row contains the ADC profiles, the third row shows the fODF profiles (original image appeared in [172]).

ADC and fODF computations using tensor models

As we previously showed, diffusion (ADC) and fODF profiles can be calculated independently of a tensor model given the limited set of gradients used in the acquisition. Alternatively, if someone has already estimated a tensor model on ADC/fODF profiles, it is possible to estimate the corresponding profile in larger sets of gradients, using the corresponding equation 2.23 for diffusion (ADC) profiles, or eq. 2.24 for fODF profiles:

$$d(\mathbf{g}) = \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\cdots\sum_{i_N=1}^{3} D^{\mathtt{diff}}_{i_1 i_2 \ldots i_N}\; g_{i_1}\; g_{i_2}\; \ldots\; g_{i_N}, \tag{2.23}$$

$$f(\mathbf{g}) = \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\cdots\sum_{i_N=1}^{3} D^{\mathtt{fODF}}_{i_1 i_2 \ldots i_N}\; g_{i_1}\; g_{i_2}\; \ldots\; g_{i_N}. \tag{2.24}$$

2.3.3 Spherical Harmonics and their Connection to Tensors

The mathematical framework of Spherical Harmonics (SH) was initially introduced by Laplace in 1782, during his investigation of Newton's law for universal gravitation. The SH of order $\ell = 0, 1, 2, \ldots$ and index $m = -\ell, \ldots, 0, \ldots, \ell$ is defined as:

$$\mathbf{Y}_\ell^m(\theta, \phi) = \sqrt{\frac{(2\ell+1)(\ell-m)!}{4\pi \quad (\ell+m)!}}\, P_\ell^m \, \cos\theta \, e^{i\, m\phi}, \tag{2.25}$$

where $\theta \in [0, \pi]$ is the polar angle, $\phi \in [0, 2\pi)$ is the azimuth angle and P_ℓ^m is the associated Legendre polynomial.

Each SH is a function on the sphere, and as a result, a set of different pairs ℓ, m can construct a set of orthonormal basis for spherical functions. Figure 2.13 exhibits a few spherical harmonics, for different ℓ, m values.

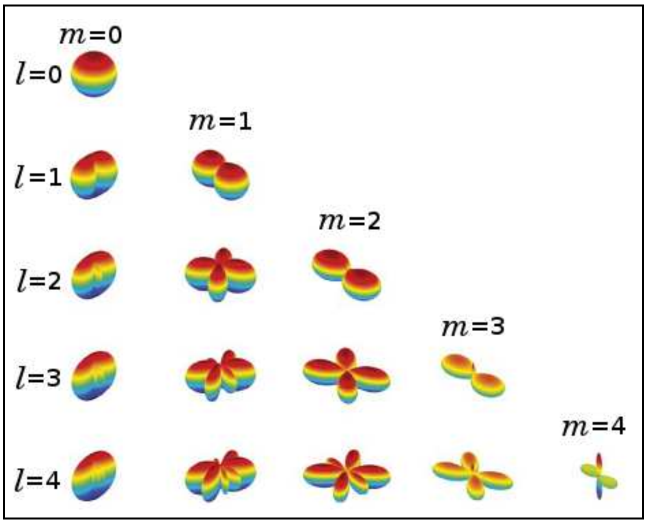


Figure 2.13: Examples of Spherical Harmonics of order ℓ and index m (image reproduced from [92]).

Given a tensor in DW-MRI that defines a symmetrical and positive real-valued function $d\,(\theta, \phi)$ (diffusion, either dODF or fODF profile) on the unit sphere, it is possible to describe $d\,(\theta, \phi)$ with the aid of K SHs:

$$d\,(\theta, \phi) = \sum_{k=1}^{K} c_k \; Y_k\,(\theta, \phi)\,, \tag{2.26}$$

where c_k are the coefficients related to $d\,(\theta, \phi)$ as they appear in the modified SH basis of [47] and

$$Y_k\,(\theta, \phi) = \begin{cases} \sqrt{2}\,\texttt{Re}\left(Y_\ell^{|m|}\,(\theta, \phi)\right) & \text{, if } m < 0, \\ Y_\ell^m\,(\theta, \phi) & \text{, if } m = 0, \\ (-1)^{m+1}\,\texttt{Im}\left(Y_\ell^m\,(\theta, \phi)\right) & \text{, if } m > 0. \end{cases} \tag{2.27}$$

In addition, the antipodally symmetric property of diffusion/dODF/fODF profiles, results into needing only even order (ℓ) spherical harmonics to describe diffusion/dODF/fODF profiles. Moreover, in DW-MRI $K = \frac{1}{2}(\ell+1)(\ell+2)$ [47, 92]. Some interesting work in the field of diffusion MRI that handles SH functions can be found in the references [47, 137, 158].

2.3.4 Tensor Metrics

Initially, Euclidean distances between tensor coefficients were considered. These distances appeared to be unsuitable to capture the precise differences of tensor data, and as a consequence Riemannian metrics were introduced. For those kind of metrics, the shortest connecting path between two points is a curve, known as geodesic curve, instead of a straight line as in an Euclidean space. The following two metrics that will be presented are defined in Riemannian spaces.

The first useful metric that we are going to need in this dissertation is called Log-Euclidean distance and was proposed by Arsigny *et al.* [6] in 2006. It is important to notice that it is only defined for second order tensors $\mathbf{D}_2^{(i)}$ as:

$$dist(\mathbf{D}_2^{(1)}, \mathbf{D}_2^{(2)}) = \left\| \log\left(\mathbf{D}_2^{(1)}\right) - \log\left(\mathbf{D}_2^{(2)}\right) \right\|, \tag{2.28}$$

where $\|.\|$ is the Frobenius norm and $\log\left(\mathbf{D}_2^{(i)}\right)$ uses the outputs of the spectral decomposition of the 3×3 matrix $\mathbf{D}_2^{(i)}$:

$$\log\left(\mathbf{D}_2^{(i)}\right) = V_i^T \log(\Lambda_i) V_i, \tag{2.29}$$

and V_i contains row-wise the eigenvectors and Λ_i is the diagonal matrix with diagonal elements the eigenvalues. Furthermore, several other metrics have been proposed for DTI and T2s in the literature [57, 99].

Inspired by the work of [153] on distances between positive functions, we define the second useful distance, between two diffusivity profiles (ADC/dODF/fODF) d_1, d_2:

$$dist(d_1, d_2) = \iint \left| \log \frac{d_1(\theta,\phi)}{d_2(\theta,\phi)} \right| \sin\theta \;\mathtt{d}\theta\; \mathtt{d}\phi, \tag{2.30}$$

where $\phi \in [0, \pi]$ is the polar angle, $\theta \in [0, 2\pi]$ is the azimuth angle that parameterize the $3D$ sphere. This distance can be used for both T2 and HOT tensors.

As can be noticed, both metrics use the logarithm. Choosing log-based distances as metrics attributes the same impact to small differences (*e.g.* 10^{-3} and 10^{-2}) and to large differences (*e.g.* 10^2 and 10^3) on the distance. A non log-based distance will be greatly affected by large differences and the contribution of small ones will be eclipsed. Such (non log-based) distances can be found in the literature *e.g.* for the fourth order tensor case [11, 15, 48, 110, 111].

Practical computation of $dist(d_1, d_2)$

The distance defined in eq. 2.30 can be approximated as the sum of N samples of the ADC/fODF/dODF profiles on the $3D$ unit hemisphere by constructing a regular grid with N_θ, N_ϕ samples in the θ and ϕ axes respectively ($\Delta\theta = 2\pi/N_\theta$, $\Delta\phi = \pi/2N_\phi$, $N = N_\theta N_\phi$):

$$dist\left(d_1, d_2\right) \simeq \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_\phi}\left|\log\frac{d_1(\theta_i,\phi_j)}{d_2(\theta_i,\phi_j)}\right|\sin\left(\theta_i\right)\Delta\theta\Delta\phi. \tag{2.31}$$

The quality of the approximation in equation 2.31 depends on the number N of samples. Due to the fact that we are interested in T4 models in this work, we performed a simulation test measuring the distances of all possible combinations between eight synthetic T4s shown in figure 2.14, in order to find the proper N that stabilizes the distances. The minimum N value derived from that simulation process (visually from the obtained distances presented in figure 2.16) is $N = 242$ samples on the unit hemisphere (see fig. 2.15).

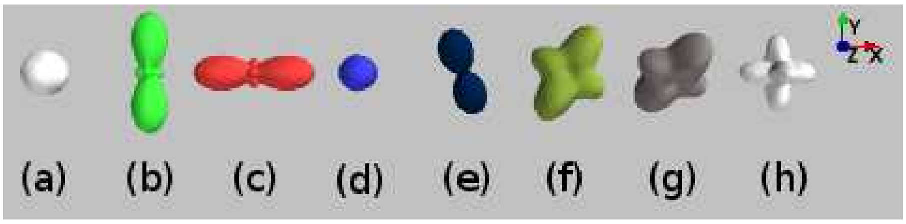


FIGURE 2.14: T4 representation of: (a) isotropic water, (b)-(d) three main fibers, while the largest diffusion occurs on y-axis, x-axis, z-axis, respectively, (e) a single fiber, (f)-(g) two crossing fibers and (h) three crossing fibers with equal diffusion in each direction. The color of the tensors highlights the orientation of the largest fiber according to the template in the top right corner of the figure.

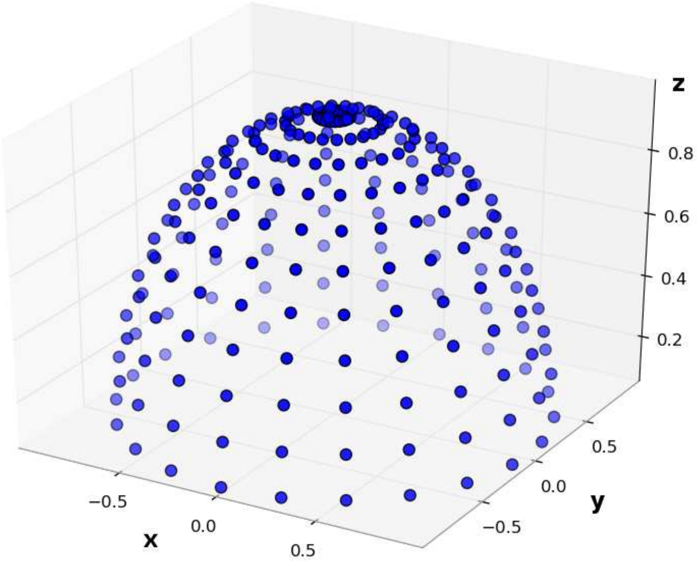
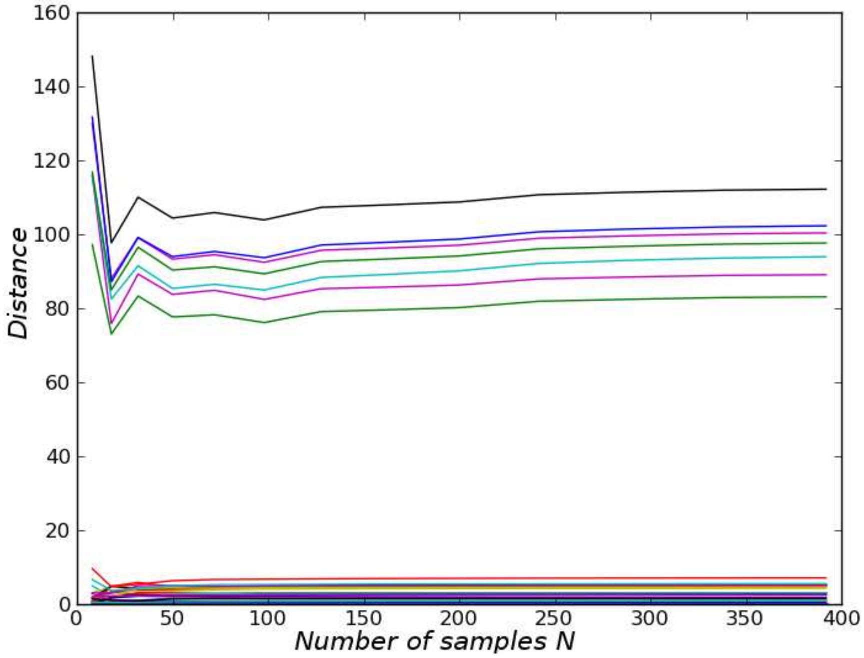
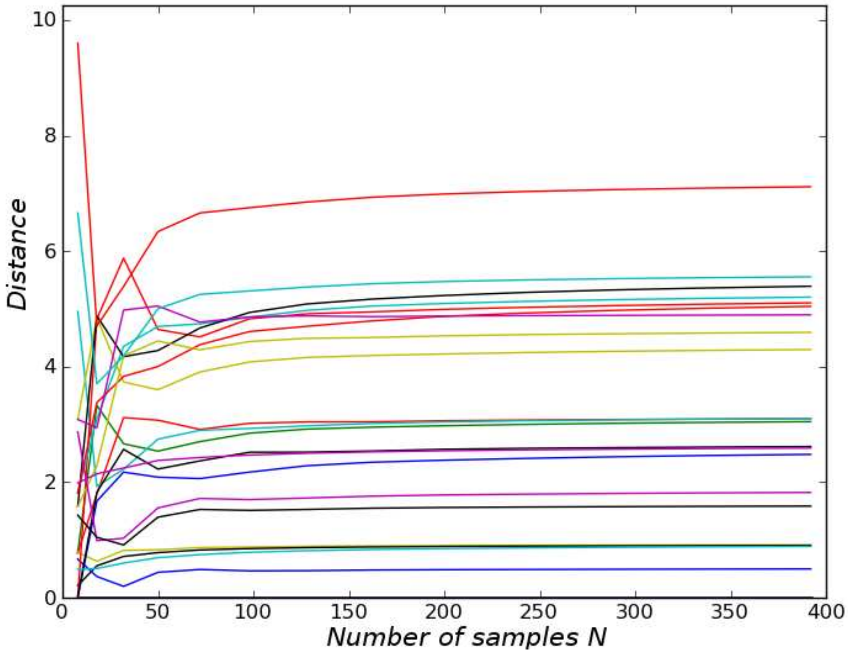


FIGURE 2.15: The figure depicts the resulting sampling scheme with 242 samples that stabilized the distances in fig.2.16.



(a)



(b)

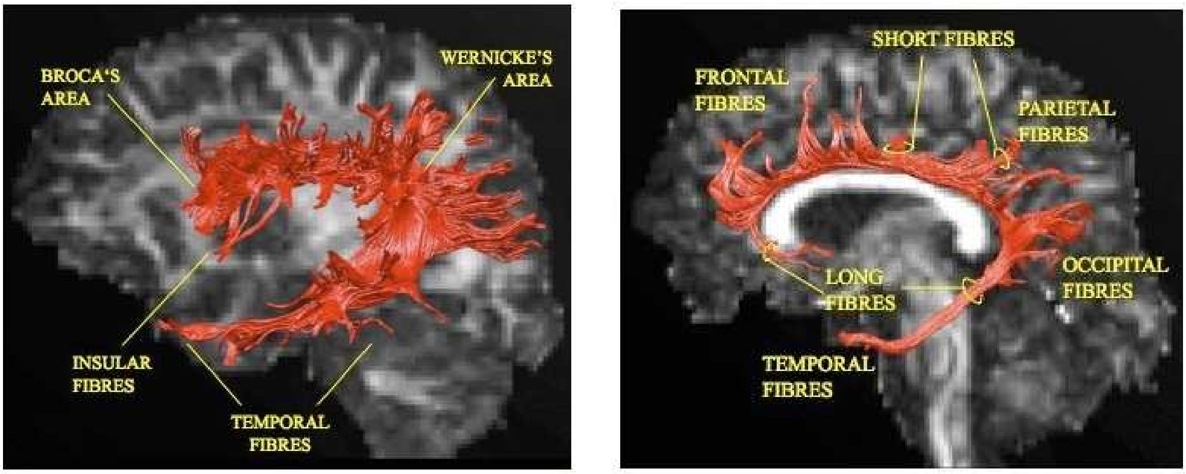
FIGURE 2.16: Simulation test in order to determine a sufficient number N of samples in the approximation of eq. 2.31. (a) Shows the distances for all pairs of combinations of T4s in figure 2.14 and (b) represents the zoom-in view of the bottom part of subfigure (a). It is noticeable that between $230 - 250$ samples in the hemisphere, the distance is stabilized.

2.4 High Level Description of DW-MRI Data

2.4.1 Fibers

Until a few years ago, mapping the connection paths between different parts of the brain had only been possible *via ex vivo* invasive techniques, *e.g.* anatomical dissection, or *in vivo* chemical tracer methods. As a consequence, non-invasive techniques to monitor and study *in vivo* brain lesions, development etc. that will affect those networks were welcome. Fitting tensor models (see section 2.3.1) to DW-MRI data permits us to approximate the underlying fiber structure and to specify the main directions of diffusion. The strategy to determine connection paths of the brain which uses information derived from tensor models is known as **tractography**. Tractography methods can be categorized into two types: (a) **deterministic** and (b) **probabilistic**. An extensive review of various white matter tractography methods can be found in [97].

Briefly, deterministic tractography exploits the information given by the principal direction of diffusion of each tensor (*i.e.* tensor major eigenvector). Figure 2.17 illustrates examples of fiber tracts, obtained using a deterministic method called "**Streamlines**" proposed by Basser in 1998 [17], of two parts of the human brain (the superior longitudinal fasciculus and the left cingulum). Moreover, figure 2.18 contains fiber tracts of one of the most extensive part of the human brain, the corpus callosum, produced by a method based on Basser's streamlines and published in 2002 [34].



(a) (b)

FIGURE 2.17: Streamlines tractographies of (a) the **superior longitudinal fasciculus** and (b) the left **cingulum** (lateral view). Images adapted from [34, 35].

On the other hand, probabilistic tractography approaches were proposed in order to bypass the weakness of deterministic tractography not being able to give information about the confidence in the fiber estimations, but also not producing all possible tracts *e.g.* due to branching. In this case, probability density functions (PDFs) of fiber orientations are

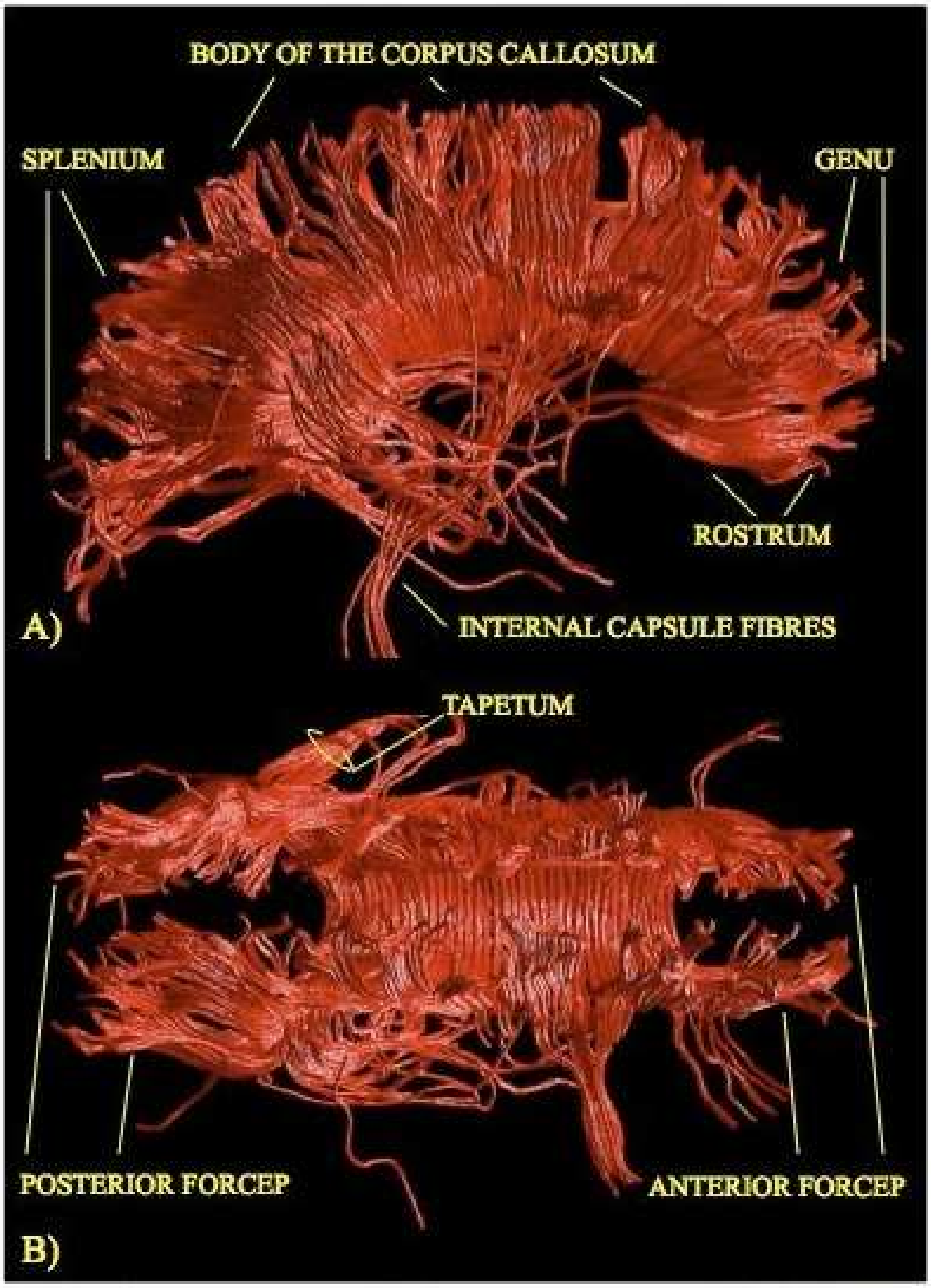


Figure 2.18: Tractography of the **corpus callosum**, a part of the brain that joins the cortex of both cerebral hemispheres. A) lateral view, B) superior view. Image appeared in [34, 35].

calculated at each voxel and several tracts can be sampled from that PDF, instead of getting a single output, such as the major eigenvector (*e.g.* in deterministic approaches). For more information about fiber PDF estimations, the reader is referred to [91, 92].

A comparison between deterministic and probabilistic tractographies is presented in figure 2.19. It is clear that the probabilistic approach yields much more trajectories than the deterministic one and besides that it is possible to know the confidence of them.

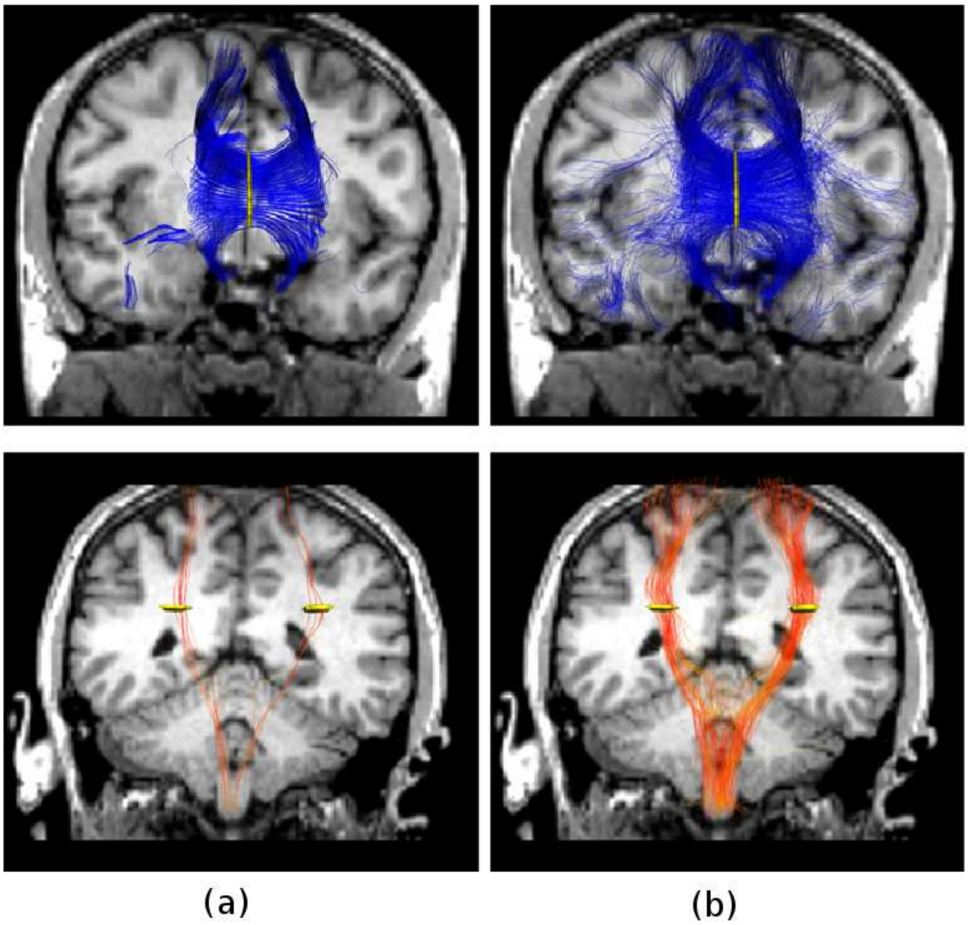


FIGURE 2.19: Comparison between deterministic and probabilistic tractographies. Tractographies of the **corpus callosum** appear in the first row and the **pyramidal fasciculus** in the second row. (a) Tracts by using deterministic approach from [113] and (b) by using probabilistic method from [130]. Images appeared in [130].

2.4.2 Connectomes

From the beginning of neuroscience, understanding the functionality and the connectivity of neural elements of the brain and identifying anatomical units has puzzled and fascinated scientists. The evolution of science and the invention of different techniques, such as fMRI, EEG etc. or even tractographies produced from previously mentioned techniques, permit us to measure the activity of neurons and to localize their connections so that interesting relational paths between them can be depicted. As a consequence, the need of a proper way to model that information came to the foreground.

The answer to the problem of representing the connectivity was found with the aid of graph theory. The nodes of the networks represent neural units, while the edges reflect the associations between neural structures. These edges are undirected and can be weighted. The weights may contain information, for example the number of fibers connecting two neural structures (*i.e.* structural properties), or the coherence of the two nodes (*i.e.* functional properties), yielding adjacency matrices after thresholding.

At this point, it is interesting to present the types of networks that we can have, as they are categorized in [148]. There are three main types of brain connectivity:

1. **Structural connectivity** is represented by a group of physical or structural connections between anatomically linked neurons. Connection scale varies from local connections of single cells to larger networks. These paths can be dynamically changed due to synaptic remodelling, development of the brain during aging or learning procedures [148]. Examples of structural connectivity algorithms of the whole human brain are given in figures 2.20, 2.21.

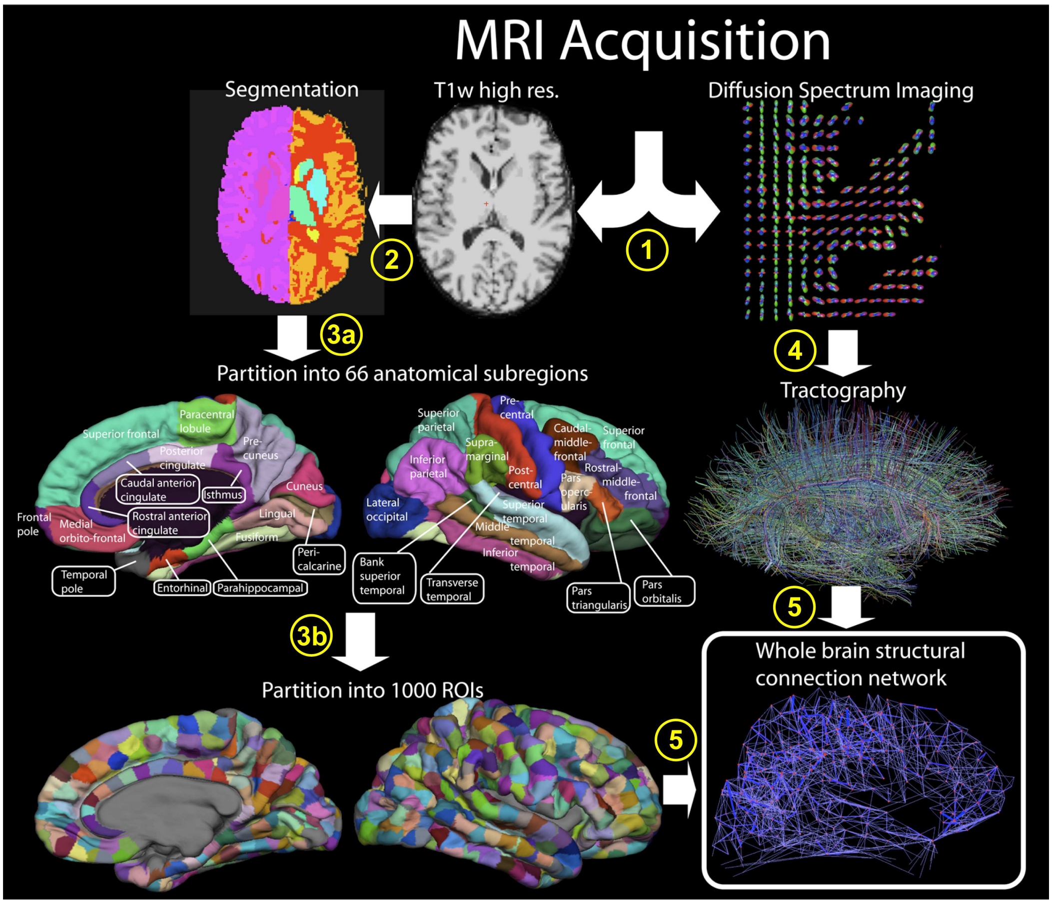


Figure 2.20: An example of the construction procedure of a whole brain structural connectivity network. (1) High-resolution T1 weighted and diffusion spectrum MRI (DSI) is acquired. DSI is represented with a zoom on the axial slice of the reconstructed diffusion map, showing an orientation distribution function at each position represented by a deformed sphere whose radius codes for diffusion intensity. Blue codes for the head-feet, red for left-right, and green for anterior-posterior orientations. (2) White and grey matter segmentation is performed from the T1-weighted image. (3a) 66 cortical regions with clear anatomical landmarks are created and then (3b) individually subdivided into small regions of interest (ROIs) resulting in 998 ROIs. (4) Whole brain tractography is performed providing an estimate of axonal trajectories across the entire white matter. (5) ROIs identified in step (3b) are combined with result of step (4) in order to compute the connection weight between each pair of ROIs. The result is a weighted network of structural connectivity across the entire brain. Figure obtained from [75].

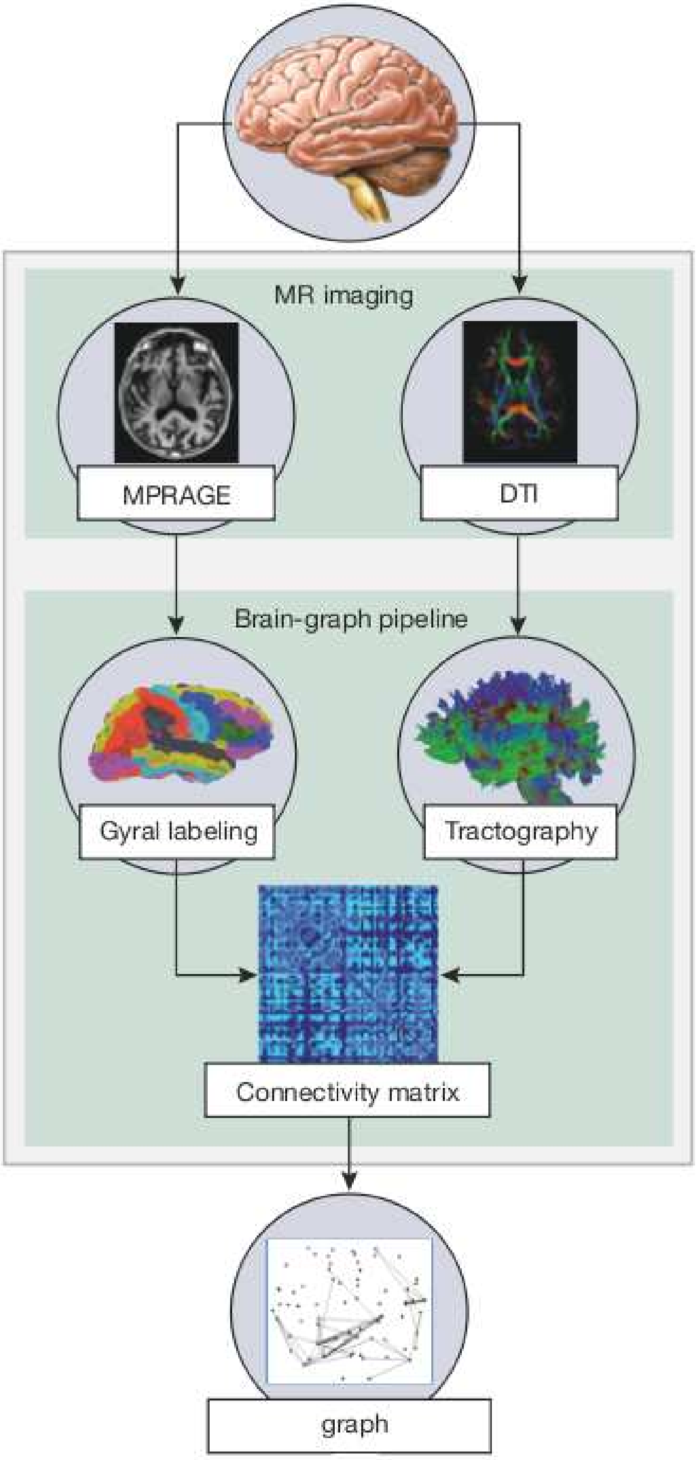


Figure 2.21: Steps of the Magnetic Resonance Connectome Automated Pipeline (MRCAP). MRCAP technique combines DTI tractography with structural MRI (MPRAGE) to construct a MR connectome. Original figure appeared in [166].

2. **Functional connectivity** draws patterns of deviations related to statistical independence between spatially remote neural structures [59, 60]. These time series data can be extracted from cellular recording techniques, fMRI, EEG, MEG, or other means. In contrast to structural connectivity, functional networks depend extremely on time and can change. In addition, the functional process can be stimulated by external factors (*e.g.* movements of the fingers, conversation with the patient etc.) that can activate different neural sensors in the brain, but also the internal state of the patient.

3. **Effective connectivity** constructs the paths of causal influences between neurons [59]. As for functional connectivity, effective connectivity is also time dependent and can be stimulated by external or internal factors.

The term "**connectome**" was originally proposed by Olaf Sporns et *al.* [148] in 2005, in order to describe the structural network of the human brain. In the same year, Hagmann [74] named as "**connectomics**" the science that studies connectome data. These terms can be used to describe any type of connectivity.

There are several obstacles that we must deal with during the construction of a human brain connectome. The complexity of the $3D$ structure of the human brain, its development and the variability of its functions, increase the number of difficulties for constructing a universal connectome of the human brain (or even particular parts of it). Moreover, invasive anatomical techniques, such as postmortem examination (which is commonly used in other species), are not always applicable, due to physical constraints of brain tissues or the absence of suitable postmortem tracing methods [148].

On the other hand, this is not the only choice. The use of non-invasive methods, such as diffusion MRI is promising. The diffusion of water molecules can reveal fiber tracts, especially in the white matter of the brain where the diffusion is anisotropic in most of the cases (in comparison to the diffusivity of the grey matter where it is closer to isotropic). Fitting diffusion models to DW-MRI data can help us to compute tractographies, which can be used to define connectomes. Figure 2.22 shows an example of the human brain connectome, resulting from the processing of diffusion MRI tractographies of the whole brain.

A simple look at figure 2.22 is enough to understand that these kind of networks are highly complex, including a large number of nodes and edges. As a result, a more compact version of them could allows us easier to derive conclusions. The calculation of **modules** (*i.e.* clusters of nodes that share more edges within the nodes of the module than with nodes of other modules) *e.g.* [36] and **hub** nodes (*i.e.* nodes that integrate a highly diverse set of signals and are capable to manage the flow of information between individual parts of the brain) *e.g.* [147] can assist in this direction. Figure 2.23 shows an example of module and hub classification.

For the explanation of abbreviations concerning the anatomical areas shown in fig. 2.22 and 2.23, the reader is referred to [75], from which the images are obtained.

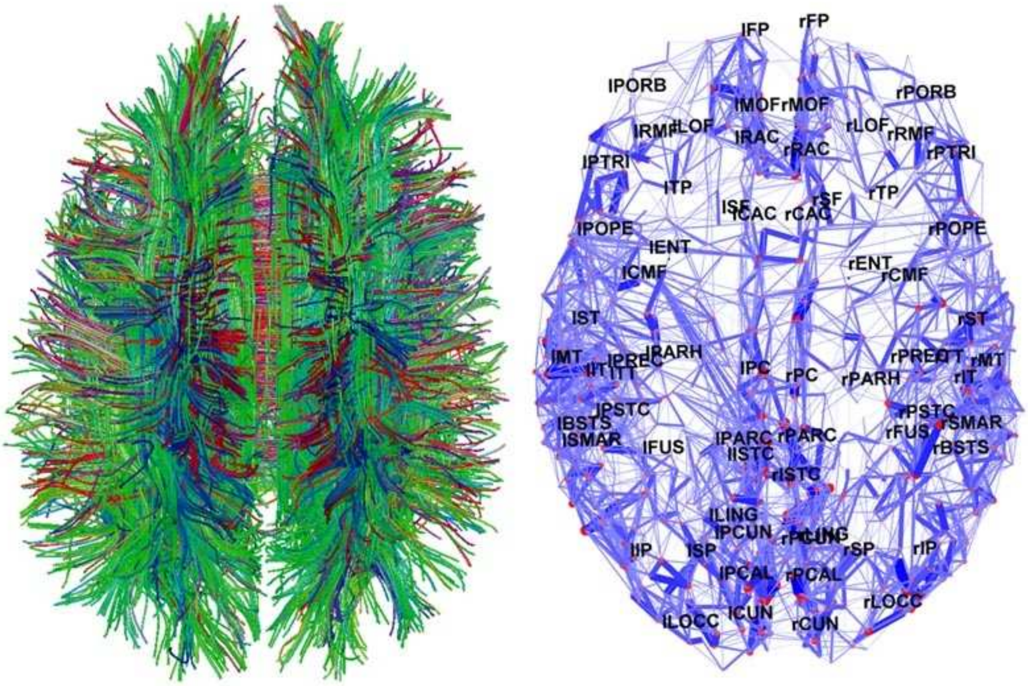


Figure 2.22: The human connectome. Images show the fiber architecture of the human brain as revealed by diffusion imaging (left) and a reconstructed structural brain network (right). Images obtained from [75]

Individual Efforts

Connectivity studies of the whole brain were performed using diffusion MRI techniques. Iturria-Medina et *al.* [87] worked on 70-90 cortical and basal grey matter areas, Gong et *al.* [69] on 78 cortical regions of the human brain and they also identified several hubs, both by using DTI methods. Hagmann et *al.* [75] combined DSI data of the whole brain, calculated by using [170], with the graph analysis of [147] allowed them to study the human cortical connectivity.

Collective Efforts

One of the largest projects whose goal was to define the anatomical and functional network of the healthy human brain started in 2009 and is called the Human Connectome Project (HCP - *http : //humanconnectome.org/*). Financed by sixteen components of the National Institutes of Health in USA, it consists of several leading institutes in the neuroscience field, such as Harvard University, Massachusetts General Hospital, Washington University in Saint Louis, the University of Minnesota and the University of California in Los Angeles. Moreover, another complementary contribution of this project can be to enlighten the study of different brain disorders that affect the connectivity paths, such as Alzheimer's disease, traumatic brain injury, stroke etc. Moreover, in 2009, the 1000 Functional Connectomes Project (FCP - *http : //fcon_1000.projects.nitrc.org/*) has been launched by leading members of the fMRI community. In 2011 the FCP idea resulted into developing the International Neuroimaging Data-sharing Initiative (INDI), initially by merging under

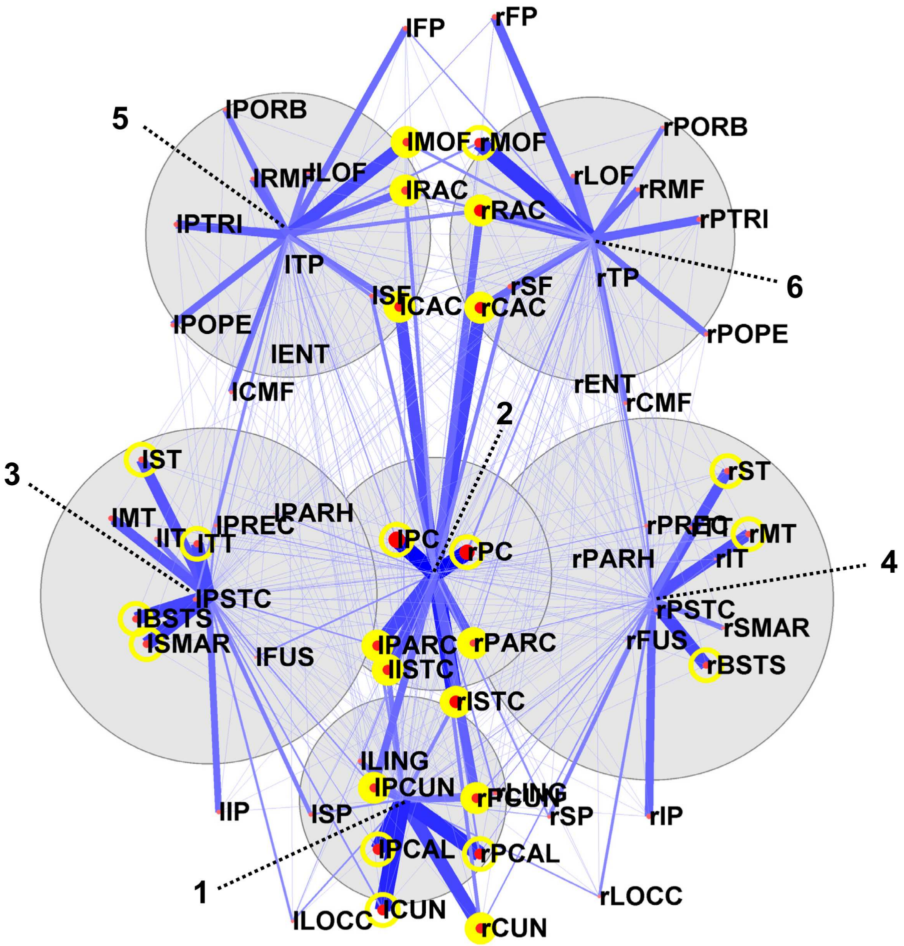


FIGURE 2.23: Modularity and hub classification. Six modules are shown as grey circles centered on their center of mass and sized according to their number of members. Edges correspond to the average connection densities of each region with the member regions of each of the six modules. Image obtained from [75]

the same name eight other individual efforts (taking place at Baylor College of Medicine, Beijing Normal University, Berlin Mind and Brain Institute, Harvard-MGH, MPI-Leipzig, NKI-Rockland, NYU Institute for Pediatric Neuroscience and the Valencia node of the Spanish Resting State Network), in order to enhance its database with global contributors, but also to establish a common sharing protocol. Furthermore, the Brain Architecture Project (*http : //brainarchitecture.org/*) aims to extract knowledge about human's, mouse's, marmoset's, zebra finch's brain architecture and structural organization.

2.5 Partial Conclusion

DW-MRI is a non-invasive technique that permits us to probe *in vivo* the diffusion of water molecules in the tissues of the human brain. It reveals the structural information of the brain. DW-MRI data correspond to raw data that should be mathematically modelled in order to perform useful calculations.

In this chapter, different kind of diffusion models were analysed, starting from DTI and its limiting T2 model, continuing to T4 models that can represent up to three clearly separated bundles of fibers and their generalization to N-order tensors. The higher the order of model, the more detailed the obtained description of the underlying fiber structure. Moreover, several measures of anisotropy per model were enumerated, and the difference between ADC, dODF and fODF profiles was highlighted. Depending on the pathology of interest, if the orientation of the fibers is changed or if only lesions on the diffusion appear, someone can choose between fODF profiles for the first case, and ADC profiles for the second case. Proper tensor metrics are useful in both cases. Finally, high levels of DW-MRI descriptions were discussed, such as fiber tracts and connectomes. Although fiber tracts are difficult to validate, non-invasive studies such as high-resolution diffusion imaging are the most promising way for mapping comprehensive structural connectivity at the macroscale. Moreover, collective efforts in constructing connectomes of the brain look very promising and auspicious.

Although the next chapters are focused on voxel-based analysis, it is clear that statistical analysis can be applied on any predefined fiber tracts and connectomes, by defining distance matrices between fiber tracts and connectomes. This aspect constitutes a remarkable perspective of the present work.

Chapter 3

Pre-processing Steps for DW-MRI Data with Emphasis on T4 Reorientation

In this chapter, the pre-processing steps of the most commonly used procedures in statistical analysis of DW-MRI data are presented. The difficulties of T4 reorientation are described and a brief overview of existing work is given in section 3.3. In section 3.4, two T4 decompositions into T2s are studied, while section 3.5 describes two T2 reorientation schemes. Section 3.6 presents our study on the fourth order tensor reorientation scheme proposed in thesis [134] and section 3.7 contains the experimental results. Finally, section 3.8 is the partial conclusion of the chapter.

3.1 Pre-processing the Raw DW-MRI Data

A few standard pre-processing steps are usually required for each datum (in normal and pathological groups), before the computation of the statistics:

- **Eddy Current Correction** is crucial due to the presence of eddy currents (also known as Foucault currents [55]) in the gradient coils that generate stretches and shears in the reconstructed volumes (*i.e.* DW-MRI data) during the data acquisition process. These distortions are different for each gradient direction. Head movements and blood pulsation are also corrected, by using affine registration to a reference volume (usually the DW-MRI image that corresponds to $b = 0$).

- **Brain Extraction** aids us to approximate the borders of the brain by excluding uninformative areas such as the skull, the eyes, the nasal and oral cavities. It will help us to gain processing time since we are interested only in the voxels of the brain, but also since the template images usually do not include that information.

- Calculating **FA images** is a usual step in the construction of an atlas, since they will be used in the estimation of the transformation between the initial patient space and the atlas space of the common template image which often corresponds to a FA image. Of course, other templates (*e.g.* DW images) can be used too.

- The estimation of a **linear/non-linear transformation** between the estimated patient's FA image and the template FA (*e.g.* MNI152, JHU-ICBM common space [114, 115]) is mandatory, in order to transform (register and reorientate *via* **data normalization**, see next section 3.2) each DW-MRI data to a common space. Typically, a non-linear transformation is used to gain accuracy, and is initialized by a linear one. A more detailed description of the data normalization problem will be presented in section 3.2.

- **Reduction of the registration error** is usually performed by smoothing the data, for example with a $3D$ Gaussian filter. Since in this case we risk to lose important information due to over-smoothing, we propose to deal with any registration errors left at the time of comparing two individuals, by searching for the best alignment between two $3D$ patches (one for each individual) in the wide extended neighborhood of the current voxel. Moreover, neighboring information can be introduced (for more details, the reader is referred to the next section 5.1.2).

3.2 Data Normalization

It is known that the anatomical structure of the human brain varies between different patients [156]. In addition, the relative position of the brain between different acquisitions can also be different. Therefore, data normalization is a crucial and mandatory step for atlas construction and population comparisons. All individuals must be aligned to a common space (*e.g.* template), usually called **atlas space**.

First of all, it is important to note that the term "*data normalization*" does not refer only to the spatial **registration** of the data, but also to the **reorientation** of the diffusion directions in order to address properly the new underlying fiber orientation, altered by the spatial registration.

Secondly, we should continue by answering the following question: "*What kind of data should be normalized?*". Since no standard guideline to follow exists in literature, most of the approaches are categorized into two kind of strategies. The first strategy registers spatially the raw DW-MRI images, and reorientates the gradient directions of the magnetic field (*i.e.* b-vectors) in order to fit tensor models on the normalized data. Normalization methods belonging to this category are, for example, [51, 152] for linear and non-linear transformations, respectively, where the b-vectors are reorientated using the rotation part of the transformation. Moreover, Hong *et al.* [82] proposed a spatial normalization of fODF for HARDI data, where the transformation is applied on the $3D$ sampling vectors of the fODF function. On the other hand, the second strategy firstly estimates a tensor model from the data and consequently normalizes the tensor images, by registering serially all tensor coefficients (considering each coefficient as a separate image), and reorientates afterwards the embedded registered tensor models. Examples of methods belonging to the second category are [2, 13, 15, 73, 135]. For more information about the second category the reader is referred to section 3.3.

3.3 Introduction to Tensor Reorientation

As was previously mentioned, applying a spatial transformation T to the tensor images, for example, that will convert the tensor field of figure 3.14(a) to 3.14(b), will result into incoherence between the main directions of the tensors 3.17(b) and the underlying fODF 3.17(a) in the new space. This phenomenon occurs due to the fact that tensors contain directional information concerning the diffusion. As a result, actions should be made to fix all impending misalignments.

Barmpoutis *et al.* proposed initially a reorientation method for T4 models in [15] and then they generalized the idea to HOTs in [13], in order to apply an estimated rotation matrix to the tensor models. Their approach is limited to linear transformations, when the whole information can be captured by an affine matrix. Unfortunately, unlike T2 models, a linear transformation that contains stretching or shearing effects is not directly applicable to T4 models on account of two fundamental reasons. Firstly, due to the fact that more than one main direction can be described by the T4 model, and secondly, each direction will be differently affected by the transformation. In that sense, Barmpoutis *et al.* fail to reorientate separately each main direction.

On the other hand, since gold standard methods for T2 reorientation exist in the literature, *e.g.* [2], HOT reorientation strategies that decompose the HOT model to several T2 components, such as [15, 20], started to appear [73, 135]. These approaches assume that each of the obtained T2s will be aligned with one fiber of the HOT model. As a consequence, the application of the transformation can be transferred to the level of each T2. Inspired by the last concept, we will focus on T4 decompositions into T2s, as an intermediate step of T4 reorientation. The following two sections give us more details concerning the decompositions proposed in [15, 20] and the T2 reorientation scheme of [2], that will help us to define Renard's method [134] to reorientate T4s.

3.4 T4 Decomposition Schemes

In this section, T4 decompositions into T2s will be presented. Firstly, the *Spectral Decomposition* (SD) (subsection 3.4.1) proposed in [20] and secondly the *Hilbert Decomposition* (HD) (subsection 3.4.2) [15] will be described, both studied also by Renard *et al.* in [135], in order to apply the transformation on the obtained T2s.

3.4.1 Spectral Decomposition

Basser and Pajevic proposed in [20] to decompose each fourth order tensor D, written as a symmetric 6×6 matrix, with the aid of eigenanalysis into six eigenvalues μ_i and six eigenvectors D_2^i, which correspond to second order tensors without any constraint of either positivity or rank:

$$D = \sum_{i=1}^{6} \mu_i D_2^i D_2^{iT}. \tag{3.1}$$

The advantage of SD decomposition is that the obtained solution is unique. On the other hand, six T2s representing up to six bundles of fibers (in contrast to three described by the T4 model) are more than we need, expecting that the remaining three eigenvectors will correspond to zero eigenvalues. Experiments showed that it is possible to obtain eigenvalues close to zero, or even negative, depicting T2s as crosses (*e.g.* second column and third row in fig. 3.1) due to the presence of noise, without any physical meanings in diffusion MRI, which constitutes an important drawback.

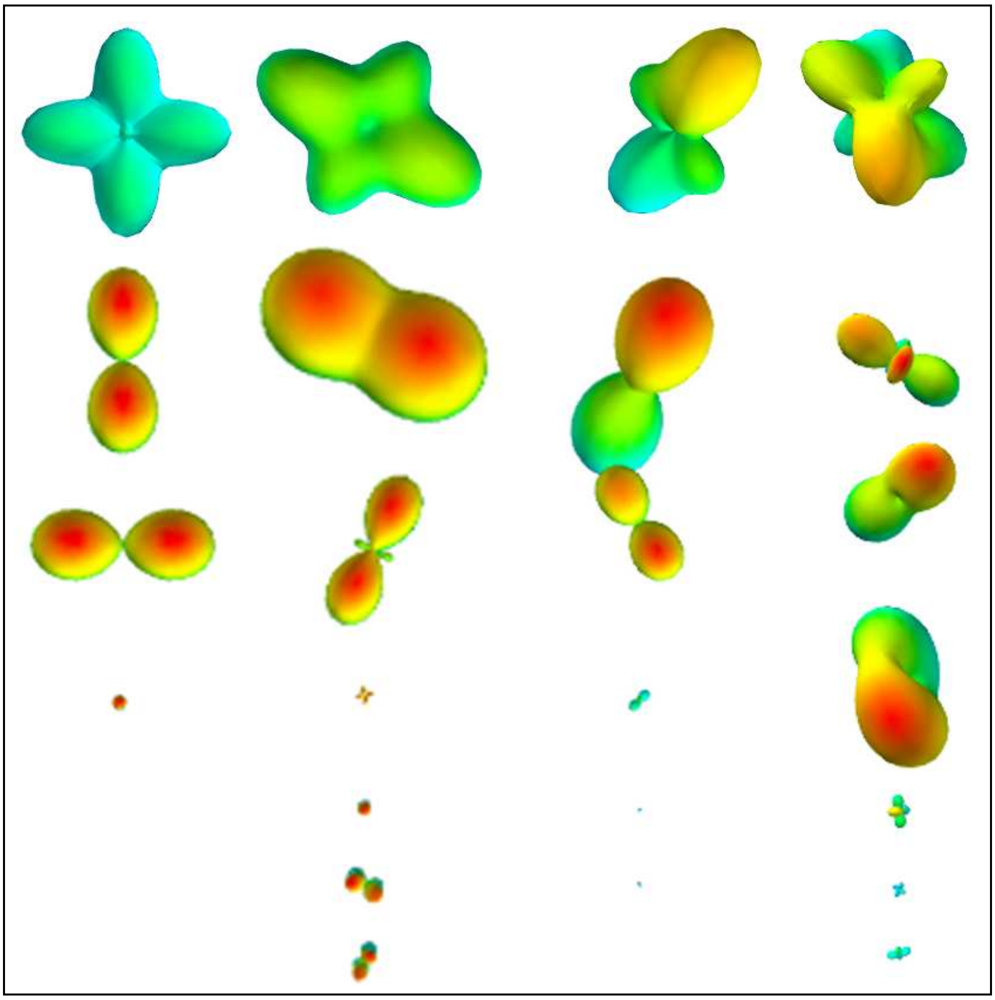


FIGURE 3.1: Four T4 **SD** decompositions. The first two are synthetic and the last two are selected from real data. The first line displays the T4 representation and the next six lines contain the six T2s produced by the **SD** decomposition.

3.4.2 Hilbert Decomposition

Barmpoutis *et al.* applied Hilbert's theorem (**Theorem 1.**) on the diffusivity function $d(\mathbf{g})$ of a fourth order tensor in [12].

The diffusivity function $d(\mathbf{g})$ of a T4 D can be written as:

$$d(\mathbf{g}) = \sum_{i,j,k,l=1}^{3} D_{ijkl}\; g_i\; g_j\; g_k\; g_l, \tag{3.2}$$

where $\mathbf{g} = [g_1, g_2, g_3]$ is a vector of the $3D$ unit sphere.

Theorem 1. (Hilbert's theorem) *Every real positive ternary quartic function d can be expressed as a sum of three squared quadratic forms*

$$d(\mathbf{g}) = \left(\mathbf{v}^T\mathbf{q}_1\right)^2 + \left(\mathbf{v}^T\mathbf{q}_2\right)^2 + \left(\mathbf{v}^T\mathbf{q}_3\right)^2 = \mathbf{v}^T\mathbf{Q}\mathbf{Q}^T\mathbf{v}, \tag{3.3}$$

where $\mathbf{v}$ is a vector of monomials $[g_1^2, g_2^2, g_3^2, g_1g_2, g_1g_3, g_2g_3]$, $\mathbf{q}_i$ are 6×1 vectors containing the coefficients of the ith quadratic function that corresponds to a second order tensor and $\mathbf{Q}$ corresponds to a 6×3 summary matrix containing the three $\mathbf{q}_i$'s.

Although the solution consists of three T2s (equal to the number of fibers in T4), Ghosh *et al.* [63] notice the non-unicity of this decomposition. As we can see in Eq. 3.4, we can obtain one solution $\mathbf{QR}$ for each 3×3 rotation matrix $\mathbf{R}$.

$$\mathbf{v}^T\mathbf{G}\mathbf{v} = \mathbf{v}^T\mathbf{Q}\mathbf{Q}^T\mathbf{v} = \mathbf{v}^T\mathbf{Q}\mathbf{R}\mathbf{R}^T\mathbf{Q}^T\mathbf{v} = \mathbf{v}^T\mathbf{Q}'\mathbf{Q}'^T\mathbf{v}. \tag{3.4}$$

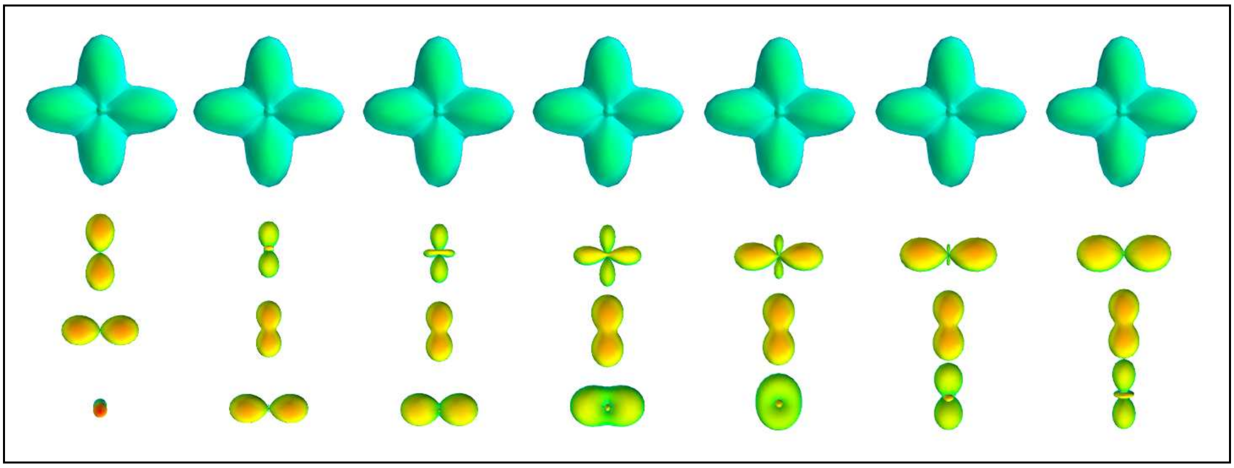


FIGURE 3.2: Different rotation matrices $\mathbf{R}$ can recompose the same fourth order tensor in **HD**. The first row shows the T4 representation. The three last rows display the three resulting T2s corresponding each to a different $\mathbf{R}$. The first column corresponds to the $\hat{\mathbf{R}}$ that minimizes L (defined by the method, see section 3.6.

In fact, another reason that clarifies the right choice of the rotation matrix $\mathbf{R}$ is the possibility that the resulting T2s in $\mathbf{QR}$ may not match in shape with the individual fibers in the T4 representation, showing us the requirement of a T2 rotation. Additionally, due to the absence of positivity constraints on T2s, they can be negative (*i.e.* related to negative eigenvalues), which someone can apparently assume as a drawback. Negative T2s are represented as crosses without any physical meaning (fig. 3.2) and T2 reorientation, as the PPD (a state-of-art method that will be described later), is not applicable in the presence of negative T2s. On the other hand, the effect of negative T2s is eliminated since the obtained signal is always positive considering the square of T2 in $\mathbf{QQ}^T$. Moreover, the maximum number of T2s is equal to the maximum number of fibers that the T4 model can represent. Furthermore, if someone desires to eliminate the appearance of crossing (negative) T2s in the decomposition (at a minimum possible level), the definition of an optimization problem with a proper cost function L is feasible. The minimization of L with respect to $\mathbf{R}$ will reduce that phenomenon. For example, in figure 3.2 the first column is related to the correct rotation matrix $\mathbf{R}$.

3.5 T2 Reorientation Schemes

It is known that every transformation T such that $\mathbf{x}' = T(\mathbf{x})$ can be locally expressed by an affine matrix $\mathbf{F}$. If T is affine or rigid (represented by a matrix $\mathbf{A}$), then matrix $\mathbf{F}$ does not depend on the position $\mathbf{x}$ ($\mathbf{F} = \mathbf{A}$). Otherwise, if T is non linear, a matrix $\mathbf{F(x)}$ can be defined at each point $\mathbf{x}$ as Eq. 3.5 shows:

$$\mathbf{F}(\mathbf{x}) = \mathbf{I_{3x3}} + \mathbf{J_u}(\mathbf{x}), \tag{3.5}$$

where $\mathbf{I_{3x3}}$ is the 3x3 identity matrix and $\mathbf{J_u}(\mathbf{x})$ is the Jacobian matrix of the vectorial field $\mathbf{u}= [u_x, u_y, u_z]^T$ at point $\mathbf{x}$ according to [2].

After all the desirable second order tensors were obtained, reorientation of each T2 is needed. For this case, two methods proposed by Alexander *et al.* [2] are presented. The first one is called *Finite Strain* (subsection 3.5.1) and the second one *Preservation of Principal Directions* (subsection 3.5.2).

3.5.1 Finite Strain

The Finite Strain (FS) algorithm belongs to the theory of continuum mechanics for distortion. In this approach, the original transformation T is approximated by a rotation matrix $\mathbf{R}$ that is extracted from the affine transformation $\mathbf{F}$ by using the polar decomposition [80] as follows:

$$\mathbf{F} = \mathbf{RS} = \mathbf{LMN}^T = \left(\mathbf{LN}^T\right)\left(\mathbf{NMN}^T\right), \tag{3.6}$$

where $\mathbf{LMN}^T$ is the singular value decomposition (SVD) of matrix $\mathbf{F}$, $\mathbf{LN}^T$ is an orthogonal matrix and $\mathbf{NMN}^T$ is a symmetric positive definite matrix. By identification, the desirable rotation matrix is $\mathbf{R} = \mathbf{LN}^T$.

Obviously, **FS** method carries a weakness, since information can be lost if a rotation matrix is extracted, for example, when the transformation contains shearing effects, leading to significant limitations and errors [73], [135].

3.5.2 Preservation of Principal Directions

To bypass the limitations of the **FS** scheme, Alexander *et al.* [2] proposed the method of Preservation of Principal Directions (**PPD**). The main idea of this approach is to apply the whole transformation **F** on each **positive** T2 and then normalize them to keep their initial form. Equations 3.7, 3.8 and 3.9 are the summary of the **PPD** steps.

$$\mathbf{n}_1 = \frac{\mathbf{F}\mathbf{e}_1}{\|\mathbf{F}\mathbf{e}_1\|}, \tag{3.7}$$

$$\mathbf{n}_2 = \frac{\mathbf{F}\mathbf{e}_2 - \left(\mathbf{n}_1^T\mathbf{F}\mathbf{e}_2\right)\mathbf{n}_1}{\|\mathbf{F}\mathbf{e}_2 - \left(\mathbf{n}_1^T\mathbf{F}\mathbf{e}_2\right)\mathbf{n}_1\|}, \tag{3.8}$$

$$\mathbf{n}_3 = \mathbf{n}_1 \times \mathbf{n}_2, \tag{3.9}$$

where $\mathbf{e}_i$ are the eigenvectors of each T2, sorted in ascending order according to their eigenvalues. Lastly, a general remark is that **PPD** is assumed to be the gold standard method for DTI reorientation.

3.6 T4 Reorientation Scheme based on HD and PPD

There are three main steps for the T4 reorientation scheme presented in [134]. Firstly, equation 3.3 is solved with respect to **Q** by using the least squares method. As we have shown, **Q** is not unique. Theoretically, the optimal solution of **R** is derived by solving an optimization problem parametrized by the three angles of the possible directions of rotation. Instead of that, in order to deal with it without losing significant time in solving another time-consuming optimization problem, a set of 1000 randomly constructed rotation matrices **R** is defined, in the second step of the method. For each matrix **R** we calculate L using eq. 3.10 and the $\hat{\mathbf{R}}$ which corresponds to the minimum L is retained.

$$L = min\left(\frac{\lambda^+}{\lambda_{tot}}, \frac{\lambda^-}{\lambda_{tot}}\right), L \in [0, 0.5], \tag{3.10}$$

where $\lambda_{tot} = (\lambda^+ + \lambda^-)$, λ^+, λ^- are the absolute values of the sum of the positive and negative eigenvalues of the T2s that **QR** contains, respectively. In this way, we force to minimize the problematic case of appearing negative and positive eigenvalues, both at the same matrix **QR**.

Both **SD** and the methods based on Hilbert's theorem, without imposing any positivity constraints to the resulting T2s, can give tensors categorized into three cases: a) all positive

T2s, b) all negative T2s and c) both positive and negative T2s. In the last two cases a pre-processing step must be done. It is important to clarify that the mix of both negative and positive T2s is the problematic case. A connection between the previously mentioned cases and the L value is the following: $L = 0$ is the case of having strictly positive or negative tensors. The greater L we get, the more close to crosses the T2s look like (case (c), please refer to fig. 3.2 for some examples).

From a mathematical point of view, this means that both positive and negative eigenvalues are obtained. Experiments showed that the last scenario occurs quite often, but the usage of eq. 3.10 reduces its occurrence, significantly (but in some cases it is not totally eliminated), so that the impact of the remaining negative eigenvalues in the decomposition is small [134].

When the final solution $\hat{\mathbf{Q}} = \mathbf{Q}\hat{\mathbf{R}}$ of the T2s is obtained, if the T2s are positive, then *PPD* reorientation (subsection 3.5.2) is applied to each second order tensor. In the presence of negative T2s, the absolute value of the eigenvalues of each T2 is used in the PPD reorientation to set the order of the principal directions.

Finally, the 15 coefficients of the reorientated T4 are extracted from the reorientated matrix $\mathbf{G} = \hat{\mathbf{Q}}_{\mathbf{reo}} \hat{\mathbf{Q}}_{\mathbf{reo}}^T$ as shown in Eq. 3.11 according to [12]. Values a,b,c,d,e,f correspond to free parameters depending on the resulting matrix $\mathbf{G}$ and they do not affect the 15 coefficients of the tensor.

$$\mathbf{G} = \begin{pmatrix} D_{1111} & \alpha & b & \frac{1}{2}D_{1112} & \frac{1}{2}D_{1113} & d \\ \alpha & D_{2222} & c & \frac{1}{2}D_{1222} & e & \frac{1}{2}D_{2223} \\ b & c & D_{3333} & f & \frac{1}{2}D_{1333} & \frac{1}{2}D_{2333} \\ \frac{1}{2}D_{1112} & \frac{1}{2}D_{1222} & f & D_{1122} - 2\alpha & \frac{1}{2}D_{1123} - d & \frac{1}{2}D_{1223} - e \\ \frac{1}{2}D_{1113} & e & \frac{1}{2}D_{1113} & \frac{1}{2}D_{1123} - d & D_{2b} & \frac{1}{2}D_{1233} - f \\ d & \frac{1}{2}D_{2223} & \frac{1}{2}D_{2333} & \frac{1}{2}D_{1223} - e & \frac{1}{2}D_{1233} - f & D_{2233} - 2c \end{pmatrix}. \tag{3.11}$$

3.7 Experimental Results

Renard's T4 reorientation method [134] (*i.e.* **HD+PPD**) was lacking of experimental results on real data. In this dissertation, a further study and analysis of both synthetic and real data was performed. Subsection 3.7.1 contains the synthetic data case and describes the evaluation scheme that is proposed and was used in this thesis, while subsection 3.7.2 corresponds to the real data case and the corresponding proposed evaluation schemes.

3.7.1 Synthetic Data

Synthetic Example 1

Figure 3.4 shows the first example of a sine transformation applied on the x parameter of a 30×30 template with T4s that represents two main bundles of fibers (one vertical and one horizontal), crossing each other in the center of the image. Figure 3.4(b) shows that after applying a spatial transformation on a tensor field, it is always needed to reorientate the embedded tensors in order to match properly with the underlying fiber orientation.

Figures 3.5- 3.6 contain the resulting reorientated tensor fields for each tested technique. As shown, **FS** reorientation fails to apply the transformation correctly, since the extraction of the rotation transformation that **F** contains produces significant errors due to loss of information. On the other hand, **SD** and Renard's method with **HD**, both using **PPD**, manage to apply the whole transformation **F** by producing notably less error than **FS**.

At this point, we must define an evaluation scheme in order to compare our results (see fig. 3.3). **D** in figure 3.3 corresponds to the estimated reorientated main directions, constructed by applying transformation matrix **F** to the extracted main directions ($3D$ vectors) of the fODF function of the T4 (**B**), with the aid of the local maxima function of *Dipy* library [62], which contains several interesting tools for analysis of diffusion-MRI data. **D** will be assumed as the ground truth (GT) solution. To continue, **D'** will be the resulting normalized tensor by registering spatially the tensor coefficients and then reorientating the registered T4s with one of the three compared methods.

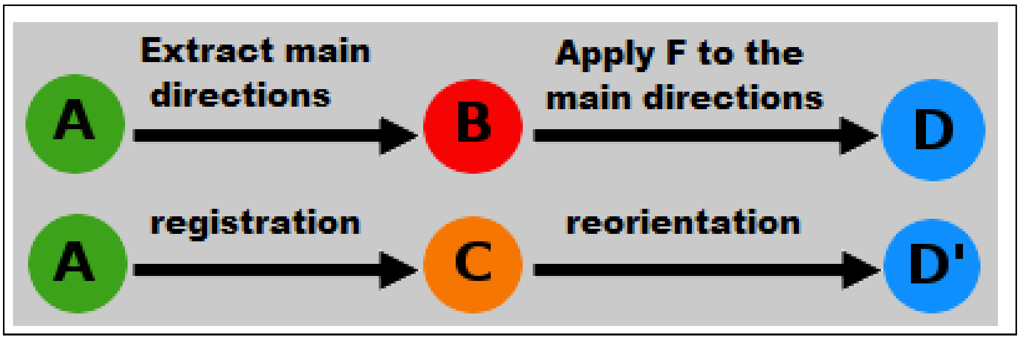
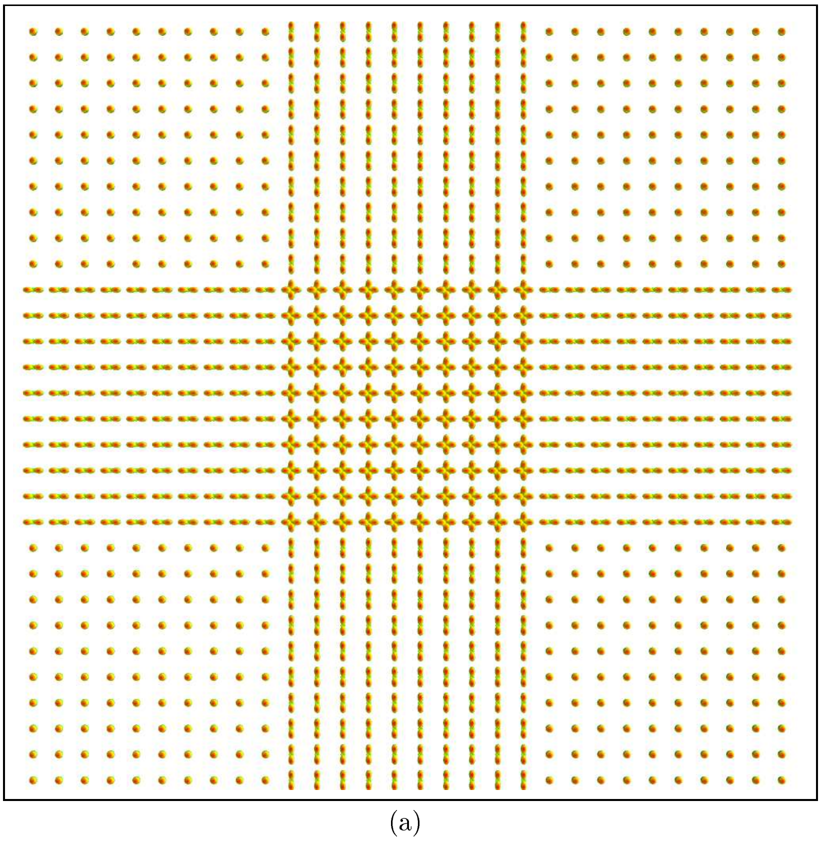


FIGURE 3.3: Evaluation scheme for synthetic data: measure the angular error (AE) between **D** and **D'**.

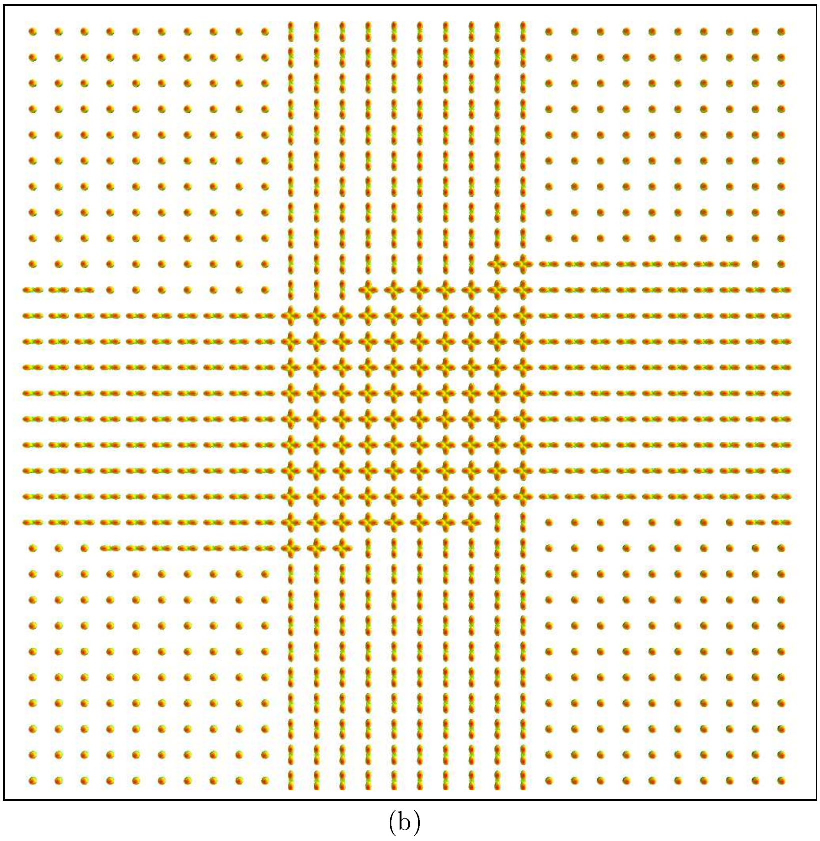
In this way, the angular error can be measured between peaks of **D** and **D'** for both horizontal and vertical bundle of fibers.

In detail, figure 3.7 shows the ground truth (GT) and the initial peaks, while figures 3.8, 3.9 contain the extracted peaks produced from each reorientation technique. Figure 3.11 shows the horizontal angular errors, while figure 3.12 presents their corresponding histograms. Moreover, figure 3.13 shows the resulting vertical angular error in the **FS** case.

As shown in table 3.1, the **HD+PPD** method gives results very close to **SD+PPD**, while **FS** produces significant errors, especially in the vertical bundle of fibers, where initially there was no angular error. This happened because **FS** rotates the whole tensor and not each main direction separately as **SD+PPD** and **HD+PPD** do.

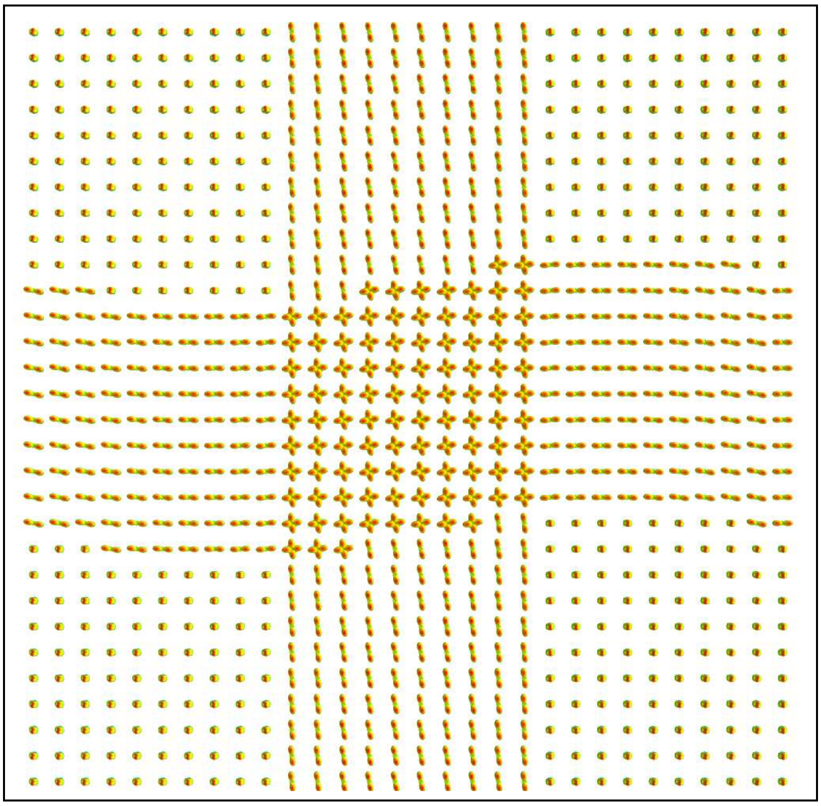


(a)

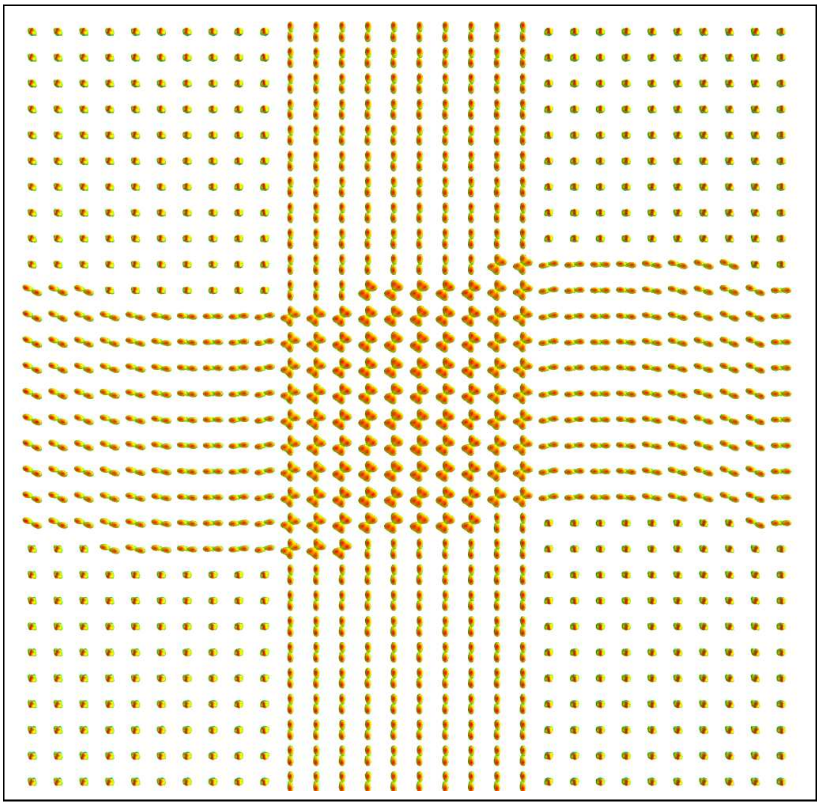


(b)

FIGURE 3.4: First example: synthetic tensor fields. 3.4(a) The template of the tensor field and 3.4(b) the initial registered template (no reorientation yet).

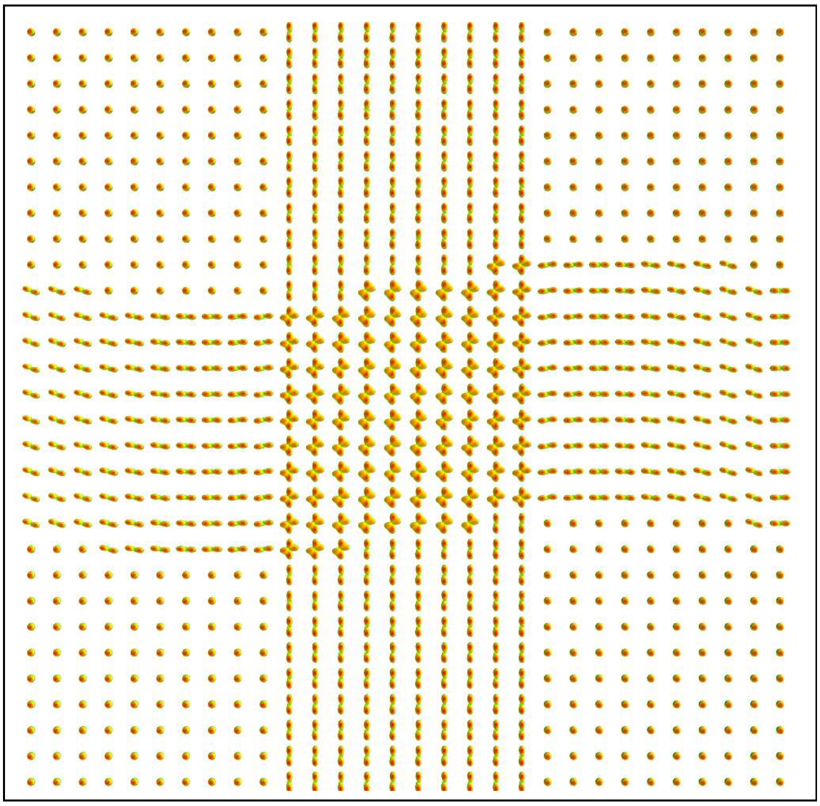


(a)



(b)

Figure 3.5: First example: synthetic tensor fields. 3.5(a) **FS** and 3.5(b) **SD+PPD** reorientated tensor fields. It is clear that **FS** gave wrong solutions in the vertical bundle of fibers.

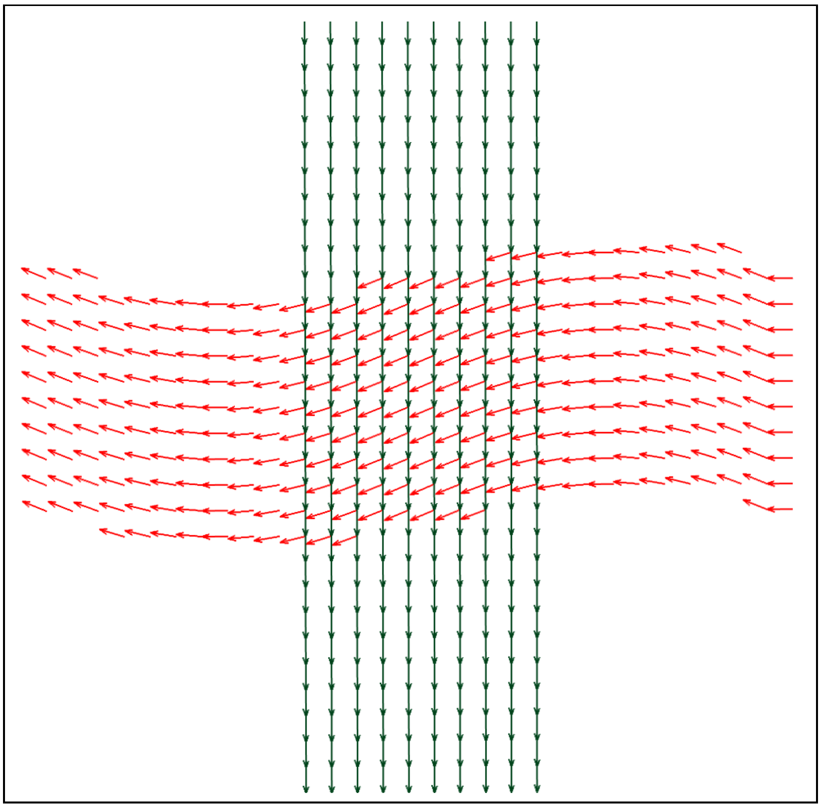


(a)

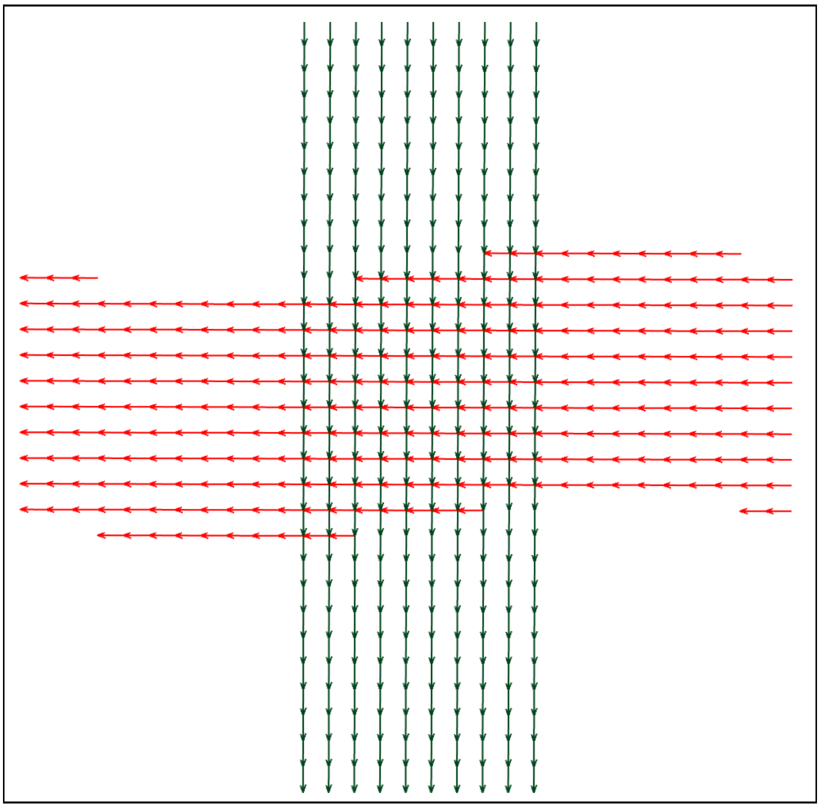
Figure 3.6: First example: synthetic tensor fields. 3.6(a) **HD+PPD** reorientated tensor field.

Method	Avg Horizontal AE	Avg Vertical AE
INITIAL (no reo)	15.36	0
FS [2]	7.58	11.82
SD+PPD [2]	3.25	0
HD+PPD [134]	3.59	0

Table 3.1: First synthetic example: Angular errors (AE) for the compared methods. **SD+PPD** and **HD+PPD** gave similar solutions, while **FS** provided solutions with larger AE than the others.

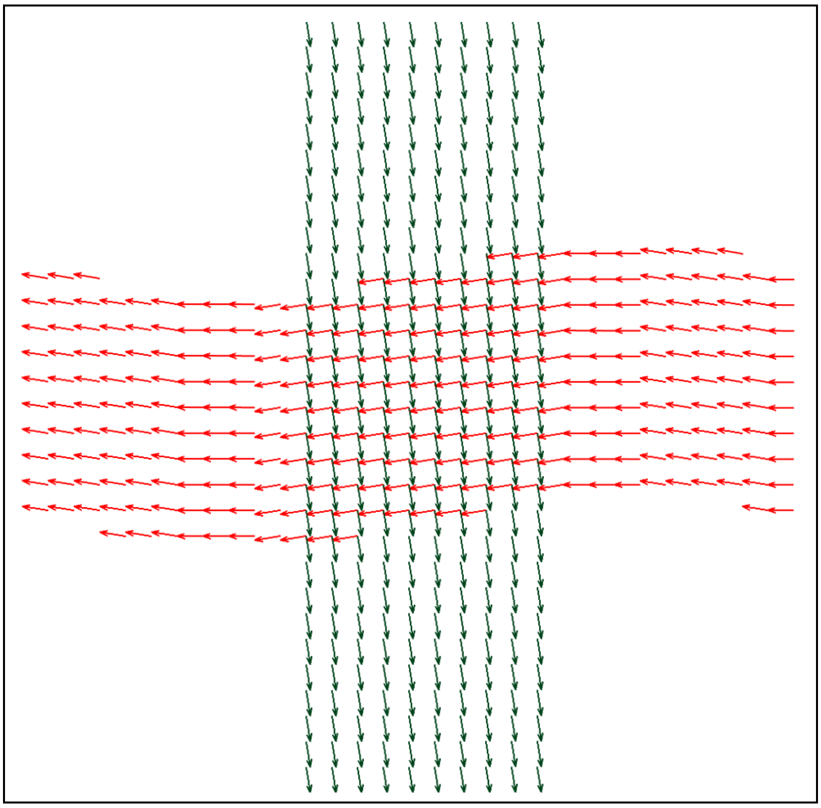


(a) Ground truth

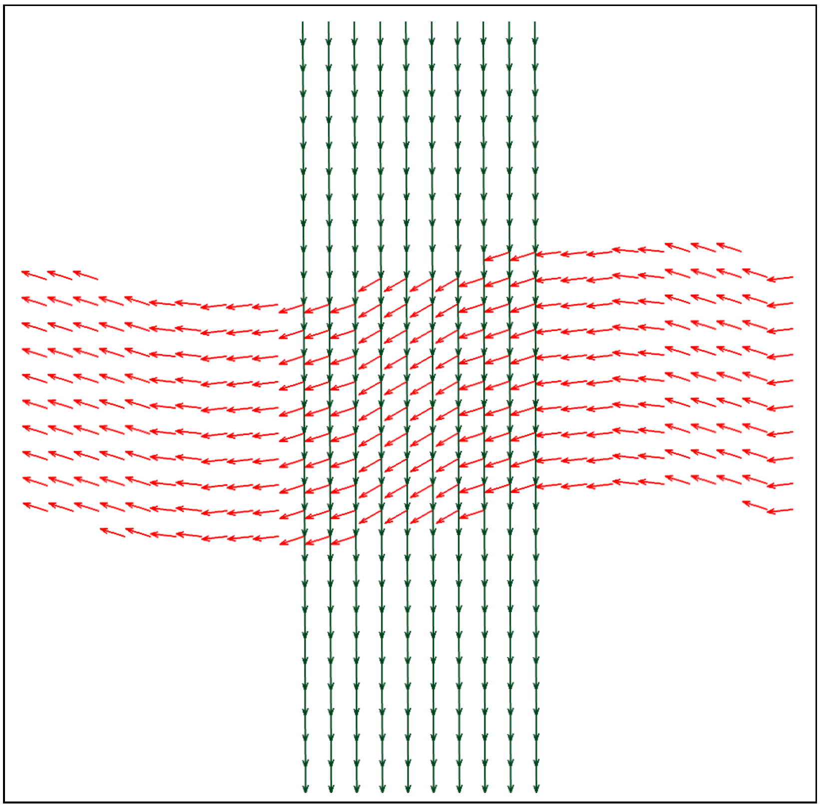


(b) Initial Peaks

FIGURE 3.7: First example: synthetic tensor's peaks. 3.7(a) GT and 3.7(b) initial peaks.

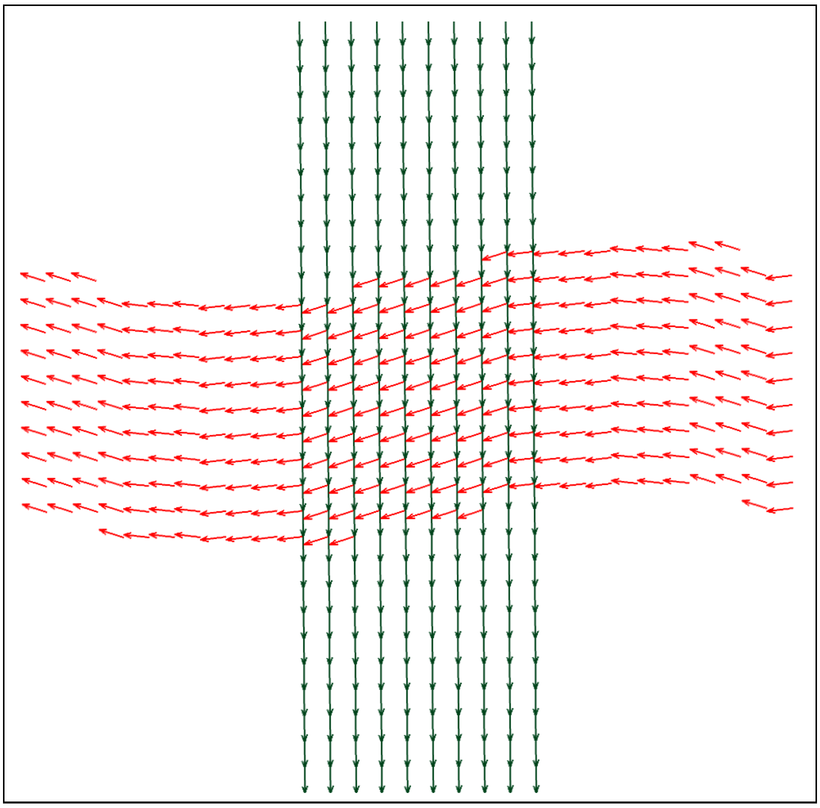


(a) **FS** Peaks



(b) **SD+PPD** Peaks

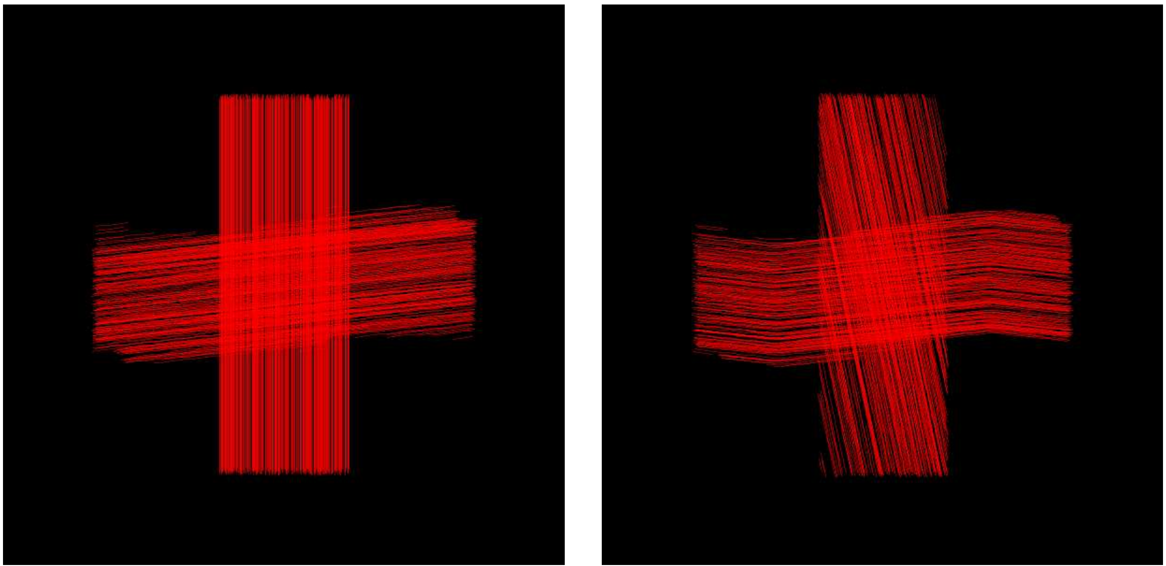
FIGURE 3.8: First example: synthetic tensor's peaks. 3.8(a) **FS** and **SD+PPD** resulting peaks.



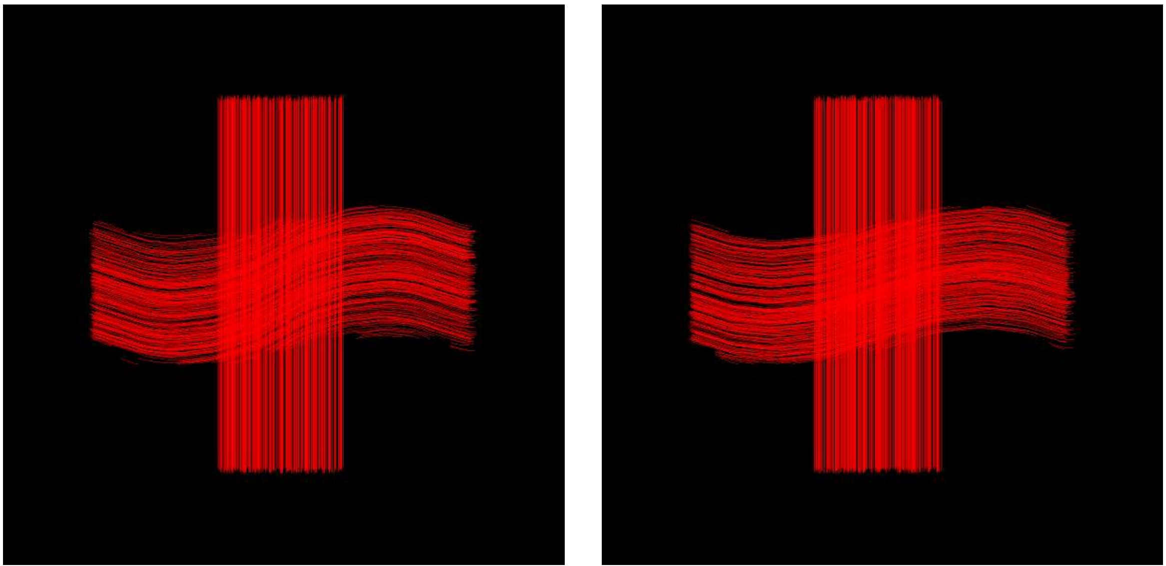
(a) **HD+PPD** Peaks

FIGURE 3.9: First example: synthetic tensor's peaks. 3.9(a) **HD+PPD** resulting peaks.

Another way to compare visually the methods is to determine fiber tracts on them. Figure 3.10 contains all corresponding tractographies, for the initial spatially registered T4 fields and the three reorientated fields. As we can notice, **SD+PPD** and **HD+PPD** gave identical tracts, while **FS** has problems especially in the vertical bundle of fibers.

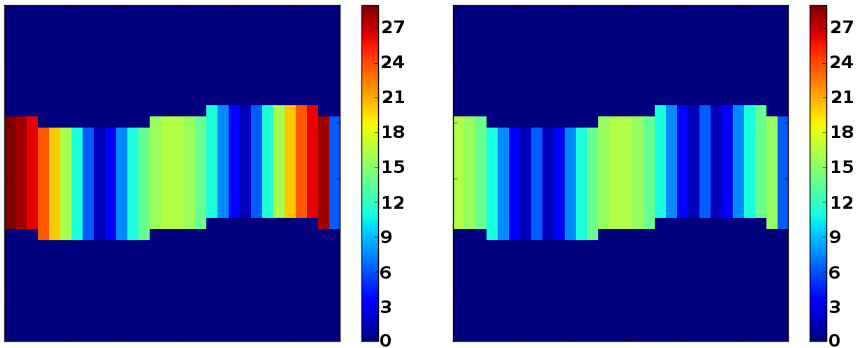


(a) Initial (b) **FS**

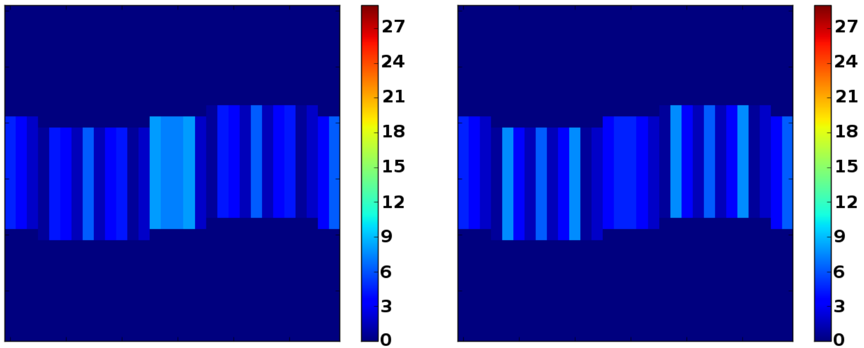


(c) **SD+PPD** (d) **HD+PPD**

FIGURE 3.10: First example: synthetic tensor's tractographies. As it is seen, **FS** produced wrong tracts especially in the vertical bundle of fibers, while **SD+PPD** and **HD+PPD** give equivalent tracts.

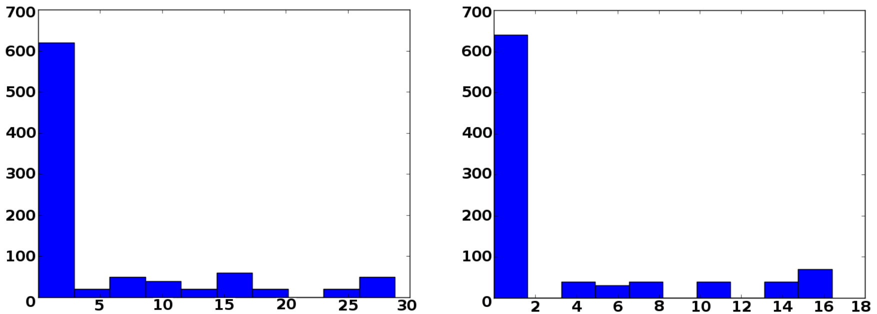


(a) Initial Horizontal AE (b) **FS** Horizontal AE

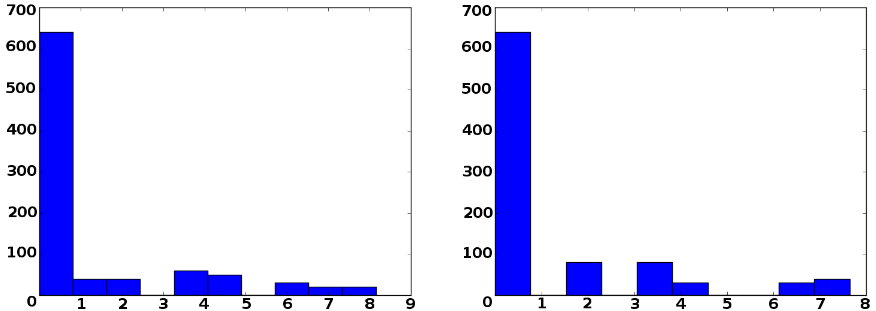


(c) **SD+PPD** Horizontal AE (d) **HD+PPD** Horizontal AE

FIGURE 3.11: First example: 3.11(a)- 3.11(d): the horizontal angular errors (AE). **SD+PPD** and **HD+PPD** perform better than **FS**.

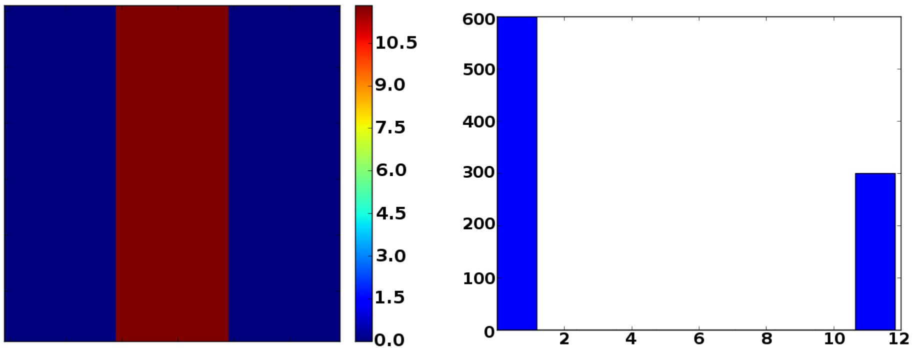


(a) Initial Horiz. AE Histogram (b) **FS** Horizontal AE Histogram



(c) **SD+PPD** Horiz. AE Hist. (d) **HD+PPD** Horiz. AE Hist.

FIGURE 3.12: First example: 3.12(a)- 3.12(d): the corresponding histograms of the horizontal AE presented in figure 3.11. Although, all methods manage to reduce the high initial horizontal error, **FS** did not succeed to reach the performances of **SD+PPD** and **HD+PPD**.



(a) **FS** Vertical AE (b) **FS** Vertical AE Histogram

FIGURE 3.13: First example: resulting vertical angular error (AE) (on the left) and the corresponding histogram (on the right) in the **FS** case.

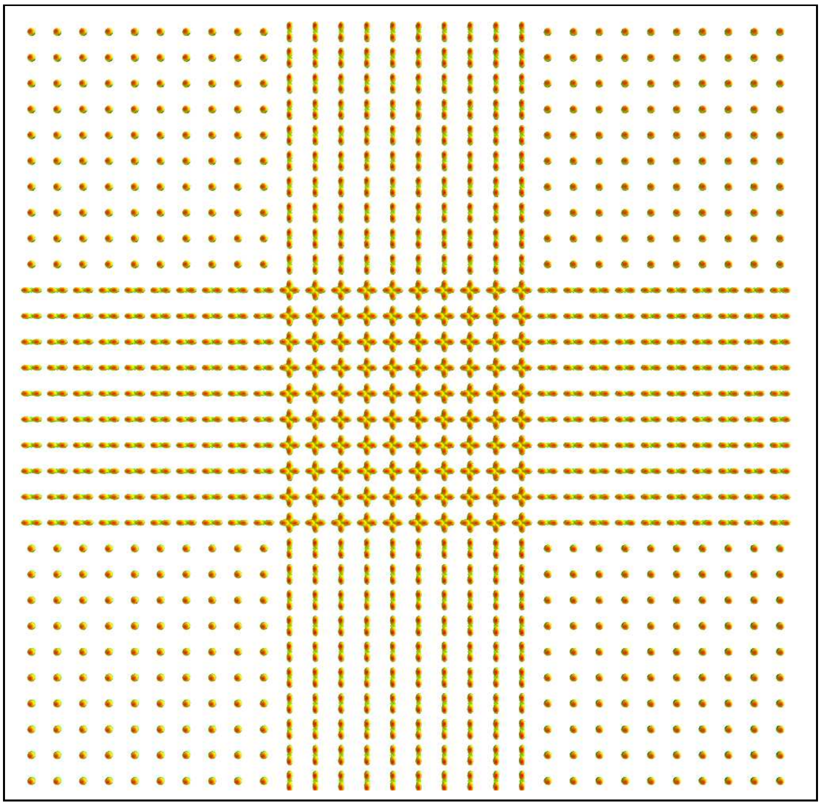
Synthetic Example 2

Another testing scenario is constructed by increasing the initial average angular error in the horizontal bundle of fibers in the first example (*e.g.* increase the angle of the sine transformation). This example highlights interesting limitations of the T4 model in representing two crossing fibers when their main directions are very close to each other. Figure 3.14 displays the initial and figures 3.15, 3.16 the resulting tensor fields. Figures 3.17- 3.19 contain the extracted peaks produced from each technique and figure 3.20 their tractographies. Observing the tractographies, it seems that **FS**'s horizontal tracts are better than **SD+PPD** and **HD+PPD**. In fact, the two last methods transformed a few tensors in the center from crosses to single fibers and as a consequence, the horizontal tracts are interrupted.

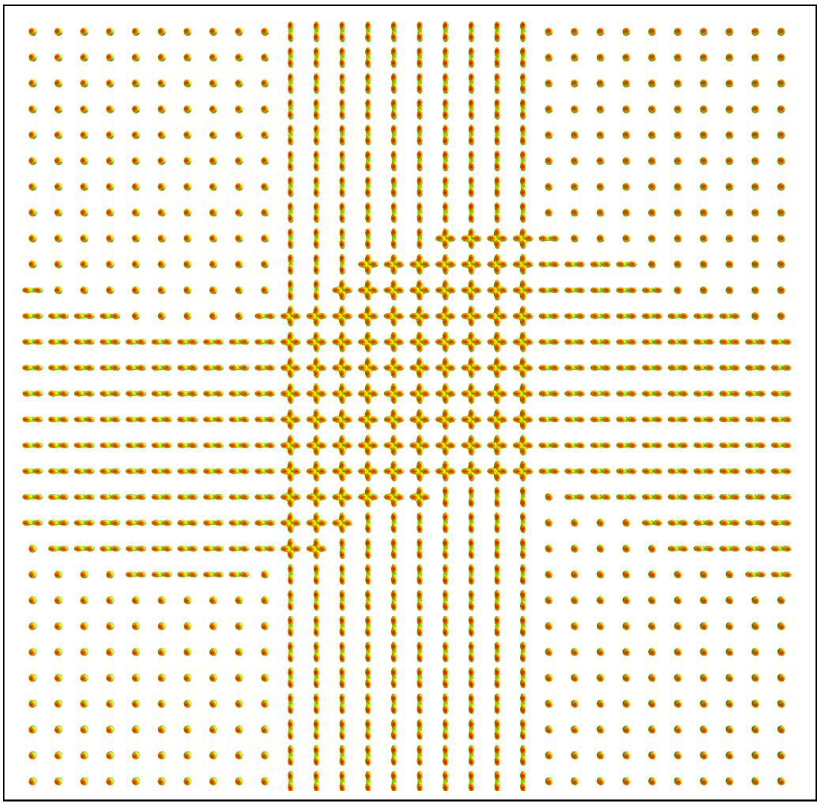
On the other hand, figures 3.21 and 3.23 show the horizontal and the vertical AE of each case, while figure 3.22 and 3.24 correspond to the histograms of the horizontal and vertical AE of the second synthetic example, respectively. **SD+PPD** and **HD+PPD** managed to reduced significantly the initial angular errors and gave identical results. On the contrary, **FS** produced vertical AE, but also did not manage to reduce a lot the horizontal AE, since **FS** does not apply the whole transformation, but uses only the rotation part of it.

To continue our previous discussion, a careful observation of figures 3.18(b) and 3.19 concludes to the existence of areas where instead of having two main directions we obtained only one (left part of the center view of the image). Figure 3.25 shows the evolution of the transformation in the center of the tensor field that contains tensors representing two crossing fibers. By plotting the fODFs of the three T4s on a 2D grid constructed by the parameters of the sampling scheme of the $3D$ unit hemisphere (the *elevation* and the *azimuth*, denoted as "*theta*" and "*phi*", respectively, in the graphs), with 0.5 degrees of sampling step, it is clear that as we move from the right part of the tensor field to the left, (and from the top to the bottom in this figure) the two initial peaks are reduced to one.

Despite the errors because of the model's limitation, table 3.2 shows that **HD+PPD** method is equivalent with **SD+PPD** and both are significantly better than **FS**.

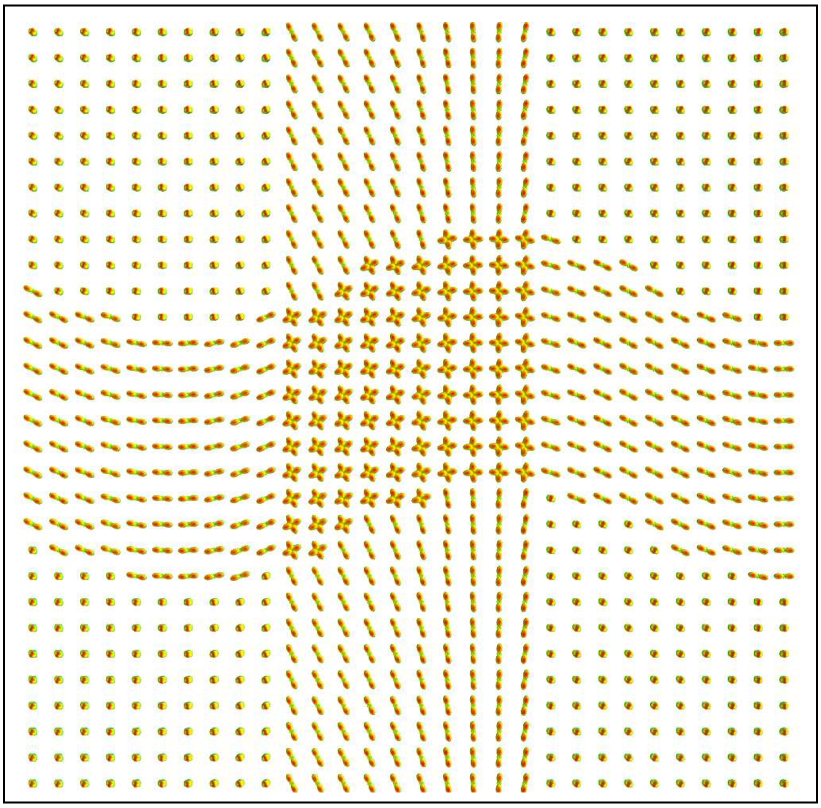


(a)

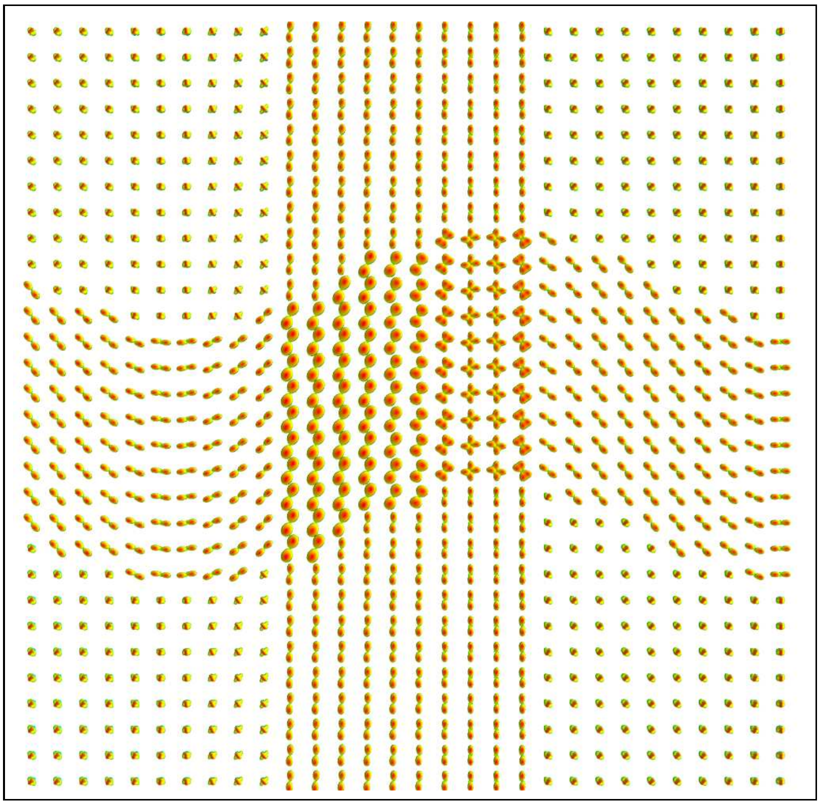


(b)

FIGURE 3.14: Second example: Synthetic tensor fields. 3.14(a) The template of the tensor field and 3.14(b) the initial registered template (no reorientation yet).

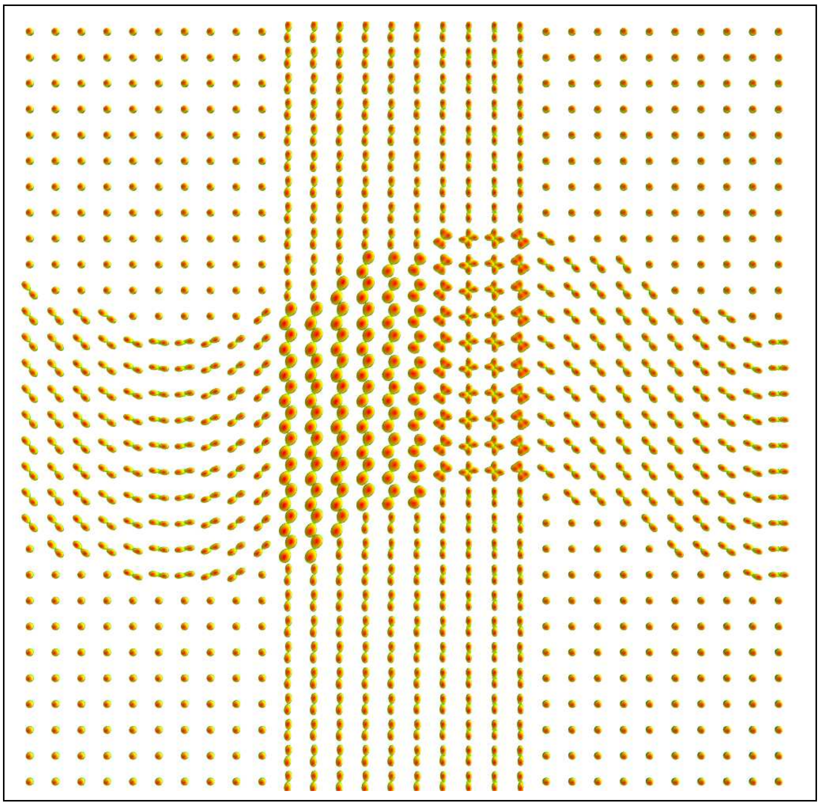


(a)



(b)

FIGURE 3.15: Second example: Synthetic tensor fields. 3.15(a) **FS** and 3.15(b) **SD+PPD** reorientated tensor fields. **FS** as it was expected had still problems in the vertical bundle of fibers.

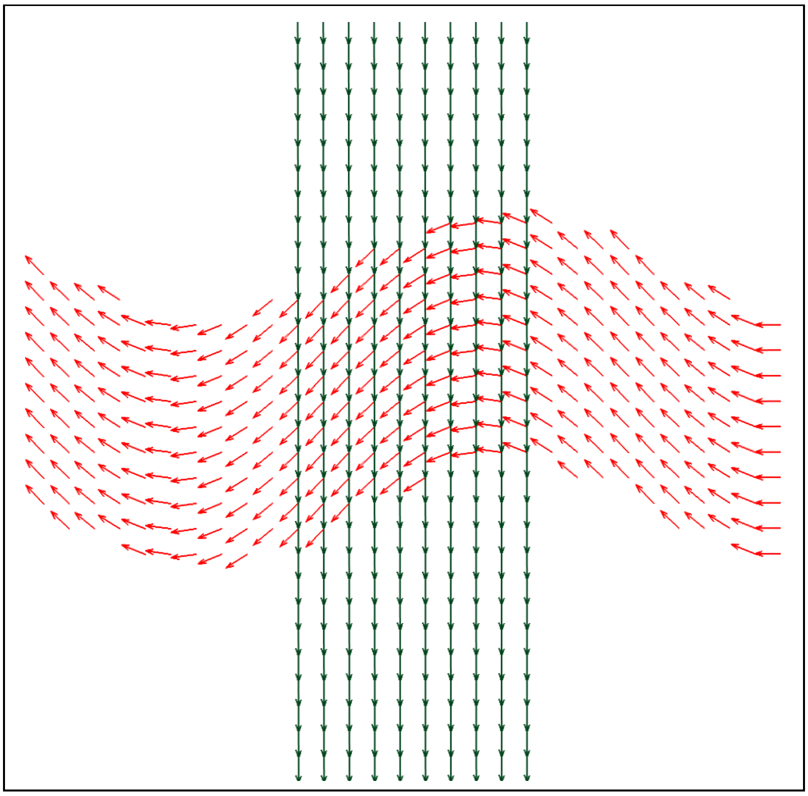


(a)

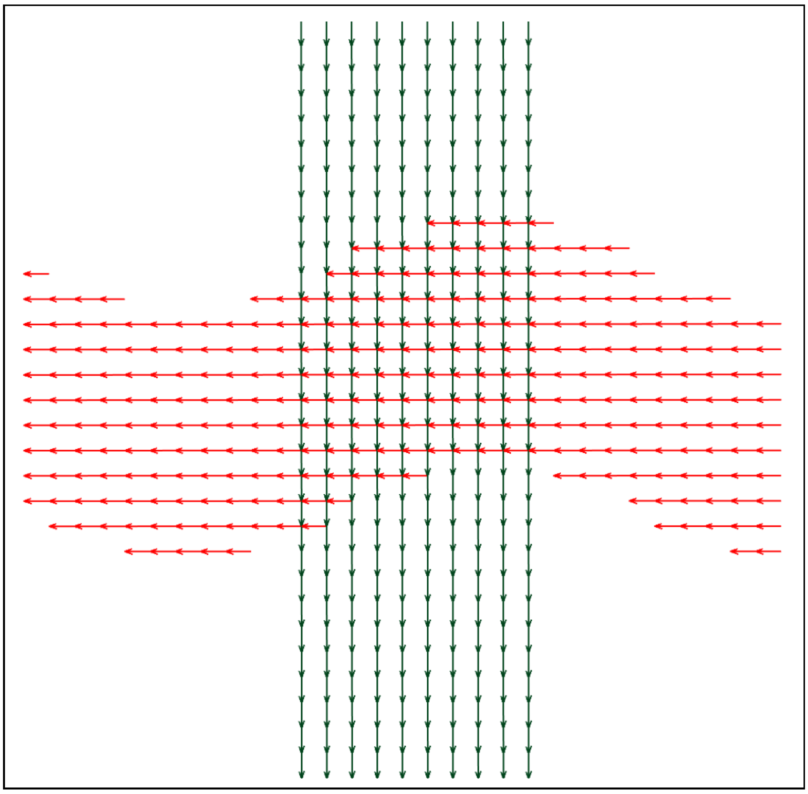
FIGURE 3.16: Second example: Synthetic tensor fields. 3.16(a) **HD+PPD** reorientated tensor field.

Method	**Avg Horizontal AE**	**Avg Vertical AE**
INITIAL (no reo)	31.95	0
FS [2]	14.62	15.66
SD+PPD [2]	3.67	$4.52e^{-06}$
HD+PPD [134]	3.67	$4.52e^{-06}$

TABLE 3.2: Second synthetic example: Angular errors (AE) for the compared methods.

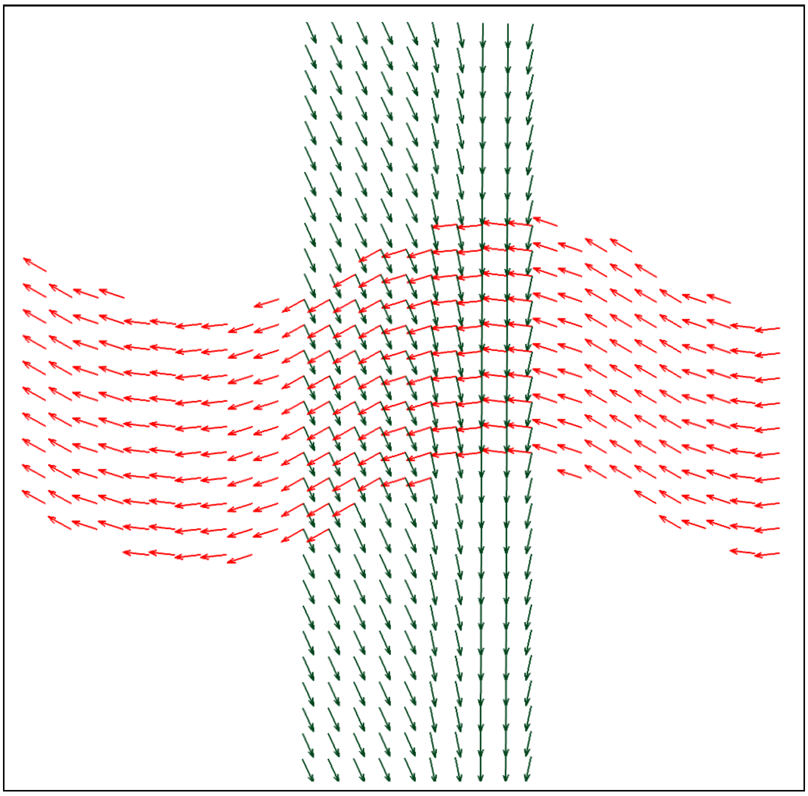


(a) Ground truth

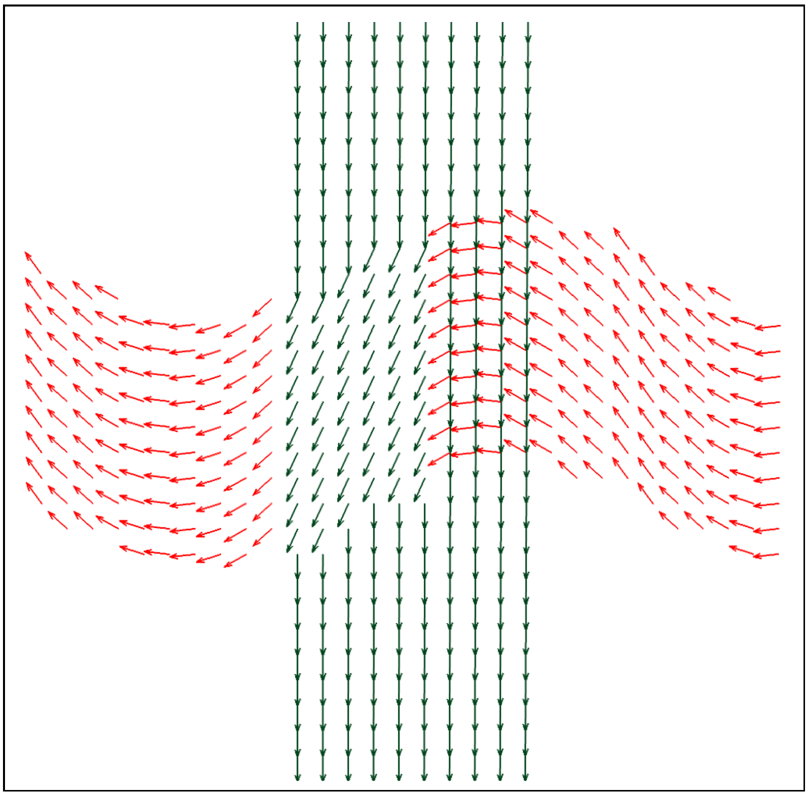


(b) Initial Peaks

Figure 3.17: Second example: Synthetic tensor's peaks. 3.17(a) GT and 3.17(b) initial peaks.

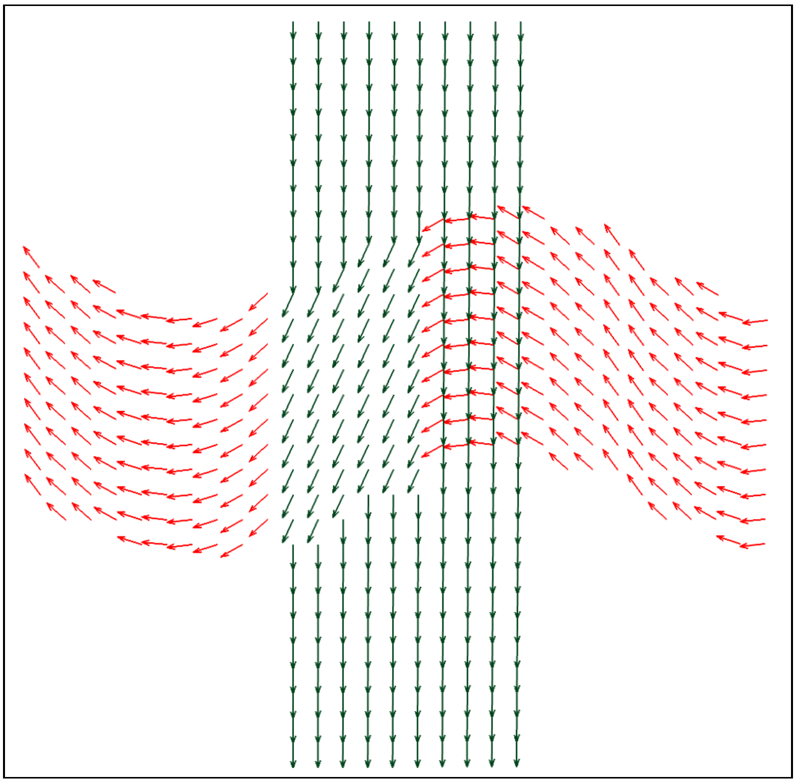


(a) **FS** Peaks



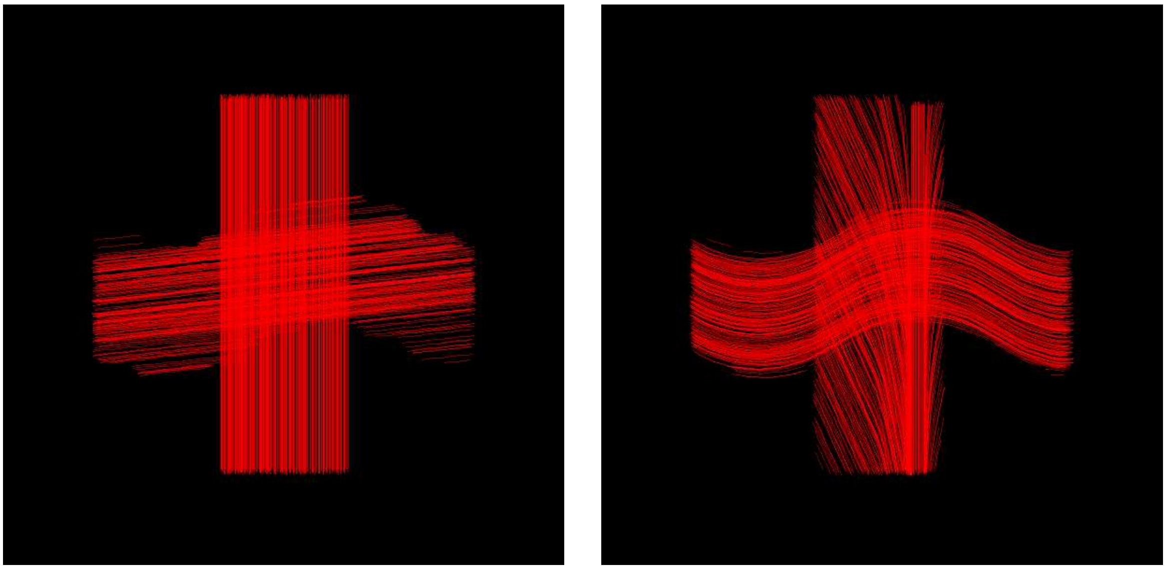
(b) **SD+PPD** Peaks

FIGURE 3.18: Second example: Synthetic tensor's peaks. 3.18(a) **FS** and 3.18(b) **SD+PPD** resulting peaks. In the **SD+PPD** case, it is noticeable that a few T4s in the center produced one principal direction instead of two, due to the T4 limitation to describe well very close to each other fibers.

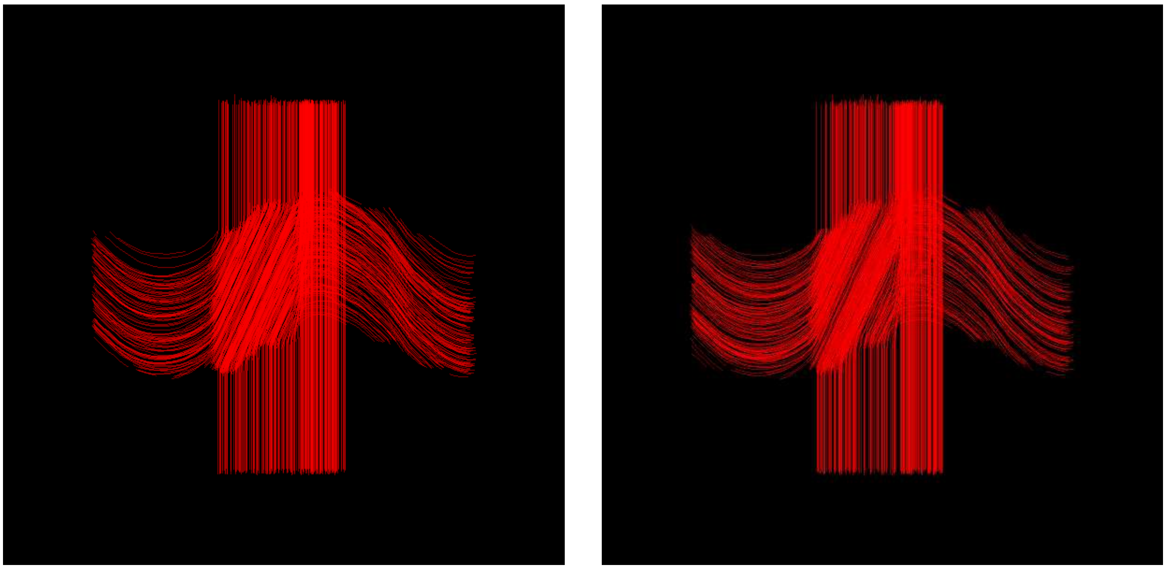


(a) **HD+PPD** Peaks

FIGURE 3.19: Second example: Synthetic tensor's peaks. **HD+PPD** resulting peaks. Similarly to the **SD+PPD** case, **HD+PPD** highlighted the T4 limitation to present well two very close to each other fibers.

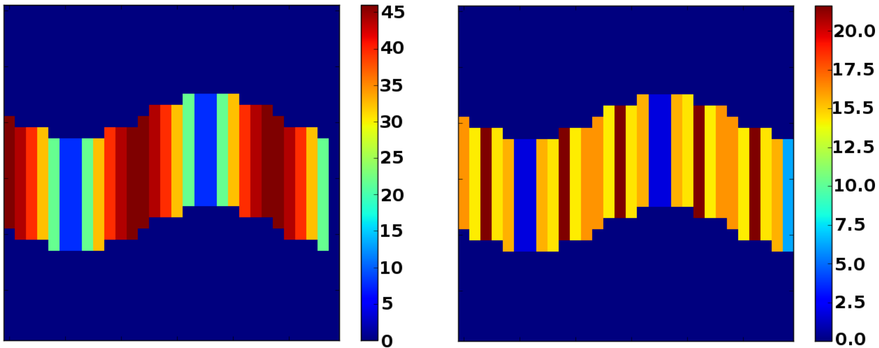


(a) Initial (b) **FS**

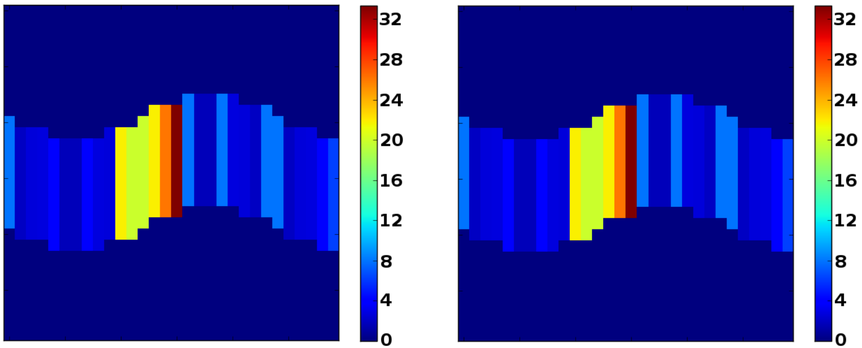


(c) **SD+PPD** (d) **HD+PPD**

FIGURE 3.20: Second example: Synthetic tensor's tractographies. **SD+PPD** and **HD+PPD** give similar results. Although **FS** horizontal tracts look better than **SD+PPD** and **HD+PPD**, a careful observation of the angular errors signify that **FS** is not close to the right answer. **SD+PPD** and **HD+PPD** horizontal tracts appear to be discontinuous as a result of getting one lobe (*i.e.* one principal direction of diffusion), instead of two, since the two initial directions got very close after the reorientation for some tensors in the center. These resulting single directions are used in vertical tracts. In addition, **FS** vertical tracts are significantly different than the correct answer.

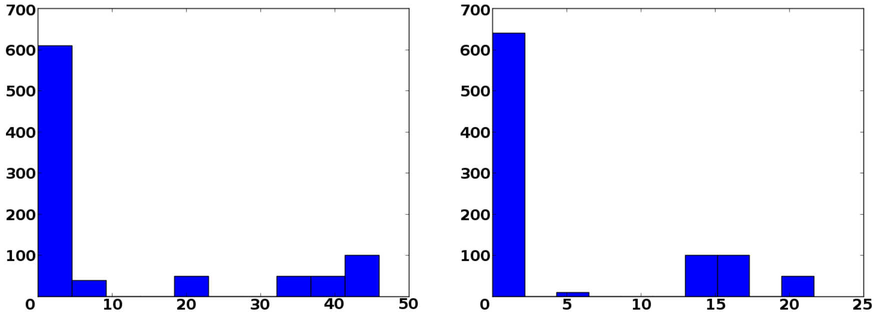


(a) Initial Horizontal AE (b) **FS** Horizontal AE

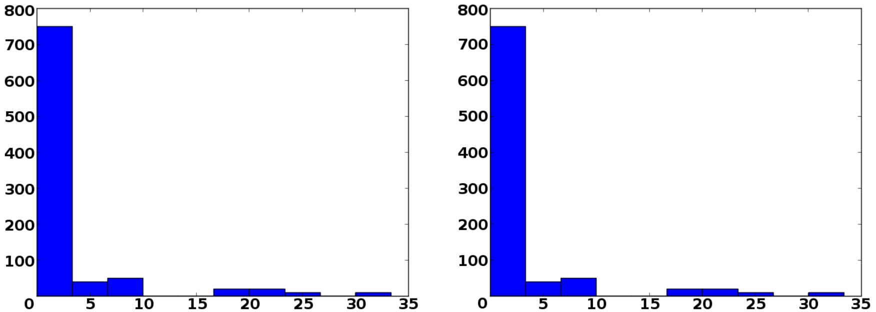


(c) **SD+PPD** Horizontal AE (d) **HD+PPD** Horizontal AE

FIGURE 3.21: Second example: 3.21(a) 3.21(d): horizontal angular errors (AE).

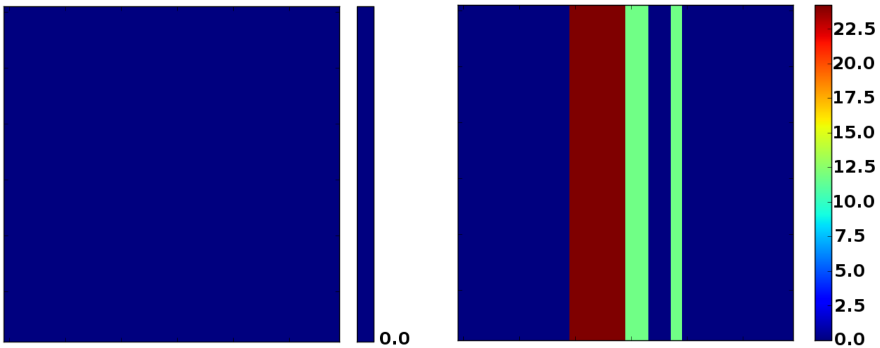


(a) Initial Horiz. AE Histogram (b) **FS** Horizontal AE Histogram

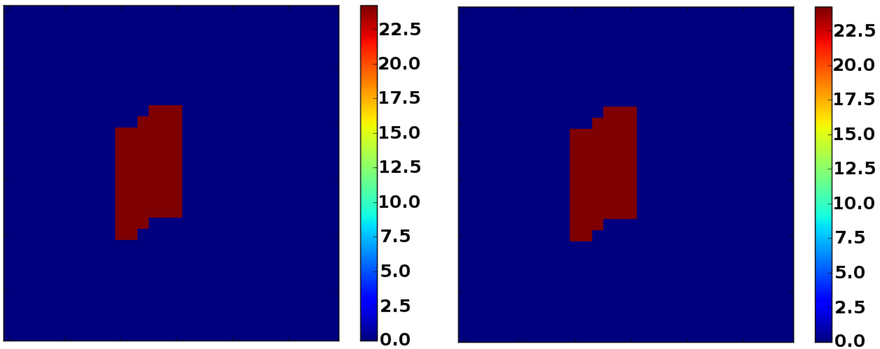


(c) **SD+PPD** Horiz. AE Hist. (d) **HD+PPD** Horiz. AE Hist.

FIGURE 3.22: Second example: 3.22(a)- 3.22(d): the corresponding histograms of the horizontal AE presented in figure 3.21.

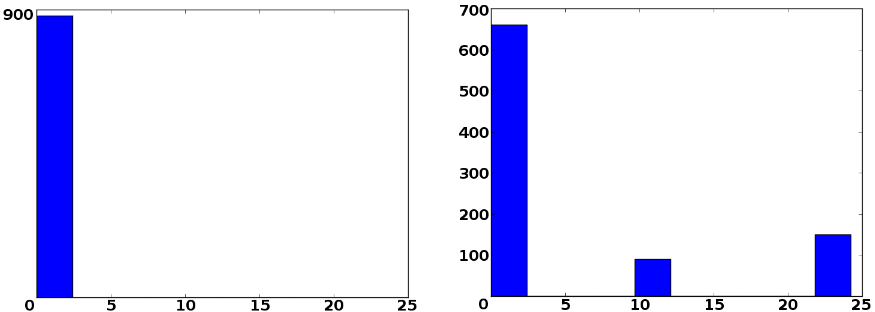


(a) Initial Vertical AE (b) **FS** Vertical AE

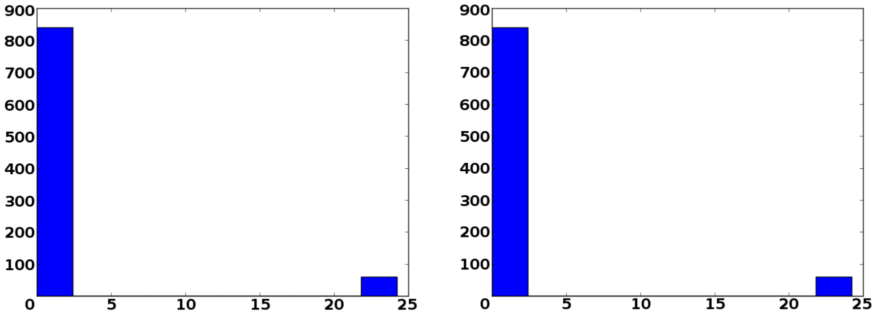


(c) **SD+PPD** Vertical AE (d) **HD+PPD** Vertical AE

FIGURE 3.23: Second example: 3.23(a)- 3.23(d): the vertical angular errors (AE).



(a) Initial Vertical AE Histogram (b) **FS** Vertical AE Histogram



(c) **SD+PPD** Vertical AE Hist. (d) **HD+PPD** Vertical AE Hist.

FIGURE 3.24: Second example: 3.24(a)- 3.24(d): the corresponding histograms of the vertical AE presented in figure 3.23.

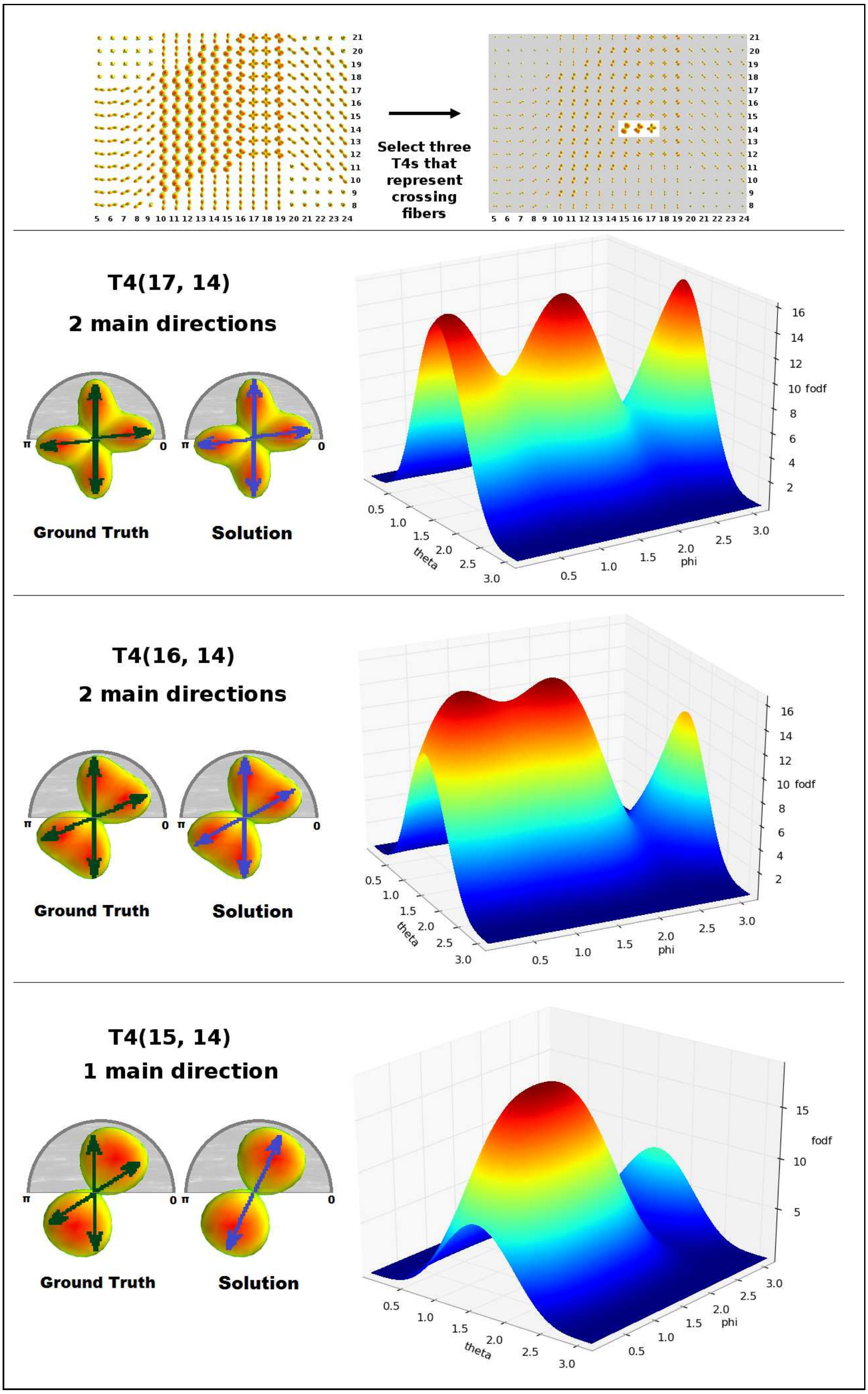


FIGURE 3.25: Influence of the transformation in areas that contain two crossing fibers, in the second synthetic example. In the left part, the ground truth and the resulting peaks are displayed, while in the right part the fODFs, estimated in the unit hemisphere of each tensor, highlight the limitation of the T4 model to represent two bundle of fibers with small angular difference, in detail.

3.7.2 Real Data

Evaluation on real data was performed by using a HARDI dataset of a healthy person. The size of the images is $128 \times 128 \times 41$ for a resolution of $1.8 \times 1.8 \times 3.5mm^3$ and 30 non colinear gradient directions (scanned twice) were used, while b value is equal to $1000s/mm^2$.

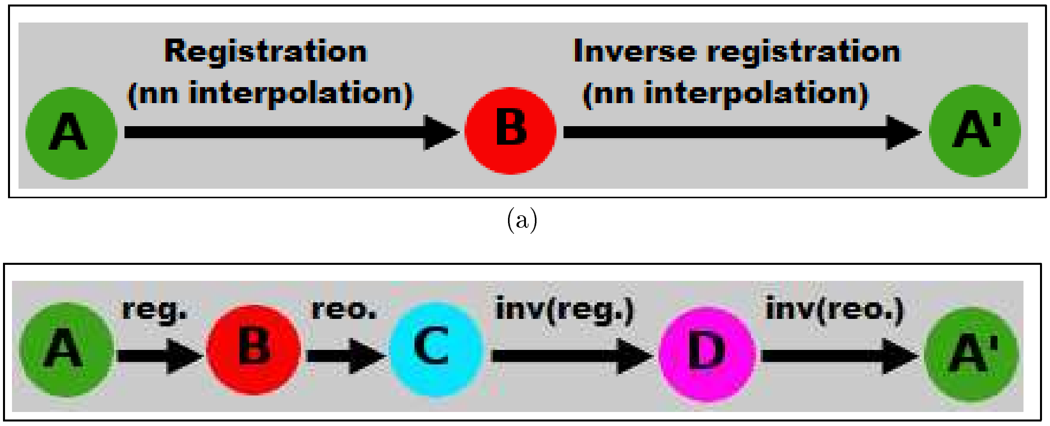
Figures 3.27, 3.28 show the reorientated T4 fields for a selected 20×20 patch containing crossing fibers. Looking at the fODF glyphs, we can claim that **SD+PPD** and **HD+PPD** produced slightly different results, while on the other hand, **FS**'s solution has many differences, as a consequence of extracting and using only the rotation part of the estimated non-linear transformation F. In fact, the **FS** reorientated T4s are very similar to the registered T4s (but not reorientated), meaning that probably the non-linear transformation contains shearing or scaling effects that cannot be included in the rotation part that **FS** uses.

Figure 3.29 shows the corresponding tractographies in a selected ROI (including the same 20×20 patch along with 3 more frames in the z-axis). **SD+PPD** and **HD+PPD** methods managed to produce more dense tractographies than **FS**, especially in the bottom part of the images. Moreover, **FS** tractography seems not to vary significantly from the initial tractography on the registered data with no reorientation step.

In contrast to the synthetic cases where we can calculate the real orientation (GT) of the main directions (given the number of them and the transformation matrix F), we cannot work similarly in the real data case. For that reason, another evaluation scheme is proposed in figure 3.26. In order to use the new evaluation process, a proper distance must be selected as an error metric. In this study, the approximation of the distance defined in eq. 2.30, between two fODF functions is selected (eq. 2.31).

At this point that the distance is selected, let us describe the current evaluation strategy. Firstly, we will measure the registration error between points **A** and **A'** of figure 3.26(a). Point **A** corresponds to the initial tensor fields, without performing either registration or reorientation. Point **A'** is constructed by applying forward and then inverse registration on the tensor coefficients. As a result, no reorientation error is included in technique 3.26(a). Equivalently, we will calculate the total normalization error (registration + reorientation) of each of the three methods *via* the strategy of figure 3.26(b). Of course, it is possible to measure the reorientation error directly, without firstly calculating the registration error.

Figure 3.30 enumerates the distances measured in frame 33 (in z-axis) of the DW-MRI data. Values of -10 in the distance images correspond to voxels outside the WM area (so that those voxels can be marked and excluded from the calculation of the average distances presented in tables 3.3, 3.4). The WM can be located using a threshold in the FA images (as in our case), or by using a template. It is observed that the majority of errors are close to zero (very low), while areas with larger errors are located in the same parts of the brain for all the three compared methods. Figure 3.31 depicts the corresponding histograms of the distances presented in figure 3.30 (z-frame 33).



(a)

(b)

FIGURE 3.26: Evaluation schemes for real data: measure the error between **A** and **A'**. 3.26(a) Registration error and 3.26(b) Registration + Reorientation error.

Observing the order of errors (10^{-1}) in tables 3.3 (of z-frame 33), 3.4 (of all z-frames) we understand that these errors are very low in comparison to the higher distances presented in the simulation test of figure 2.16 ($10 - 10^2$), concluding that the obtained solutions contain a very small amount of error.

Theoretically, we would expect to notice no errors in the **FS** method, since the shape of the tensors is intact, but in practice we obtained errors which can be justified due to the presence of registration errors and the estimation of the rotation part of the transformation that **FS** uses.

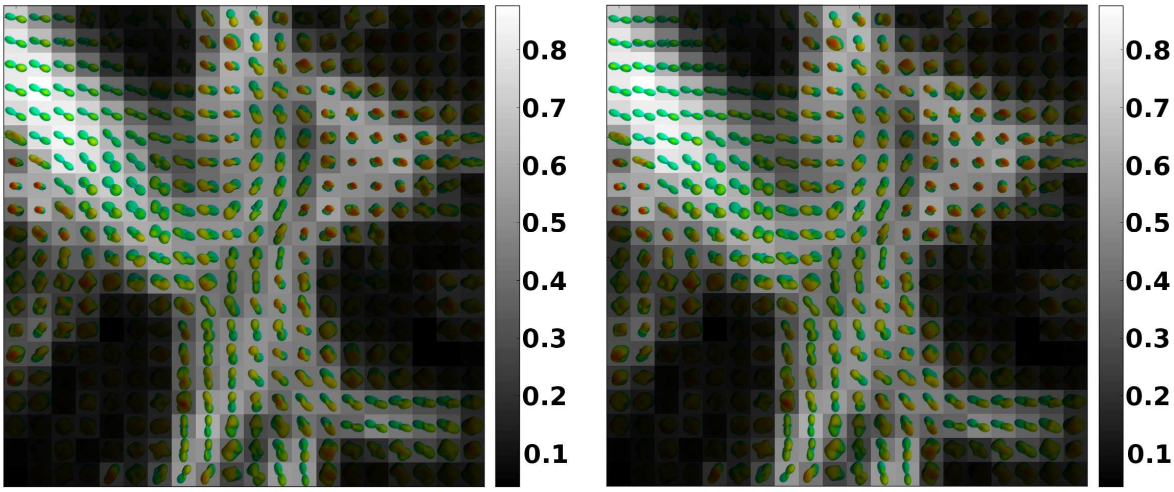
On the other hand, checking only the average errors can potentially hide any sparse larger errors. For example, maximum errors in fig. 3.30 are close to 8, similarly to the bottom part of fig. 2.16. The appearance of such errors does not ensure us that they are as large as to produce false biomarker detections (or in the contrary to cover the really differences) in statistical analysis. For that reason, the registration and reorientation of tensor coefficients should be carefully tested.

Method	Avg Distance (frame 33)
INITIAL (no reo)	0.02
FS [2]	0.51
SD+PPD [2]	0.57
HD+PPD [134]	0.6

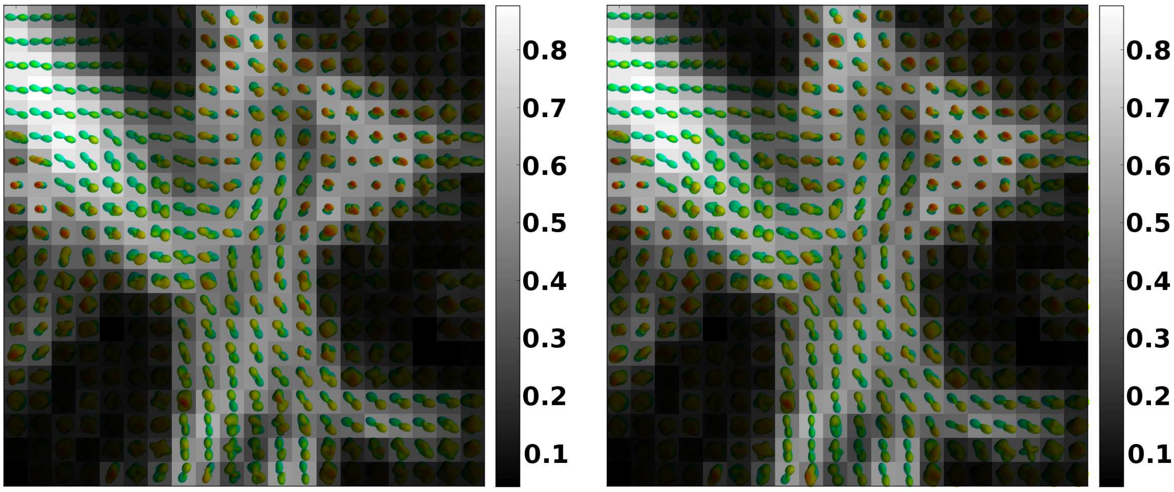
TABLE 3.3: Distances of the compared methods in the real data case of frame 33. The calculated errors are significantly low. **FS** errors should not be excepted, however they exists due to the presence of registration errors and the approximated rotation part of the transformation.

Method	Avg Distance (all frames)
INITIAL (no reo)	0.02
FS [2]	0.70
SD+PPD [2]	0.75
HD+PPD [134]	0.77

Table 3.4: Distances of the compared methods in the whole real data. The calculated errors are significantly low.

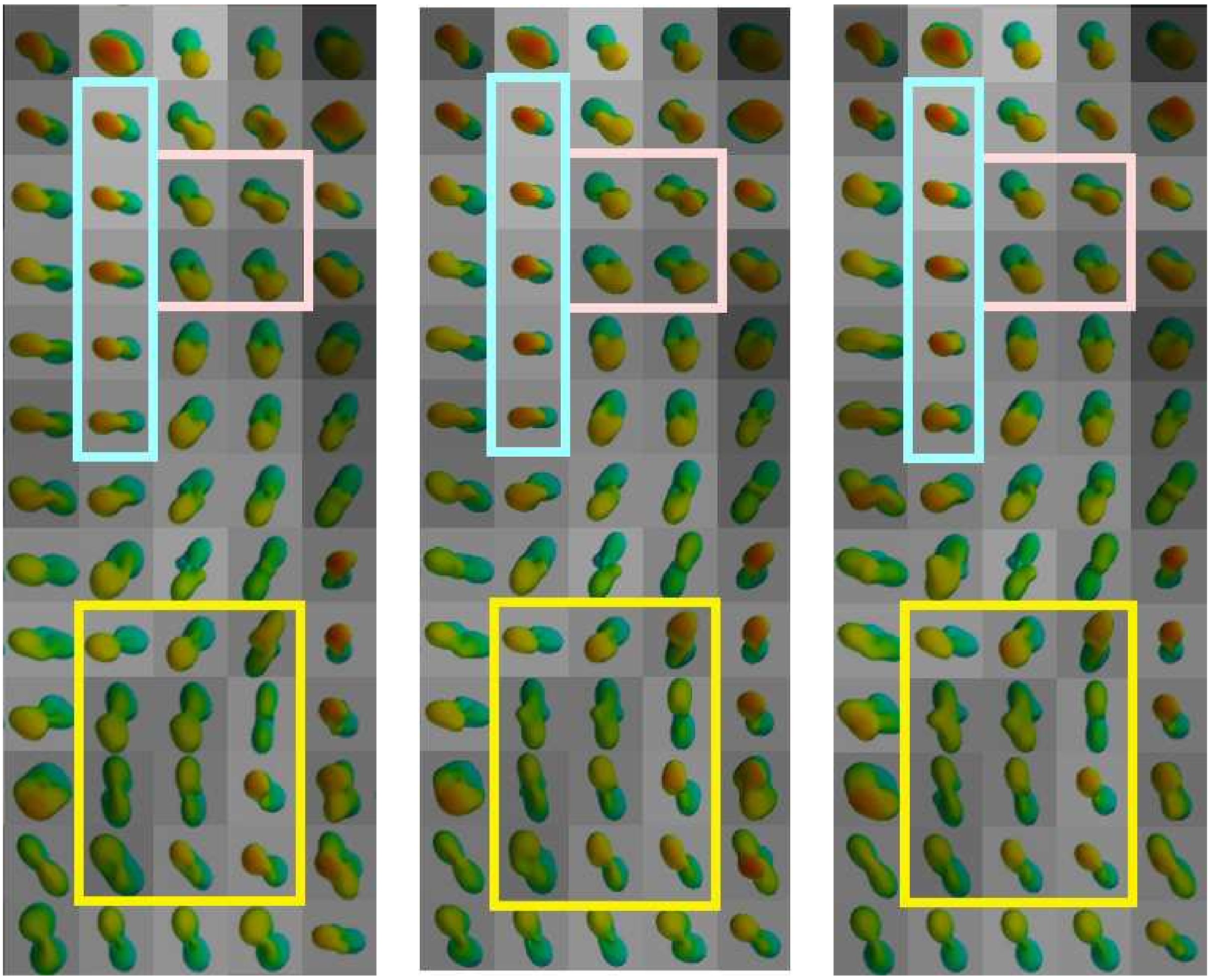


(a) reg. T4s NO reo. (frame 46)   (b) **FS** (frame 46)



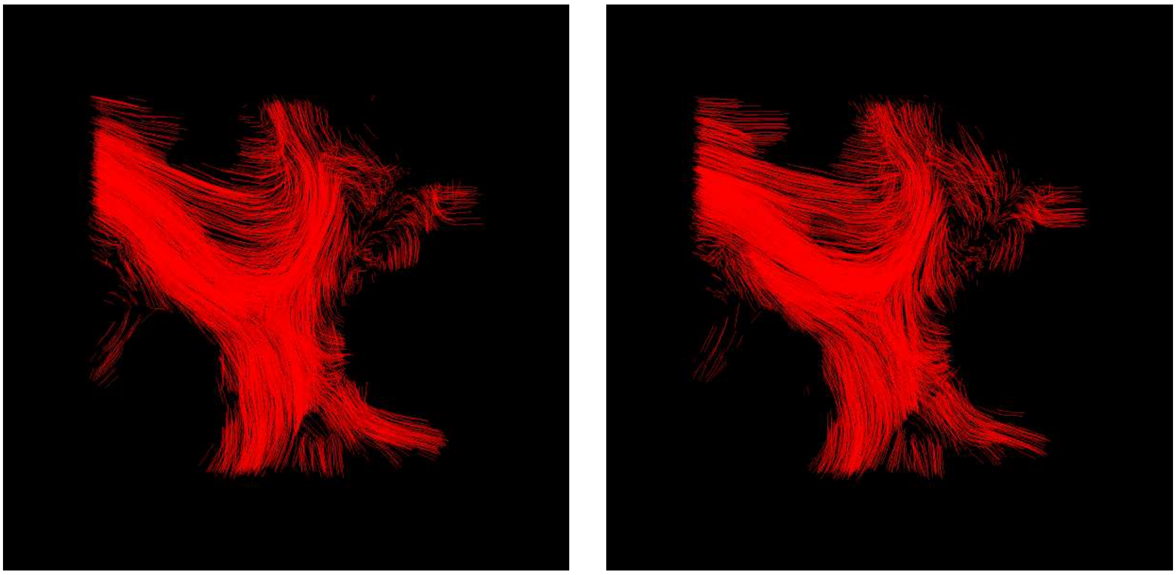
(c) **SD+PPD** (frame 46)   (d) **HD+PPD** (frame 46)

Figure 3.27: Resulting tensor fields (in a patch of 20×20 size) of the compared methods in a ROI with both single and crossing fibers. The fODF glyphs are plotted on the estimated FA images obtained by DTI analysis with the FSL toolkit [89].



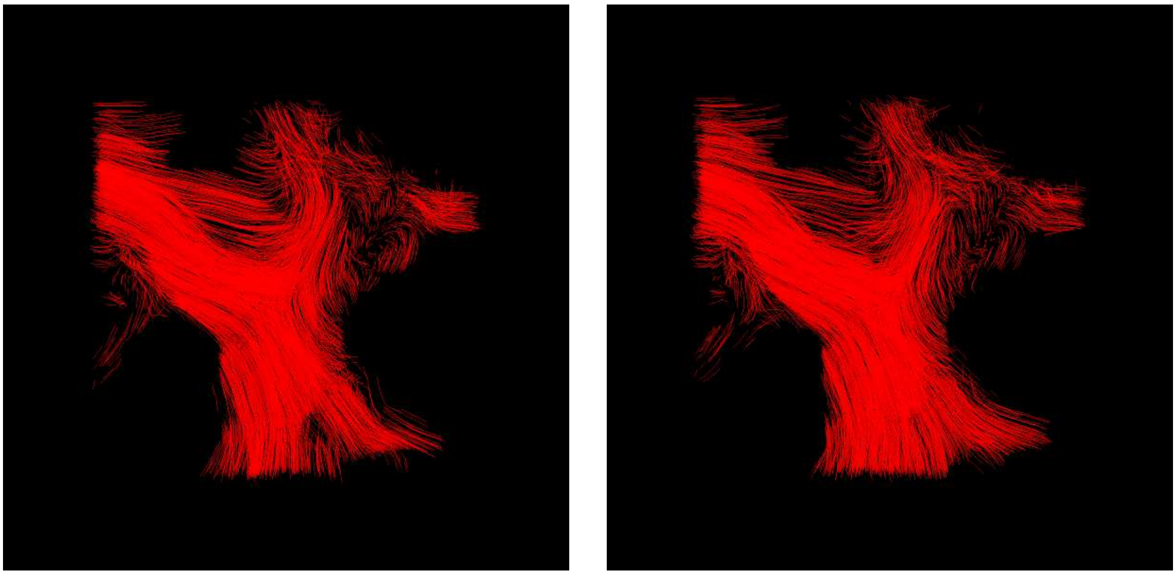
(a) **FS** (frame 46) (b) **SD+PPD** (frame 46) (c) **HD+PPD** (frame 46)

FIGURE 3.28: Zoom in particular areas of figure 3.27 in order to locate the differences. It seems that all the methods differ (more or less) from each other. **SD+PPD** and **HD+PPD** are more similar than **FS** which differs significantly in many areas.



(a) reg. T4s NO reo. Tractography (frames 44-47)

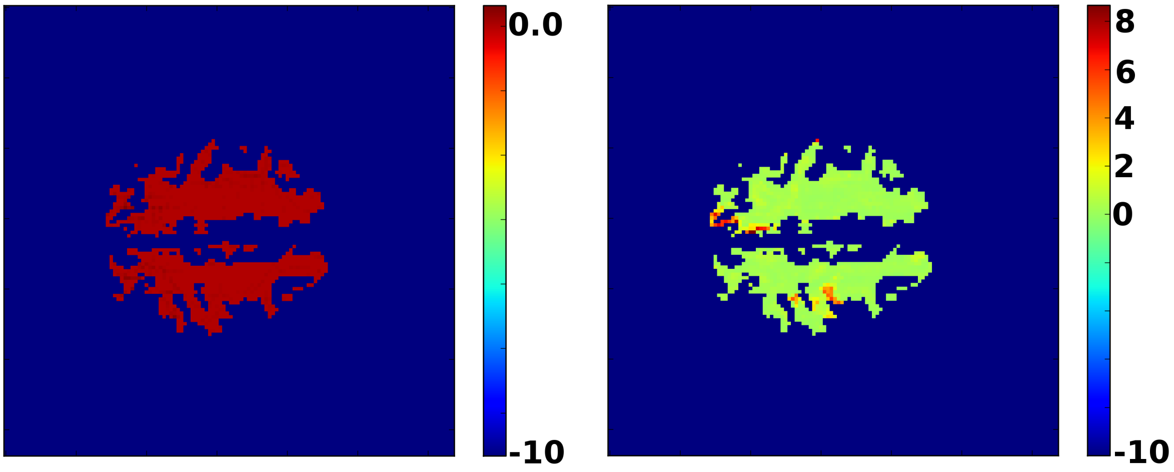
(b) **FS** Tractography (frames 44-47)



(c) **SD+PPD** Tractography (frames 44-47)

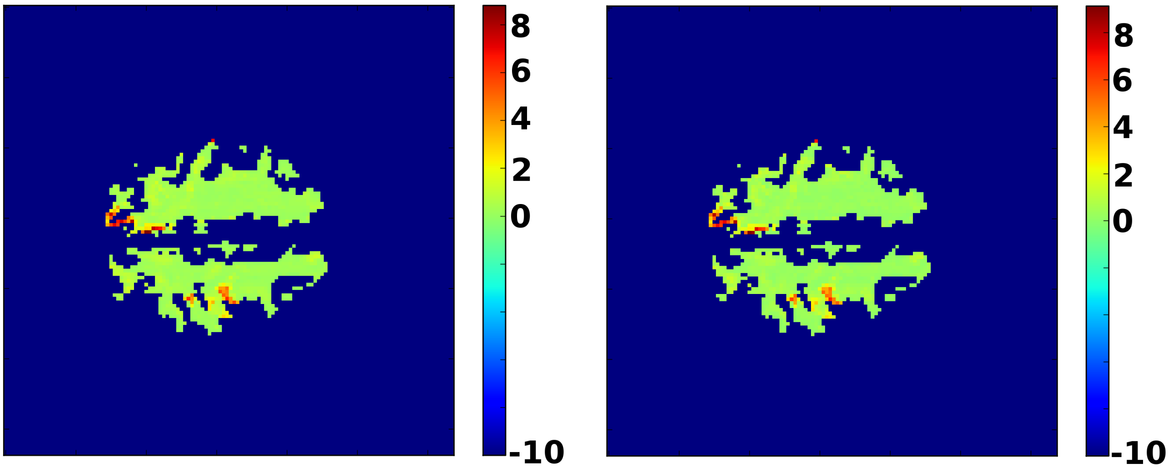
(d) **HD+PPD** Tractography (frames 44-47)

FIGURE 3.29: Resulting tractographies of the compared methods in a ROI with both single and crossing fibers.



(a) Registration Error

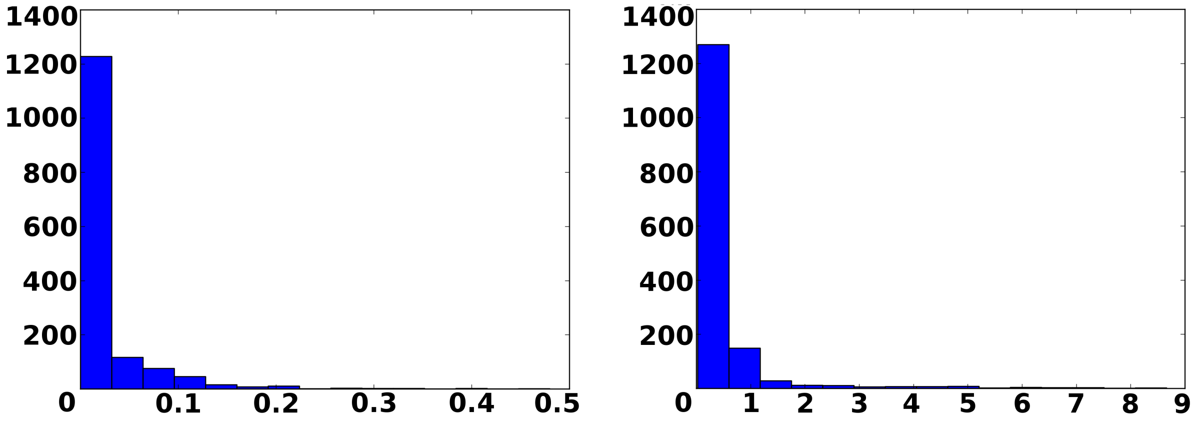
(b) **FS** Error



(c) **SD+PPD** Error

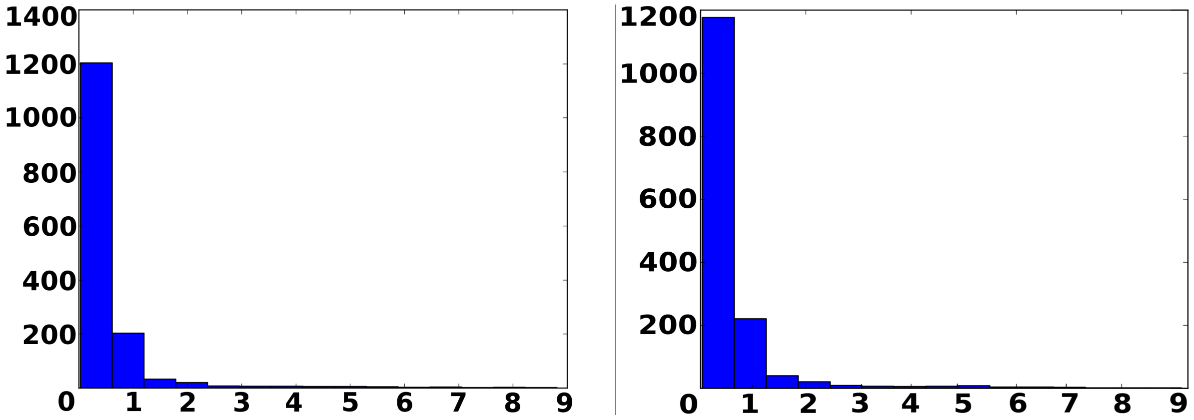
(d) **HD+PPD** Error

Figure 3.30: Distances of frame 33 (size of image: 128 × 128). Note: maximum value of subfigure 3.30(a) equals to 0.48.



(a) Histogram of Registration Error

(b) Histogram of **FS** Error



(c) Histogram of **SD+PPD** Error

(d) Histogram of **HD+PPD** Error

FIGURE 3.31: Histograms of the distances of frame 33.

3.8 Partial Conclusion

In this chapter, several methods for T4 reorientation are studied, discussed and evaluated in both synthetic and real data. Moreover, the performance of **HD+PPD** (the particular method of our interest) was further analysed, especially due to the lack of testing real cases in Renard's thesis [134]. **HD+PPD** is competitive with respect to **SD+PPD** method and better than **FS** which uses only the rotation part of the transformation.

Although in this chapter T4 normalizations were studied, afterwards DWI normalization was chosen to be used in the proposed statistical analyses for two main reasons. Firstly, a quite-promising method for non-linear DWI normalization was proposed in 2013. Secondly, the reorientation of a crossing T4 can potentially reduce the number of the main directions of diffusion, resulting into totally different fiber structure than the underlying one (as was previously shown, see fig. 3.25).

As part of future work, someone could try to spatially register the raw DWI data (instead of registering the tensor images), then fit tensor models on the registered DWI data and finally reorientate the tensor models with **FS**, **SD+PPD** and **HD+PPD** (instead of rotating the b-vectors as the DWI normalization is performed in [51, 152] and then fitting tensor models). In this way, we avoid tensor image registration, which is more risky than DWI registration (errors in tensor coefficients can affect many more directions of diffusion than errors in DWI images). Moreover, it is efficient to use the whole transformation matrix F in order to reorientate the tensors, instead of using the rotation part of F, as it is the case for **FS** (tensor reorientation) and DWI normalization [51, 152].

In the next chapter, the construction of statistical atlases for the case of DW-MRI data is introduced, along with some interesting statistical models applicable to our case that will be used to compare their performance against the proposed statistical approaches.

Chapter 4

DW-MRI Data Statistical Analysis - a Review

Generally in medical imaging, the term "*atlas*" (*i.e.* collection of maps) refers to an anatomical $3D$ representation of an area/organ (*e.g.* brain). An atlas is constructed by accumulating data of one or more subjects in a common coordinate system [145]. Moreover, a map can be useful, as a standard-space template, in the alignment of many subjects/measurements on the same space, for example in order to compare individuals/populations statistically [145]. Of course, aligning different subjects to a common space requires a mandatory and specific step of reorientation depending on the type of the data (*e.g.* raw DWI data, tensors etc. see section 3.2). In DW-MRI, an atlas can contain information concerning the brain's structure (*e.g.* FA template, fiber tracts templates etc., such as in the JHU atlas [114, 115]).

As the number of considered patients is increasing, statistical atlases may be devised by extracting patterns characterizing a particular property/disease etc. [145]. For example, these patterns can be calculated given a set of individual DW-MRI data that can be split into two groups, the normal population (or control group) and the abnormal population (or pathological or testing group) that contains patients of a specific disease (common between all abnormal individuals). Statistical atlases capture the variability of specific patterns in each population and are useful to determine **biomarkers**. For the rest of the dissertation we will refer to the term "*probabilistic atlas*" simply as "*atlas*".

In the case where the number of individuals is large enough to completely capture the variability of both populations, then the project of constructing atlases related to a disease, by measuring the variability of the groups, is linked to the biomarker extraction problem. This is addressed through **population *versus* population** comparisons. On the other hand, when the data are sparse (often happening in the case of the abnormal population) or when there is no common disease patterns between patients, it is hard (or meaningless) to construct the disease's atlas, but it is possible to compare the state of each abnormal individual to the normal population *via* an **individual *versus* normal population** test.

4.1 Categories of DW-MRI Data Analysis

Voxels (or region of voxels in the brain) which are affected by the disease and appear significantly different (*e.g.* lesions) in comparison to the normal population are called biomarkers. Those biomarkers are usually determined by performing populations comparisons (normal *versus* abnormal population). A collection of such biomarkers offers a powerful tool to the physicians, which can aid them to extract diagnostic and prognostic factors of predisposition of the disorder (not only in already known regions of lesions, but also in potential new regions) so that for example patient's treatment can precede the expansion of the inflammation. In addition, they can be useful to monitor the patient's condition.

Populations comparisons (and individual against normal population comparisons) can be performed either by voxel-based, or by ROI-based, or by tract-based analysis, similarly as mentioned in [92] for the DTI case. Historically, **analysis based on region of interest (ROI)**, such as anatomical volumes (*e.g.* tapetum, hippocampus etc.), or geometrical shapes (*e.g.* rectangular, ellipsoidal etc.) were first developed. One advantage of ROI-based analysis is the sensitivity to slight variations (especially for small ROIs) [92]. Ideally, it is applicable when the study is related to particular areas of the brain that can be defined easily and when there is no limit to the computation time. On the contrary, it is typically affected by inter-observer variability [54, 107, 136], which can be reduced by a manual positioning of the ROIs of a single person, but cannot totally be eliminated [136]. Nowadays, the existence of templates with ROIs assists us to reduce more that effect. Once the ROIs are defined, we can calculate the standard deviation and the mean of our measurements (*e.g.* FA images) or we can perform histogram analysis. In general, due to the fact that large ROIs tend to reduce the standard deviation, ROI analysis is proposed for the detection of subtle differences in well-defined small ROIs [92]. More examples can be found in the references of Park *et al.* [125].

Voxel-based analysis (VB) was originally proposed to compare the mean grey matter volume of two populations by Ashburner and Friston [7] in 2000. It is based on spatially registering all individual datasets into a common template, in order to calculate statistics voxel-by-voxel in an unbiased way. According to Park *et al.* [125], VB analysis is more exploratory and suitable for identifying new areas with lesions without any prior knowledge of their existence. Foong *et al.* [58] in 2002 and Buchel *et al.* [32] in 2004 applied VB analysis on DTI. The main advantage of VB is that it does not require any prior knowledge for the localization of the disease, since specific areas with significant differences will be automatically extracted [92]. On the other hand, the selection of a common template and any registration errors left may affect the quality of the results.

Tract-based analysis (TB) does not explore the whole human brain, but it is assisted with user pre-defined tracts to locate voxels with lesions. Pagani *et al.* in 2005 [123] used Diffusion Tensor MRI tractography to construct a probability map for the pyramidal tract by measuring the changes in MD and FA images in patients with early multiple sclerosis (MS). In the same year, Gong *et al.* [70] identified the cingulum *via* DTI tractography. One year later, Lin *et al.* [100] located the pyramidal tracts in order to perform quantitative analysis based on DTI measurements (such as FA images, primary diffusivity, or transverse diffusivity based on the eigenanalysis of the tensor model) for neuromyelitis optica (NMO) disease. Jones *et al.* presented in [93] that registering the individual data in a common template can be avoided. On the contrary, brain atrophy can possibly effect the results [91].

4.2 Recent Related Work

Atlases can be classified into two groups: a) scalar-based statistical atlases and b) statistical atlases on multidimensional data.

Methods belonging to the first category exploit the information provided by scalar measurements, such as fractional anisotropy (FA) and apparent diffusion coefficient (ADC) as employed in [53], or possibly mean diffusivity (MD), relative anisotropy (RA) [16], scalars derived from T2 models, and possibly generalized anisotropy (GA) [122] or generalized fractional anisotropy (GFA) [159] for HOTs. Moreover, tract-based spatial statistics (TBSS) was proposed in 2006 [144] and is available in FSL [89]. TBSS calculates voxelwise statistics on FA across the estimated skeletons (tracts). Furthermore, working in the same direction by processing scalar measurements, Ghosh *et al.* in 2012 [65], inspired by the work of [20, 22], expanded the proposed invariants of [61] to T2 and T4 models.

In the second group of statistical atlases, multidimensional informative models, more complex than scalar images, were taken into consideration. For instance, in 2005, Daurignac *et al.* [43] designed disease-specific probabilistic atlases in order to study alcoholism and to identify patterns of functional and structural lesions due to alcoholism using MRI and DTI. In the same year, Schwartzman *et al.* [139] proposed a method for voxelwise analysis calculating F statistics to address the problem of populations comparison, by studying the principal eigenvector of the T2 model modelled by the bipolar Watson distribution on the unit sphere [103]. Their statistical test verifies whether both populations have the same mean direction (derived from the bipolar Watson distribution). False discovery rate (FDR) was used to overcome the multiple comparisons problem and to correct the false positive detections. In 2008, Schwartzman *et al.* [140] extended their previously mentioned idea by calculating T statistics and assuming global parametrization of their statistical test (common for all the voxels). Spatial smoothing was used to reduce locally the noise variance and increase the effectiveness of FDR analysis.

In 2007, Whitcher *et al.* [174] proposed a set of non-parametric and parametric multivariate tests for populations comparisons that could benefit from the whole information included in T2 models. In this way, they have shown that processing T2s using Log-Euclidean metrics could extract more differences than working with scalar images (*e.g.* FA). In addition, Commowick *et al.* in 2008 [38] proposed a concept for statistical comparison of individual against normal population in MS disease relying on T2 models, too. The authors presented how to compute an unbiased atlas of T2s derived from the set of normal individuals, through

DTI normalization and finite strain T2 reorientation (FS - see section 3.5.1) followed by tensor resampling and averaging using Log-Euclidean metrics in order to obtain the mean T2 healthy template. Finally, they compared statistically each abnormal individual to the mean normal template, by calculating z-scores (*i.e.* Mahalanobis distances) and their corresponding p-values. In 2013, Osborne *et al.* [120] presented a non-parametric bootstrap method for two-sample tests applied to DTI on homogeneous Riemannian manifolds. Osborne *et al.* tested the equality of the generalized Frobenious means of the two populations on the space of symmetric positive matrices (*e.g.* T2 matrix).

Moreover, multivariate regression models based on T2 coefficients [178] and general linear models on T2 coefficients [28] (both including covariates, such as age and gender) have been proposed recently (2009 and 2014, respectively). In 2014, Naylor *et al.* [116] proposed two different concepts for voxelwise analysis of multiple MRI modalities. Their first method is based on fitting multiple univariate linear regression models (one for each modality) and the second approach is described by a single multivariate linear regression model (without assuming independence of the modalities). The multivariate linear regression model appeared to be more efficient than fitting multiple univariate linear regression models.

In early 2015, Caruyer and Verma [33] proposed to study the coefficients of the SH representation of ADC profiles based on HARDI data by computing 12 (for rank-4 SHs) or 25 (for rank-6 SHs) rotationally invariant markers in order to better describe the WM of the brain. Although all these invariants are informative, it is hard to physically explain them.

Furthermore, in 2011 Ingalhalikar *et al.* [86] proposed a high-dimensional non-linear SVM classification methodology for regional features extracted from DTI data. This approach can be also used to assign a probabilistic abnormality score per patient (*i.e.* individual *vs* normal population). Application to autism spectrum disorder (ASD) was presented. In the same year, Bloy *et al.* [25] extended the idea of [86], by using the following variations. Firstly, diffusion ODF (dODF) models [158] were defined (*via* spherical deconvolution of the DW-MRI data, without any prior tensor model, or SH etc.), instead of DTI. For each individual, several ROIs with homogeneous WM structure were determined and orientation invariant features of each ROI's average fODF are included into a feature vector. To continue, PCA is used to reduce linearly the dimension of the data and a linear SVM classifier is trained on the resulting coefficients. Lastly, the trained SVM classifier calculates a probabilistic score per testing individual referring to its likelihood given each group. In addition, in 2012 Bloy *et al.* [26] used again the dODF model of [158] in order to construct a WM dODF atlas that consists of automatically clustered regions according to the homogeneity of the embedded fiber structure and orientation. In the same year, Grigis *et al.* [72] presented their longitudinal study on NMO and MS diseases, formulated as population comparisons, by detecting statistical differences in DWI signals using a multivariate statistical test based on bootstrap technique. In March 2015, Commowick *et al.* [39] generalized and extended their idea presented in [38], for voxelwise individual *versus* normal population statistical analysis applied to MS disease, by altering the DTI model (limited to describe single and not crossing fibers) with the orientation distribution function (ODF) produced by any HOT model. Although, statistical comparisons of any given abnormal ODF (formed as a vector) against any/all normal ODFs are possible and straightforward, they chose to increase the robustness of the method by performing principal component analysis (PCA), in order to treat any artifact or registration error left in the ODF values, prior to statistical analysis.

Another innovative idea is included in methods which determine the manifold of their data and perform geodesic analysis (in non-Euclidean or Riemannian spaces), such as [56, 109, 129] for tensor models or [49, 50, 67, 68] for ODF profiles (derived from HARDI data) represented as PDF functions defined on the unit sphere without any need of fitting tensor models, or SHs etc. In this direction, Verma *et al.* [164, 165] performed voxelwise T2 statistical analysis, firstly by introducing the concept of estimating the non-linear submanifold that 2nd order tensors lie on (*via* dimensionality reduction - Isomap [154] and the estimated geodesic T2 distances), and secondly, by applying multivariate statistics (such as, the Hotelling T^2 test) on the estimated Euclidean submanifold. An analytical description of this method is presented in subsection 4.3.1.

Finally, populations comparisons can be set by solving classification problems (with two classes, *i.e.* the normal class and the abnormal class, or more). Training a classifier and then evaluating its performance by measuring the generalization error (GE) in unseen data can provide evidence if the two populations are similar or not. A variety of classifiers can be found in the literature such as linear (*e.g.* perceptron), quadratic, non-linear (*e.g.* SVM), non-parametric or parametric statistical classifiers (*e.g.* k-NN, Decision Trees, Random Forests, Bayesian). For the purposes of this dissertation, Random Forest Classifiers will be studied and tested (for more information see subsections 4.3.3, 6.3.4).

To sum up, it seems that tensor model analysis is more efficient than working with scalar images. Moreover, reducing the dimension of the working space is really interesting and assists to calculate statistics robustly. To the best of our knowledge, statistical analysis based on T4 models do not exist in literature. As a result, in this study we will focus on voxel-based statistical atlases that encapsulate and compare the information provided by T4 tensor models.

4.3 Application of a Suitable Test

To begin with, prior to the selection of a suitable test, we should specify the problem's type. Is it the "populations comparison" problem, or the "individual *versus* normal group" approach in order to compare profiles (ODF/diffusion)?

According to Verma *et al.* [165], lesions induced by white matter disorders are better captured by statistical population comparisons. Given two quite large groups of healthy and pathological individuals, we can construct control patterns on which we can measure the variability of any pathological population.

In this section, three selected approaches are further discussed and presented that will be needed in chapter 6 to be compared with the proposed method of chapter 5. The first one can be found in the literature, the second is synthesized by the combination of two separated methods and the third one is constructed by using the theory of Random Forests.

4.3.1 Representing and Analyzing T2s in a Reduced Space

Verma *et al.* [165] proposed their method for voxel-based analysis. They noted that when working with diffusion tensors (T2s), which lie on a non-linear submanifold of the space $\mathbb{R}^6$, it is not safe to apply directly any standard statistical model, due to the fact that T2s do not follow multivariate Gaussian distributions in $\mathbb{R}^6$. In addition, most of DTI statistical analysis at that time was based on scalar (*e.g.* FA images) or vectorial diffusivity measurements (*e.g.* principal diffusivity) that require prior knowledge of the pathologically affected areas (which is not always available) and they do not enclose any information about the embedded submanifold structure or they do not introduce any geodesic distance metrics. A clear visual comparison between the Euclidean distance between two T2s (*i.e.* green dotted line) and the underlying geodesic distance (*i.e.* red curve) is shown in figure 4.1.

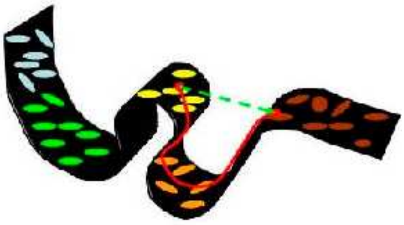


FIGURE 4.1: The choice of a proper geodesic distance is mandatory. Each ellipsoid corresponds to a T2 located in a non-linear submanifold in $\mathbb{R}^6$. The green dotted line shows the Euclidean distance between T2s, which does not correspond to the ideal geodesic distance (denoted by the red line) which is calculated throughout the T2 submanifold. Image obtained by [165].

For all previously mentioned reasons, Verma *et al.* [165] chose to estimate a dense reduced subspace of the initial sparse T2 space, by using proper geodesic distances between T2 tensor models, in order to apply standard multivariate statistics and measure the mean and the variance of the populations in a reduced dimensional space.

Determination of the reduced space using Isomap

Verma *et al.* [165] chose to determine the reduced space by using Isomap [154], a non-linear dimensionality reduction technique which combines the well-known multidimensional scaling (MDS) method [96] with graph theory, and particularly shortest paths calculations based on geodesic distances. In this case, Isomap calculates an inter-point distance matrix for all couples of individuals *via* a graph representation and a shortest path calculation (*via* the Floyd-Warshall algorithm [40]) which includes the k neighbors' tensor Riemannian distances (*e.g.* tensor metric appeared in [56]) for each couple of individuals. Once the distance matrix is calculated, MDS is used to determine the reduced space in a way that distances between points in the reduced space mimic the corresponding distances in the initial space. Finally, by plotting the residual variance for different values of dimension ($d \leq 6$), the resulting L-shaped elbow plot concluded that working in the reduced dimension of 2 is safe for the T2 case.

Hotelling T^2 statistical test

At this moment that all individual points (normal and abnormal) are mapped into the reduced Euclidean space, the Hotelling T^2 test can be applied to compare the means of the two groups (since groups' covariances are assumed to be equal). For further information about that specific statistical test, the reader is referred to Appendix A. Although calculating statistics in a reduced dense space can be significantly helpful, comparing the means of two Gaussian distributions can hide important information or can lead to significant errors that could probably be avoided if someone could study the whole embedded information derived by assuming a more complex model per group. We will set some comparisons in this direction in the chapters with the experimental results that will be presented later.

4.3.2 Analyzing the Inter-point Distance Matrix in High Dimensional Space

Hotelling T^2 test and many parametric statistical tests are based on strong assumptions for data modelling. For example, Hotelling T^2 test works properly only for Gaussian distributed data with the same covariance for both groups and uses the location (*i.e.* mean) information to compare the two groups. To circumvent these restrictive assumptions, multivariate nonparametric tests have been proposed in the literature (see the review [118]). However, many parametric and nonparametric tests may not be used when the dimension of data is greater than their number, or may show poor performance for high dimensional data.

One of the most important steps in populations comparison is the calculation of an inter-point distance matrix Δ which contains the distances between all possible combinations of individual data (M normal and N abnormal). As a result, Δ corresponds to a symmetric matrix of size $(M+N) \times (M+N)$ with zero diagonal elements. Working with tensors (as is the case in this dissertation), the computation of each non-diagonal element can be accomplished by using one of the tensor metrics presented in section 2.3.4, for instance equation 2.30 which is defined independently of the order of the tensor. A well-defined and suitable tensor metric yields matrix Δ with significant information about the separability between the two populations.

Statistic of interest

Working in this direction, Biswas and Ghosh [23] proposed in 2014 a nonparametric two-sample test, applicable to high dimensional data and formed on any type of inter-point distances. Moreover, a variety of other statistical tests based on inter-point distances can be found also in [23].

Given two populations related to two distributions F, G with $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\} \sim F$ and $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\} \sim G$ *i.i.d.* observations of each distribution, Biswas and Ghosh proposed to reject the Null Hypothesis (*i.e.* $F = G$) for high values of the following statistic:

$$T_{M,N} = \|\widehat{\boldsymbol{\mu}}_{\Delta_F} - \widehat{\boldsymbol{\mu}}_{\Delta_G}\|^2, \tag{4.1}$$

where $\|.\|$ is the Euclidean norm and $\widehat{\boldsymbol{\mu}}_{\Delta_i}$, $i \in \{F, G\}$ represent 2D vectors defined as follows:

$$\widehat{\boldsymbol{\mu}}_{\Delta_F} = [\widehat{\mu}_{FF}\,,\ \widehat{\mu}_{FG}]\,, \qquad \widehat{\boldsymbol{\mu}}_{\Delta_G} = [\widehat{\mu}_{FG}\,,\ \widehat{\mu}_{GG}]\,, \tag{4.2}$$

embedding the following coefficients:

$$\widehat{\mu}_{FF} = \binom{M}{2}^{-1} \sum_{i=1}^{M} \sum_{j=i+1}^{M} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_\Delta\,,\ \widehat{\mu}_{GG} = \binom{N}{2}^{-1} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \|\boldsymbol{y}_i - \boldsymbol{y}_j\|_\Delta$$

$$\widehat{\mu}_{FG} = (MN)^{-1} \sum_{i=1}^{M} \sum_{j=1}^{N} \|\boldsymbol{x}_i - \boldsymbol{y}_j\|_\Delta. \tag{4.3}$$

The symbolic norm $\|.\|_\Delta$ in equations 4.3 corresponds to the inter-point distance elements of matrix Δ. Qualitatively, $\widehat{\mu}_{ii}$, $i \in \{F, G\}$ represents the average inter-point distance between all couples of points in the same group i, whereas $\widehat{\mu}_{FG}$ represents the average inter-point distance between all couples of points belonging to different groups.

Biswas' and Ghosh's statistical test is rotation invariant, free of distributional assumptions, simple and computationally efficient [23]. It is applicable to any high dimensional data which provide a distance function. In high dimensional well-posed (*i.e.* the number of the measurements is equal or greater than the number of unknowns) problems, this test outperforms other tests for the location, the scale and the scale and location problems. Furthermore, it performs well in low number of samples, even in ill-posed problems, where many methods perform poorly and practically cannot be used. Finally, it can be generalized for multi-sample tests (see [23]).

Proposed extension: Highest Probability Density (HPD) interval of the embedded p-value

At this point that the statistic of interest is selected, we propose to estimate the p-value (associated to the statistic), along with its corresponding HPD interval, using a permutation test, especially designed for distance matrices, proposed by Reiss *et al.* in 2010 [133]. Working with the same concept of label shuffling, Reiss *et al.* proposed the following way to permute the entries of a given inter-point distance Δ:

$$\Delta_\rho = E_\rho \Delta E_\rho{}^T, \tag{4.4}$$

where $E_\rho = \left(e_{\rho(1)} \dots e_{\rho(M+N)}\right)^T$, corresponding to a $(M+N) \times (M+N)$ matrix containing the permutation binary vectors of permutation function $\rho()$. More precisely, e_i is defined as the (M+N)-dimensional vector with 1 in the i-th element and 0 in the other elements. Therefore, a new statistic $T^{(\pi)}_{M,N}$ can be calculated for each permutation iteration π.

Furthermore, given the real statistic $T^{(R)}_{M,N}$ (corresponding to the initial distance matrix Δ) and a set of Π statistics $T^{(\pi)}_{M,N}$, $1 \leqslant \pi \leqslant \Pi$ resulting from the permutation test, we can approximate the p-value $\nu^\star$ of getting statistics equal or more extreme than the reference one (the effective statistic) through randomly sampling the distribution of the statistic.

$$\begin{aligned} \nu^\star = P\left(T^{(\pi)}_{M,N} \geqslant T^{(R)}_{M,N}\right) &= \int q_\pi \left(T^{(\pi)}_{M,N} \geqslant T^{(R)}_{M,N}\right)\ p\left(T^{(\pi)}_{M,N} = x\right)\ \mathtt{d}x \\ &\simeq \frac{1}{\Pi} \sum_{\pi=1}^{\Pi} q_\pi \left(T^{(\pi)}_{M,N} \geqslant T^{(R)}_{M,N}\right) = \widehat{\nu}, \end{aligned} \tag{4.5}$$

where $p\left(T^{(\pi)}_{M,N} = x\right)$ is the distribution of $T^{(\pi)}_{M,N}$ under the hypothesis of indiscernible populations.

In addition, the credibility interval of the p-value $\nu^\star$ will be determined as follows. A set of binary values $Q = \{q_1, \dots, q_\Pi\}$ is collected by comparing the reference statistic $T^{(R)}_{M,N}$ with each $T^{(\pi)}_{M,N}$. If $T^{(\pi)}_{M,N} \geqslant T^{(R)}_{M,N}$ then $q_\pi(.) = 1$ and 0 otherwise. As a consequence, the set Q contains samples of the Bernoulli distribution of parameter the unknown p-value $\nu^\star$. Assuming as a prior that $\nu^\star$ is uniformly distributed, we can estimate the credibility interval of the unknown p-value $\nu^\star$, as the 99% of the *a posteriori* mass of $p(\nu \mid q_1, \dots, q_\Pi) \sim Beta(\alpha+1, \beta+1)$, where α is the number of 1's and β the number of 0's in Q.

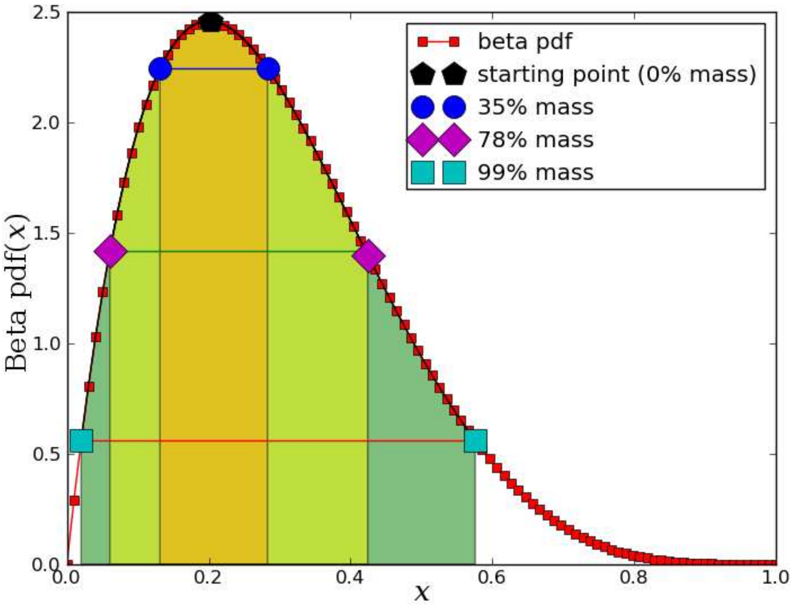


FIGURE 4.2: Illustration of the HPD interval estimation by calculating the 99% of the distribution mass with the aid of Dichotomy.

That interval is also known as Highest Probability Density (HPD) interval and is extracted *via* Dichotomy (actually, three steps of Dichotomy, one for the PDF value and two for the lower and upper bound of the HPD interval) as presented in figure 4.2.

At this point we have to mention that HPD interval's length depends on the number of the permutations. If someone wants to reduce that length, more iterations should be included in the permutation test.

4.3.3 Analyzing Classification Errors using Random Forest Classifiers

Populations comparisons can be solved as classification problems. For example, if two populations are very mixed together (*i.e.* similar), then the classification task fails to discriminate the two groups.

A Brief Introduction

Random Forest (RF corresponds to its registered trademark) proposed by Leo Breiman in 2001 [30] is a versatile and competitive tool for statistics. Classification, regression, abnormality detection (*via* density estimation), manifold learning, semi-supervised applications, such as image segmentation etc. are some among several of its applications [41, 42].

A RF is an ensemble of T randomly trained *Decision Trees* (DT) [31]. The attribute of **randomness** is gained *via* bootstrap sampling, since each tree uses two random **subsets** of **samples**, one random (bootstrapped) subset for training and the other one containing the unseen samples for testing (evaluation). In this way, the out-of-bag (OOB) score (*e.g.* the percentage of correctly classified points), or the converse of OOB score, the generalization error (GE), can be calculated on previously unseen data (*i.e.* data in testing set). Moreover, each tree node is split according to a random **subset** of the **sample's features**.

Another interesting key point that should be mentioned is the fact that each decision tree is expanded and left **unpruned** without having to deal with over-fitting problems due to the randomness property.

To continue, a RF is defined by a set of parameters (for example, the number T of the trees, the maximum depth D of the trees etc.). For a full list of them the reader is referred to [41, 42], where the effect of each parameter is also tested and discussed.

Every DT will be separately evaluated on its own testing set, producing a local GE. The global representative GE of the RF will be the average of the individual GEs.

Introducing a point $\mathbf{v}$ into the RF means that every DT will examine the point $\mathbf{v}$ concluding to an individual prediction. The final, single, prediction of the RF is obtained by averaging (or voting) the individual tree predictions.

For the purposes of this study, we will focus on RF Classifiers. A simple example of RF classifier can be found in figure 4.3.

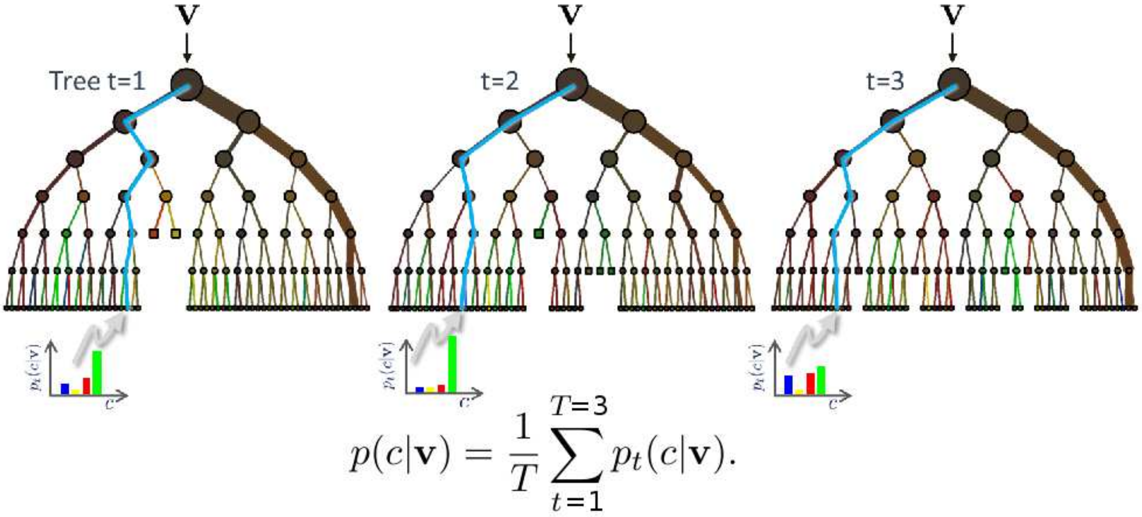


FIGURE 4.3: A RF classifier with three DTs. Assigning sample **v** to a class is achieved by averaging the three posterior (not binary) $p_t(c|\mathbf{v})$ of each tree. Image appearing in [41].

Applications of Random Forests in DW-MRI

RF were very recently applied to DW-MRI. For example, DTI tractography analysis using RFs for the MS disease were proposed in 2011 [95] and in 2012 [106]. Predictors of clinical impairment between the cerebellum and the cerebral hemispheres were analysed using RFs on DTI data for MS disease in 2014 [131]. In addition, the effect of lesions detected using DTI WM tractography to global disease severity and cognitive and behavioural disturbances were studied using RFs for the progressive supranuclear palsy disease in 2014 [1]. In the same year, RFs were applied to measure the independent contribution of the **FA** and the **MD** to language impairment detection in a TB analysis of pediatric epilepsy patients [124].

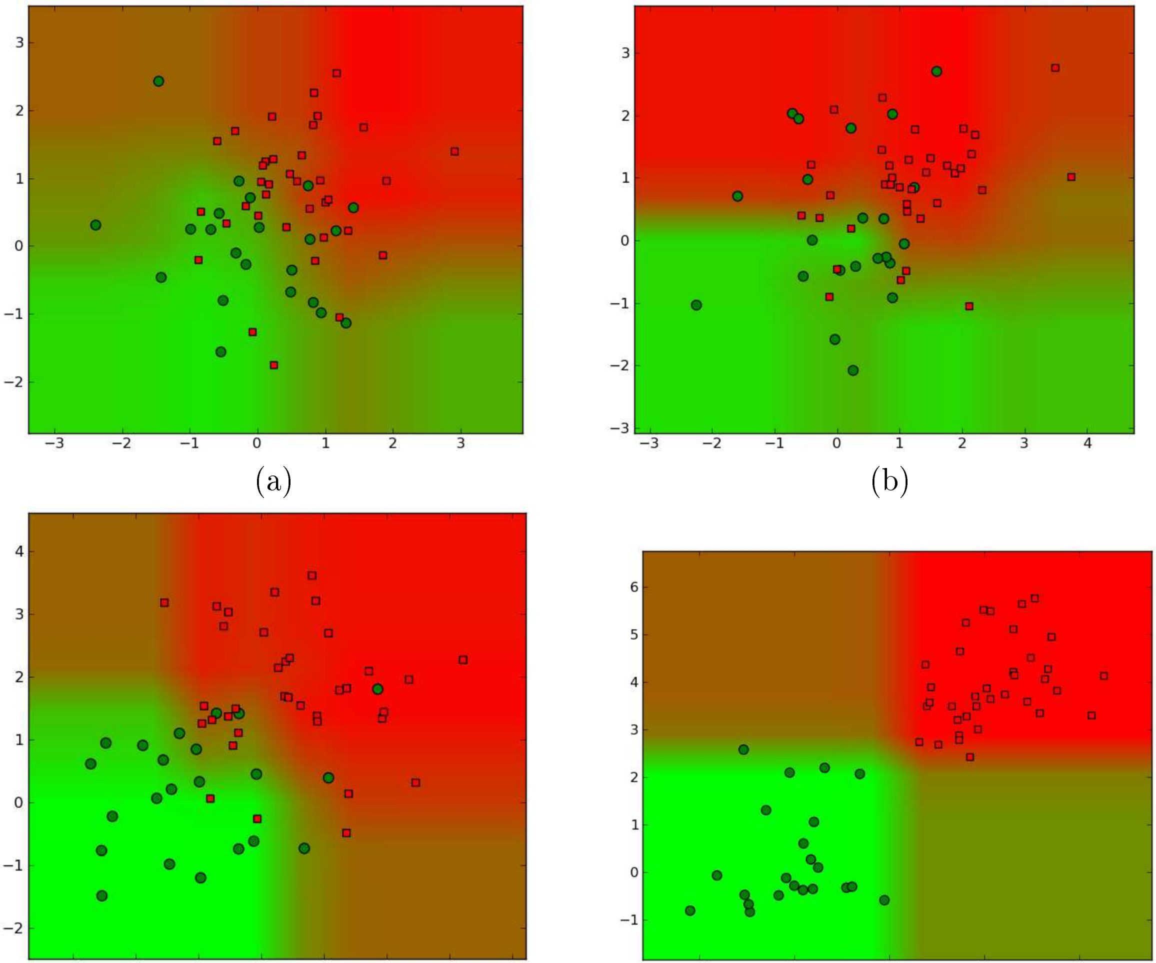
Furthermore, a certain group of studies benefits from the RF ability to analyse high dimensional data. Multivariate RFs on multimodal MR Imaging (Diffusion, Perfusion, and Spectroscopy) were defined to determine which criteria could differentiate between grades and genotypes of oligodendroglial tumors in 2013 [52]. Lesion segmentations for ischemic stroke were implemented using RF Classifiers on multimodal MRI data, such as TI-weighted, T2-weighted, FLAIR, and ADC MRI images in 2014 [108], or functional, anatomical and diffusion data for stroke in 2015 [37]. Segmentation of thalamus (a crucial task during the evaluation of many brain disorders) using T1-weighted MRI data and nuclear parcellation on DW-MRI data were proposed in 2014 [150]. In the latter study, FA images, fiber orientation and connectivity between the thalamus and the cortical lobes were selected as feature vectors to define RF Classifiers. A few more applications are the following [3, 29, 81].

Finally, the existence of numerous recent approaches based on RFs for DW-MRI, proposed for problems similar to ours, highlights the potentiality of the method to achieve competitive results.

Statistical Model

Populations comparison can be achieved using Random Forest Classifiers. A RF Classifier measures the generalization error (GE) of all unseen data and concludes to similarity if GE is very high, while on the contrary it results to dissimilarity when GE is very low.

In practice, it can be explained that if the classification task fails to discriminate the two groups (*i.e.* similar groups), it will result into high values in GE. If the two populations are well separated (*i.e.* dissimilar), the GE is low. Figure 4.4 depicts the RF classifications of four different synthetic cases, given a RF with 500 decision trees and maximum tree depth 4. As the distance between the two populations is increased, the GE is reduced.



(a) (b)

(c) (d)

FIGURE 4.4: Examples of resulting classifications given a RF with 500 decision trees and maximum tree depth equal to 4, on synthetic $2D$ data. (a) GE = 0.38, (b) GE = 0.29, (c) GE = 0.14 and (d) GE = 0.02. We can see that as the two populations are moving away from each other, (a)→(d), the GE decreases. The colorful background is related to $p(c|\mathbf{v})$, labeled as RGB color vectors $[p(c = \text{"}\texttt{abnormal}\text{"}|\mathbf{v}), p(c = \text{"}\texttt{normal}\text{"}|\mathbf{v}), 0]$. Areas with green color correspond to points with higher probability to belong to the normal group than belonging to the abnormal one, and red otherwise. Finally, brown levels correspond to areas with high levels of uncertainty.

As a consequence, we will consider the GE as our statistic of choice, and moreover we will also calculate its own HPD interval, similarly to what was done for the previous statistical method. In fact, there is no need to define a permutation test in order to estimate the HPD interval of the GE, since we can benefit from the randomness of each tree's testing set (in size and in samples).

Given a set of T trees along with their corresponding testing sets V_t, $1 \leq t \leq T$ and $|V_t| = l_t$ (*i.e.* the number of samples in V_t), a list $Q = \{q_{11}, \dots, q_{l_1 1}, \dots, q_{1T}, \dots, q_{l_T T}\}$ can be constructed, containing binary values for all testing points of each tree, where q_{it} is equal to 1 if the point i is misclassified in tree t and 0 otherwise. Therefore, GE is related to the probability ϱ, which is the probability of getting a point i misclassified in tree t:

$$P(q_{it}) = \begin{cases} \varrho & \text{, if } q_{it} = 1, \\ 1 - \varrho & \text{, if } q_{it} = 0, \end{cases} \tag{4.6}$$

Moreover, two useful lists can be obtained, $L_{wrong} = \{w_1, \dots, w_T\}$ and $L_{length} = \{l_1, \dots, l_T\}$, where w_t corresponds to the number of wrong classifications and l_t the total number of tested samples in tree t, $1 \leq t \leq T$. In this way, we can find the HPD interval of the GE, as the 99% of the *a posteriori* mass of $P(GE|Q, \varrho) \sim Beta(\varrho; \alpha + 1, L_{tot} - \alpha + 1)$ (see fig. 4.2)

$$P(GE \mid Q, \varrho) \propto \varrho^{\alpha}(1 - \varrho)^{L_{tot} - \alpha}, \tag{4.7}$$

where $\alpha = \sum_i^T w_i$ contains the total number of wrong classifications in $L_{tot} = \sum_i^T l_i$ total number of tested samples.

4.4 Partial Conclusion

In this chapter, the concept of statistical atlases was presented along with references to existing state-of-art techniques. In addition, three methods to perform statistical analysis were selected and further described (each one for specific reasons) with the ultimate goal to compare them with the proposed methods in the next chapters. The first one, although based on Hotelling T^2 test (a quite weak statistical test due to its draconian assumptions), benefits from the idea of working in a reduced space, which is favorable and promising. The second method was selected as an innovative method which can handle high dimensional data with no need to perform dimensionality reduction. Finally, the third approach is based on RF classifiers, which are assumed to be flexible in statistical calculations even in high dimension. Experimental results for all of these methods are included in chapter 6.

In the next chapter, the definition of the proposed statistical model, applicable to population *versus* population comparison, will be presented and discussed.

Chapter 5

Population *VS* Population Comparison: Proposed Statistical Model for T4s

In this chapter an innovative statistical model is proposed, aiming to offer efficient early diagnosis, prognosis and patient follow up for a given disease. The proposed statistical test gains in sensitivity due to the use of the T4 fODF parameterization to describe the data, which produces better representation of the fiber structure than the T2 fODF model. Moreover, due to the high dimension of the data, we selected to reduce the dimension, in order to calculate robust statistics in a dense space. In this thesis, application of the proposed statistical test to populations comparison was achieved in the case of the NMO disease (experimental results are available in chapter 6).

5.1 Preliminary Steps

Before calculating the statistics, some preliminary steps are required. Firstly, data normalization is a crucial step in order to transform all data into the same space. Secondly, T4 fODF model parametrization and the definition of a proper metric that compares two T4 fODF profiles are chosen to calculate an inter-point distance matrix. Finally, the considered part of neighboring information is defined, assisting to eliminate any registration error left.

5.1.1 Selected Data Normalization

Mentioning the term "normalization", we refer to the procedure of transferring all data to the same common reference space. To achieve that, one can calculate a spatial transformation (*e.g.* linear, non-linear) and apply it on each datum. Due to the fact that our data (tensor images or raw DW-MRI data in our case) contain information about the orientation of the diffusion at each point, their spatial registration will result into important errors on not accounting of the new underlying fiber's structure. To bypass that obstacle, different normalization techniques were proposed in the literature, as described in section 3.2. DWI normalization uses only the rotation part of the whole transformation, but tensor image normalization can produce much more significant errors (due to the registration step on tensor's coefficients), than normalizing the raw DWI data. In other words, a small variation in the tensor coefficients can produce a totally different fODF profile, while a small variation in the DWI signal may not affect notably the estimated tensor model.

Initially in this study, tensor normalization (see section 3.3) was chosen in the absence of a non-linear DWI normalization method. In the recent years, a competitive method for DWI normalization was proposed in 2013 [51] and our initial choice altered to the new approach. Calculating the FA image of each datum (by fitting voxelwise T2 tensor models) allowed us to estimate the non-linear transformations between each FA image (eq. 2.8) in the initial space and a reference FA template (*e.g.* JHU-FA-2mm) by calling standard procedures from the FSL toolkit [89]. The estimated transformation is applied on the DWI data and their reorientation is achieved by extracting the local rotation component of the transformation and applying it to the spatially registered DWI data with the aid of a proper FSL patch available on the web and proposed in [51].

5.1.2 Selection of a T4 fODF Parametrization, a Proper Metric and the fODF Patches

Equation 2.24 in chapter 2 shows that a fODF function $f(\mathbf{g})$ can be described, for example by a T4 tensor model. In this study, the coefficients of the T4 fODF model were estimated according to [172] by minimizing a quadratic cost function under the positivity constraint of the estimated model.

T4 tensor model parameters belong to $\mathbb{R}^{15}$. We must define a proper metric in that space. Choosing an Euclidean distance between the coefficients of two T4s would not be pertinent, since it would not take into consideration properly the whole information provided by the corresponding profiles defined on the sphere. For example, a small variation in $\mathbb{R}^{15}$ will not correspond to the same variation in the profiles on the sphere. A proper distance should be defined between two positive valued fODF functions on the $3D$ unit sphere. Experiments showed that selecting an Euclidean distance integrated on the sphere is not sufficient, since for example, differences between diffusivity values 10^2 and 10^3 will overlap any significant differences between values 10^{-2} and 10^{-3}. For that reason, log-based distances became popular, such as the log-Euclidean distance for T2s in [6], or distances described in [153].

For the purposes of this study, we chose to work with the distance defined in equation 2.30. An illustrative comparison between the Euclidean distance (L_2 norm) and the proposed one (both integrated on the sphere) is depicted in figure 5.1. Each point corresponds to a T4 fODF tensor represented in the $2D$ reduced space by using an inter-point distance matrix containing distances between all possible pairs of tensors in a feature extracting algorithm (*e.g.* Isomap [154]). The red point represents an outlier corresponding to an fODF related to a T4 model for which one of the coefficients was divided by 60. In the case of the proposed distance, the outlier is recognized and penalised by separating it from the mass of the other tensors (*i.e.* ideal case), while in the case of the Euclidean distance the outlier is not well separated.

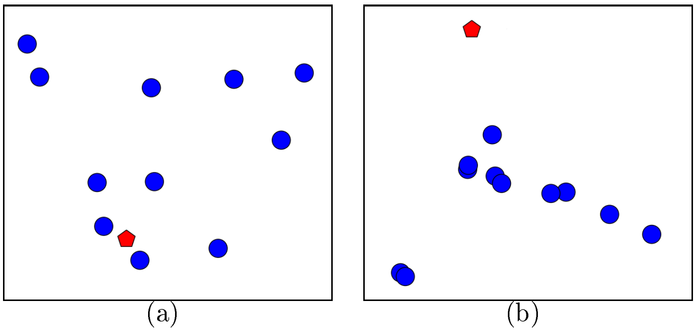


FIGURE 5.1: Comparison between the Euclidean distance and the proposed distance of eq.2.30. The outlier (red point) should be separated from the mass, as is the case for the selected distance and not for the Euclidean distance.

At this point, we should note that voxelwise distances are not efficient enough, especially when data contain potential registration errors. In order to deal with these registration errors, many approaches choose to smooth data, risking to lose useful information. Another option could be to rely on skeletons of white matter bundles of fiber [144], but the risk of losing information is still not totally eliminated. Introducing information contained in the neighborhood of each voxel is another solution, which seems more suitable and less risky. As a result, we selected to sum all the distances between two selected $3 \times 3 \times 3$ patches per voxel. The selected patches are defined by searching among all the possible coupled patches in two $5 \times 5 \times 5$ neighborhoods for the one that minimizes the sum of all the distances in the smaller patches. Figure 5.2, illustrates the idea of searching all possible 3×3 coupled patches in a 5×5 neighborhood, for simplicity.

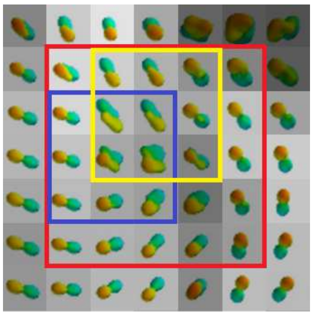


FIGURE 5.2: The choice of the best 3×3 patches between two 5×5 neighborhoods (one per individual dataset that is included in the comparison).

5.2 Feature Extraction (ISOMAP)

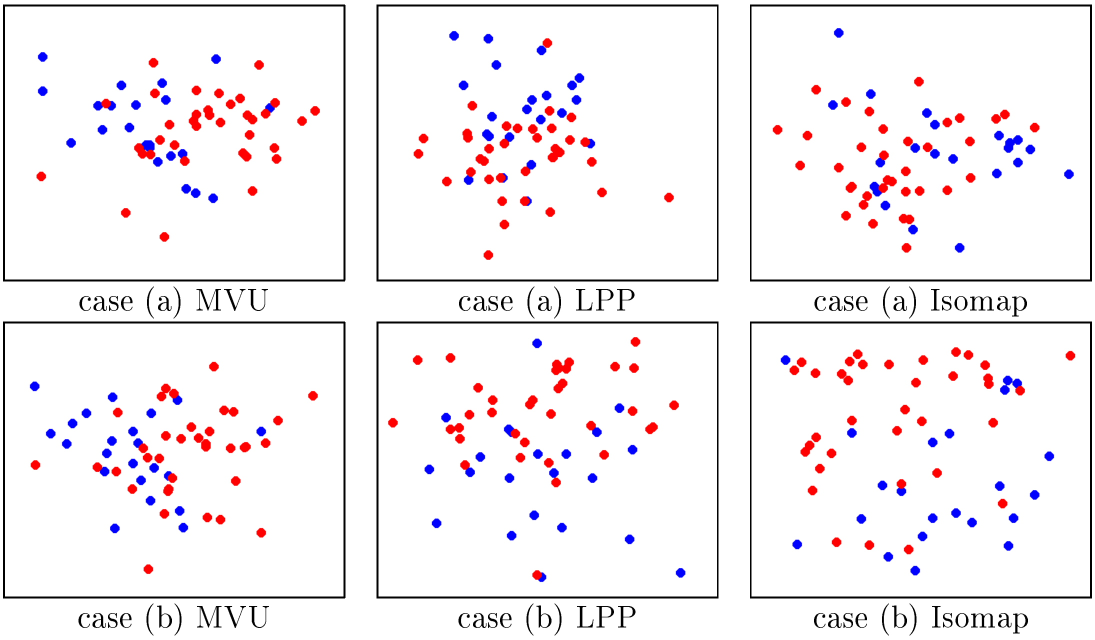
For each voxel (or a $5 \times 5 \times 5$ neighborhood referring to that voxel), an inter-point distance matrix can be calculated. The dimension of the current space is 15 (for a single T4) or

$5 \times 5 \times 5 \times 15 = 1875$ for a $5 \times 5 \times 5$ neighborhood of voxels. Assuming a number of data (corresponding to pathological and healthy patients) close to the order of tens or hundreds, it can be directly concluded that this space will be sparse and not suitable for calculating coherent statistics with robustness. As a consequence, reducing the dimension of the space will provide a more dense space to work with.

To address the fact that the data lie on a manifold and consider geodesic distances, we resort to non-linear dimension reduction techniques. We tested several methods, such as Isomap [154], maximum variance unfolding (MVU) [171] and locality preserving projection (LPP) [77, 78]. We chose to work with Isomap, similarly to [165] since in general, there were no particular differences from a discriminative point of view (see table 5.1 and figure 5.3).

case / p-value's HPD interval	**Non-linear Method** **MVU**	 **LPP**	 **ISOMAP**
(a) Almost Dissimilarity / $[0.04, 0.065]$	9	12	9
(b) Clear Dissimilarity / $[0.0, 0.0046]$	12	9	10

Table 5.1: Comparison between different non-linear methods, such as MVU, LPP and ISOMAP. The table contains the number of wrong classifications between 22 individuals of the normal group and 36 individuals of the abnormal group (totally 58 samples). Two characteristic cases are extracted from real data calculations, where the estimation of the p-values signify in (a) almost different populations while in (b) significantly different populations. As we can see Isomap gave slightly better solution in average than the other approaches.


case (a) MVU case (a) LPP case (a) Isomap

case (b) MVU case (b) LPP case (b) Isomap

Figure 5.3: Plots of the 2D reduced space for the 58 samples of the two cases presented in Table 5.1 which contains 22 normal individuals (denoted by the blue color) and 36 abnormal individuals (corresponding to red color).

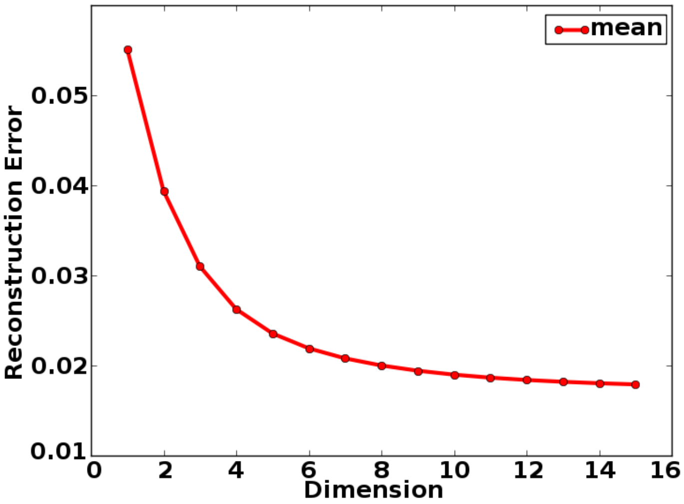


FIGURE 5.4: Scree plot of the reconstruction error in function to the reduced dimension. On the basis of this plot, we chose to work with a reduced space of dimension 2 for the T4 case (similar to the T2 case as in [165]).

Other important reasons which explain the need to perform dimensionality reduction are the following. Firstly, working in high dimensional spaces is not efficient, since they can be sparsely filled with data. Secondly, lacking of statistical tests for multidimensional data is another obstacle. As a consequence, constructing a reduced dense space where statistics can be calculated robustly is crucial. This reduced space should be built by incorporating geodesic distances in the initial space. Verma *et al.* [165] highlighted that specific property in their study for T2 models (see figure 4.1). In this direction, Isomap correlates the structure of the data points in the initial space, with the new structure of the points in the reduced space by retaining the geodesic inter-point distances. In other words, the distances between points in the reduced space mimic the distances between the corresponding points in the initial space. Scree plot presented in figure 5.4 lead us to chose $d = 2$ as the dimension of the reduced space, which is similar to the T2 case that Verma *et al.* analyzed.

5.3 Statistic of Interest

Given a normal and an abnormal population consisting of several points in the reduced space (determined in the previous step), we propose to represent the probability densities p_i, $i = 1, 2$ of each population i by using kernel density estimation, assuming that each point is associated to a Gaussian kernel [76]. Each population's PDF has one free parameter, the covariance matrix connected to each population's point. It is determined with the aid of Scott's rule [142].

At this point that populations' PDFs are defined, the selection of a proper metric to compare the two populations is needed. In our first attempt, we chose to work with the popular symmetrized Kullback-Leibler divergence. Unfortunately, there is no closed formula known for the case of mixture of Gaussian distributions associated to the definition of kernel density estimations. As a result, only a numerical approximation could give an answer to our problem and in practice it is very time consuming.

As a consequence, for all previously mentioned reasons, we chose to build our statistical model on another measure of discrepancy $\mathcal{P}$ proposed by Sfikas *et al.* in [143], suitable to compare mixtures of Gaussian distributions, which provides efficient results in much less computational time. In fact, our experimentations indicated that $\mathcal{P}$ is calculated 150 times faster than our previous numerical approximation (*e.g.* with 5600 samples of the working space) of the symmetrized Kullback-Leibler (sKL) divergence. To be more precise, computation of $\mathcal{P}$ needed 10^{-3} minutes to compare two PDFs in a single permutation, in a computer with 4 processors at 3.20 GHz and 8 GB of RAM memory. On the other hand, sKL required 0.15 minutes in a single permutation (meaning that the corresponding computational time for a typical set of 1000 permutations will be 150 minutes). In addition, the discrepancy $\mathcal{P}$ and the sKL approximation produced equivalent characterizations of the populations (similar/dissimilar).

The discrepancy $\mathcal{P}$ is our statistic of interest and is generally defined as follows [151], [143]:

$$\mathcal{P}\left(p_1, p_2\right) = -\log\left[\frac{2\displaystyle\int p_1\left(\boldsymbol{x}\right) p_2\left(\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x}}{\displaystyle\int \left(p_1(\boldsymbol{x})\right)^2 \mathrm{d}\boldsymbol{x} + \int \left(p_2(\boldsymbol{x})\right)^2 \mathrm{d}\boldsymbol{x}}\right]. \tag{5.1}$$

This metric is symmetrical and becomes zero if p_1, p_2 are equal and positive otherwise.

According to [143], if we consider the following two mixtures of Gaussian distributions (one for each population, normal and abnormal):

$$\left\{\begin{array}{lcl} p_a\left(\boldsymbol{x}\right) & = & \displaystyle\sum_{i=1}^{I} \pi_i^{(a)} \mathcal{N}\left(\boldsymbol{x}; \boldsymbol{\mu}_i^{(a)}, \boldsymbol{\Sigma}_i^{(a)}\right) \\ p_b\left(\boldsymbol{x}\right) & = & \displaystyle\sum_{j=1}^{J} \pi_j^{(b)} \mathcal{N}\left(\boldsymbol{x}; \boldsymbol{\mu}_j^{(b)}, \boldsymbol{\Sigma}_j^{(b)}\right) \end{array}\right., \tag{5.2}$$

we will come up with the following straightforward relation

$$\int p_a\left(\boldsymbol{x}\right)\ p_b\left(\boldsymbol{x}\right)\ \mathrm{d}\boldsymbol{x} = \sum_{i,j} \pi_i^{(a)} \pi_j^{(b)} \sqrt{\frac{\exp\left(k\right)\left|\boldsymbol{V}\right|}{\left(2\pi\right)^{N_x}\left|\boldsymbol{\Sigma}_i^{(a)}\right|\left|\boldsymbol{\Sigma}_j^{(b)}\right|}}, \tag{5.3}$$

where N_x is the dimension of the point $\boldsymbol{x}$ and $\boldsymbol{V}, k, \boldsymbol{\mu}$ are given by the following equations

$$\boldsymbol{V} = \left(\left(\boldsymbol{\Sigma}_i^{(a)}\right)^{-1} + \left(\boldsymbol{\Sigma}_j^{(b)}\right)^{-1}\right)^{-1}, \tag{5.4}$$

$$k = \boldsymbol{\mu}^T \boldsymbol{V}^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}_i^{(a)T} \left(\boldsymbol{\Sigma}_i^{(a)}\right)^{-1} \boldsymbol{\mu}_i^{(a)} - \boldsymbol{\mu}_j^{(b)T} \left(\boldsymbol{\Sigma}_j^{(b)}\right)^{-1} \boldsymbol{\mu}_j^{(b)}, \tag{5.5}$$

$$\boldsymbol{\mu} = \boldsymbol{V}\left(\left(\boldsymbol{\Sigma}_i^{(a)}\right)^{-1} \boldsymbol{\mu}_i^{(a)} + \left(\boldsymbol{\Sigma}_j^{(b)}\right)^{-1} \boldsymbol{\mu}_j^{(b)}\right), \tag{5.6}$$

for a in $\{1,2\}$ and b in $\{1,2\}$. Indices i, j, a, b were omitted from $\boldsymbol{V}$, k, and $\boldsymbol{\mu}$ for simplicity. In addition, the previous formulas may be simplified, since $\boldsymbol{\Sigma}_i^{(a)}$ does not depend on i and similarly, since $\boldsymbol{\Sigma}_j^{(b)}$ does not depend on j.

5.4 Estimation of the p-value and its credibility interval

For the purposes of performing population comparison, the distance $\mathcal{P}$ is handled as a statistic. Therefore, our statistical problem is formulated as checking if the reference distance $\mathcal{P}_0$ (*i.e.* the one that compares the populations' PDFs given the initial-true labelling of the individual points) is an extreme value with respect to the distribution $p(\mathcal{P})$ of the distances.

Due to the fact that $p(\mathcal{P})$ is generally unknown, we produce random samples of the distribution of $\mathcal{P}$ *via* permutation testing. We select to approximate the probability of getting distances $\mathcal{P}$ equal or greater than the reference one $\mathcal{P}_0$, $\mathrm{P}(\mathcal{P}_\pi \geqslant \mathcal{P}_0)$, with the aid of Monte Carlo distance samples $\mathcal{P}_\pi$, under the Null Hypothesis that both populations are indiscernible. This probability is also known as p-value and we will denote it $\nu^\star$ and is equal to:

$$\nu^\star = P\left(\mathcal{P}_\pi \geqslant \mathcal{P}_0\right) = \int_{\mathcal{P}_0}^{+\infty} p\left(\mathcal{P}\right) \ \mathtt{d}\mathcal{P}. \tag{5.7}$$

Many statistical approaches are complacent to calculate only a single value for estimating $\nu^\star$. For example, a possible solution $\widehat{\nu}$ could be to divide the number of random distances that satisfy $\mathcal{P}_\pi \geqslant \mathcal{P}_0$, with the total number $\varPi$ of label shufflings.

On the other hand, each estimated p-value $\widehat{\nu}$ has its own credibility interval, which depends on the number of permutations. For that reason, in order to have insight in the precision reached, we propose to calculate a credibility interval of the approximated p-value $\widehat{\nu}$. The length of the interval can be reduced by increasing the number of the label shufflings.

The steps of the proposed permutation test are the following: initially, we permute the labels of the individual points $\varPi$ times and at each iteration of the permutation test we estimate the value of $\mathcal{P}$. Each comparison of $\mathcal{P}_\pi$ with the reference distance $\mathcal{P}_0$ produces a binary value q_π ($q_\pi = 1$ if $\mathcal{P}_\pi \geqslant \mathcal{P}_0$, and 0 otherwise). Each q_π consists in a sample of the Bernoulli distribution parametrized by the unknown p-value $\nu^\star$. At this point, the problem is translated into estimating $\nu^\star$ by using the binary samples $q_1, \ldots, q_\varPi$.

$$P(q_\pi) = \begin{cases} \nu & \text{, if } q_\pi = 1, \\ 1-\nu & \text{, if } q_\pi = 0. \end{cases} \tag{5.8}$$

Assuming a uniform prior for $\nu^\star$, we can calculate the posterior $p\left(\nu \mid q_1, \ldots, q_\varPi\right)$ which follows a Beta distribution $Beta(\alpha+1, \varPi-\alpha+1)$, where α is the number of 1's in $\{q_1, \ldots, q_\varPi\}$.:

$$P(\nu \mid q_1, \ldots, q_\varPi) \propto \nu^\alpha (1-\nu)^{\varPi-\alpha}, \tag{5.9}$$

The smallest interval enclosing the 99% of the *a posteriori* mass $p(\nu \mid q_1, \ldots, q_\Pi)$, *i.e.* the Highest Probability Density (HPD) interval, is considered to be the interval of ν that we are searching for.

More details about how to calculate the HPD are shown in figure 4.2. The main idea is that we start from the maximum point and by performing dichotomy on both axes we can locate the interval corresponding to the 99% of the mass. In this way, we have two characteristic values for each p-value, the upper and the lower bound of the HPD interval.

A summary of the proposed statistical pipeline can be found in figure 5.5.

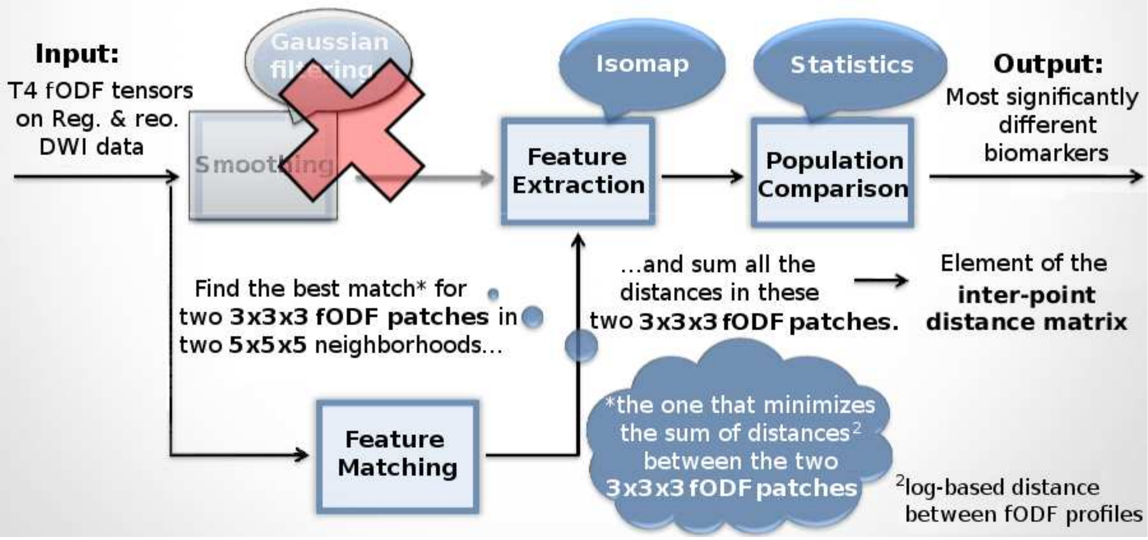


FIGURE 5.5: The steps of the proposed approach.

5.5 Partial Conclusion

A statistical model based on tensors (particularly T4s which present better accuracy than T2s, but any order tensor can be used) for the population comparison problem was proposed in this chapter. After the enumeration of all needed pre-processing steps, the importance of dimensionality reduction was highlighted and explained.

The proposed statistical test is based on the comparison of Gaussian kernel density PDFs by resorting to permutation testing. Moreover, instead of limiting our method in calculating only a single corresponding p-value for each voxel, we continue further by estimating the HPD interval of each p-value. To fastly compare the two kernel density estimations, a rapid and effective discrepancy, proper for mixtures of Gaussian distributions, was derived from the literature and used in this study.

It is common that many methods proposed for statistically medical analysis stop to the point of locating areas with small p-values (assisted by a selected threshold, usually equal to 0.05). In this study, we consider that constructing a list of sorted p-values will extract the most different voxels in the top of that list, since the ranking of each voxel is much more informative than its p-value, especially when we want to compare if different methods give the same result (*e.g.* ranking).

In the following chapter, the experimental evaluation of the proposed method is presented along with many other comparisons to methods presented in chapter 3.

Chapter 6

Group Comparisons: Evaluation on NMO disease and synthetic cases

Applying the proposed statistical approach within a region of interest for a specific disease can reveal an interesting list of p-values sorted in ascending order. It can highlight the most significantly different voxels (*i.e.* biomarkers) in the top of that list, which finally can help us to define regions of interest for each particular disease.

Neuromyelitis optica (NMO) disease, or Devic's Syndrome, is an inflammatory neurodegenerative autoimmune pathology that results in simultaneous Wallerian degeneration in regions which are directly connected to the spinal cord and to the optic nerves. Moreover, NMO causes gradual demyelination due to inflammation in several regions rich in aquaporin-4 of the human brain such as the periventricular, the hypothalamic and the periaqueductal regions and the bottom part of the fourth ventricle. The main symptoms of NMO disease are the optic neuritis and the transverse myelitis that cause blindness and paralysis of the extremities, respectively. Due to the fact that there is no standard cure for NMO, the objective is to stop or delay the evolution of the disease. Moreover, the development of cutting edge tools that could provide early diagnosis and prognosis are crucial.

Population comparisons of Normal and Abnormal groups aid us to extract interesting biomarkers, or regions of them, that will signify that those specific regions are characteristic areas affected by the disease. In this way, we can provide useful information by guiding the doctors through their examination or to properly adjust the patient's treatment.

In order to set our experiments, 22 normal (healthy individuals) and 36 abnormal (pathological individuals) DW-MRI datasets were used. The brain scanning procedure provided us with HARDI data, where 30 non colinear gradient directions (signal is measured twice in each direction) and b-value= $1000\ \mathrm{s/mm}^2$ were used, resulting into images of size of $128 \times 128 \times 41$ and resolution $1.8 \times 1.8 \times 3.5\ \mathrm{mm}^3$. The proposed statistical model was implemented in python/cython, while special routines from the FSL toolkit [89] were used for the registration and reorientation steps. Moreover, all other methods used for comparison were also implemented in python.

6.1 Application of the Proposed Method to the T4 fODF case

Selecting a ROI with 2741 voxels in the brain, such as the left and right anterior limb of the internal capsule, the left and right posterior limb of the internal capsule and the left and right posterior thalamic radiation (including optic radiation), that are already known as directly affected by the NMO disease according to medical knowledge and the literature, allowed us to verify that the proposed method is capable to highlight the region as pathologically affected by NMO, too. Figure 6.1 illustrates the histogram of the obtained p-values (HPD's upper bound), concluding into the characterization of the region as a ROI for NMO, since a peak appears in the bin with the lowest p-values ($[0, 0.05]$). On the other hand, a region with no lesion would produce a "flat" histogram over the interval $[0, 1]$.

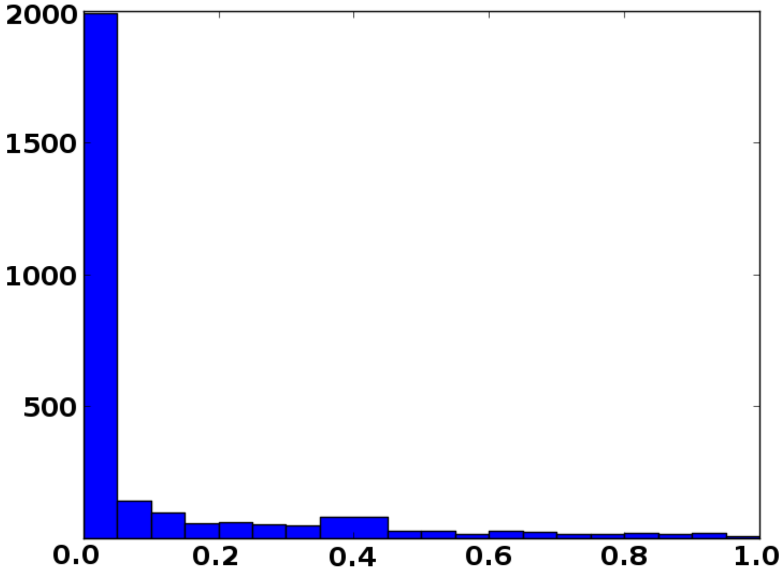


FIGURE 6.1: The histogram of the resulting p-values (HPD's upper bound) of the proposed method applied on T4 models in a ROI with 2741 voxels (bin size = 0.05). The peak in low p-values signifies that the selected region is pathologically affected by NMO disease.

Figure 6.2 depicts three characteristic cases of voxels that were found throughout the experiments. Firstly, the case of getting dissimilar populations when the probability density of the normal population is totally different than the abnormal probability density. In this voxel, if we smooth our data, instead of finding the best match of the fODF patches, the obtained result would signify that the populations are similar. The second case represents an intermediate case that seems to converge to similarity, and finally, the third voxel corresponds to a similar case where both probability densities look almost identical.

Table 6.1 summarizes the resulting HPD intervals of the three presented cases of fig. 6.2. The number of permutations was 1000. The length of each interval can be reduced, by increasing the number of the label shufflings. Each statistical test took 3 minutes on a standard computer. Most of the computation time was devoted to determining the best matched patches in the calculation of the inter-point distance matrix.

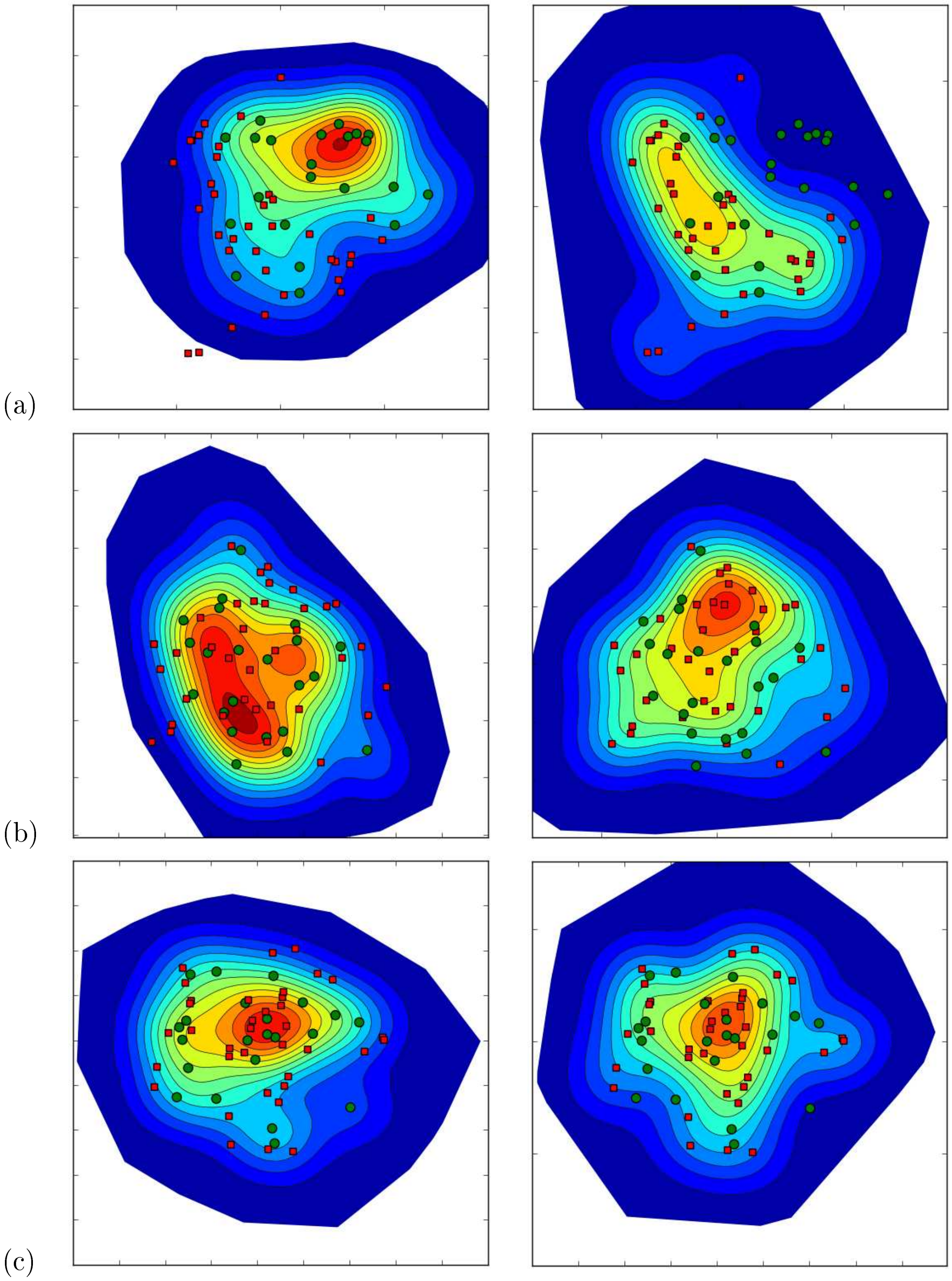


FIGURE 6.2: Visualization of probability densities, based on Gaussian kernel density estimation, in the reduced space. Green points: normal individuals; red squares: abnormal individuals. Presentation of three characteristic cases: (a) dissimilar populations, (b) and (c) similar populations. Left column: representation of the probability density corresponding to the normal population. Right column: representation of the probability density corresponding to the abnormal population. Blue: low density, red: high density. The same color scale is used across all subfigures.

Fig.6.2 Cases	**p-value HPD interval**	**width of interval**	**Populations are**
(a)	$[0, 0.0046]$	0.0046	Dissimilar
(b)	$[0.35, 0.43]$	0.08	Similar
(c)	$[0.988, 0.999]$	0.011	Similar

TABLE 6.1: HPD intervals of the p-values for the cases depicted in Fig. 6.2, by performing 1000 label shufflings.

In the next sections of the chapter, the application of the proposed statistical model on T2 tensors is presented, along with some other benchmark comparisons.

6.2 Application of the Proposed Method to the T2 fODF case

The application of the proposed statistical model on T2 tensors and the comparison with the obtained results of the T4 case was the first point that we would like to test. It is known that the T4 model can describe with much more accuracy a fODF that represents a complex structure of fibers (up to 3 main bundles, see figure 2.9) than the T2 model, and as a result, a more representative model could potentially produce much more biomarkers.

Except from comparing the fODF profiles of the T2 models by using equation 2.30 (similarly to the T4 case), another popular metric exists in the literature for T2 tensor models, the log-Euclidean distance (eq. 2.28). As a consequence, we will present both resulting statistics by constructing the inter-point distance matrix at each voxel using both distances. Then, this matrix will be introduced into the Isomap step, in order to produce the points in the reduced space.

To begin with, figure 6.3 shows the distribution of the obtained p-values (HPD's upper bound) in the T2 case by comparing the T2 fODFs on the unit sphere (the same distance as for T4 case), in the same ROI as in T4 case. Directly, we can notice that T2 fODF case also concludes that the ROI has increased interest as being pathologically affected by NMO. On the other hand, we can see that the peak in the lowest p-values contains less biomarkers in the T2 fODF case than in T4 fODF case, meaning that our initial thought for the amount of biomarkers is validated and T4 fODFs is able to detect biomarkers that cannot be highlighted with T2 fODFs.

To continue, plugging the log-Euclidean distance in our procedure would result into substituting the comparison between fODF profiles with the comparison between the 6 coefficients of the T2 model. The obtained distribution of the resulting p-values can be seen in figure 6.4. Although, in this case a peak in the lowest p-values was obtained too, we observe that the number of biomarkers is reduced a lot, much more than the T2 fODF case. As a consequence, if someone choose to work with T2 models, it will be significantly better to resort to fODF comparisons instead of comparing the tensor's coefficients.

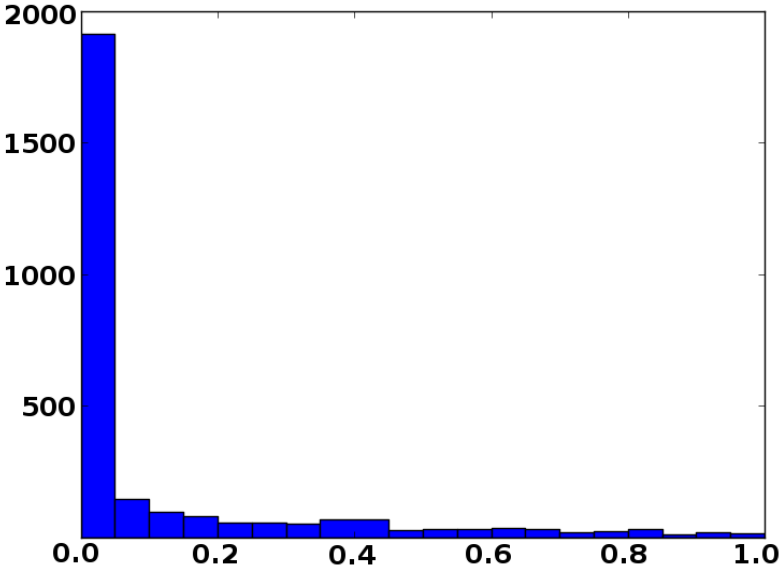


FIGURE 6.3: The histogram of the resulting p-values of the proposed statistical model applied on T2 fODF case in a ROI with 2741 voxels (bin size = 0.05). The peak in low p-values signifies that the selected region is pathologically affected by NMO disease, but it produced less biomarkers than the T4 case.

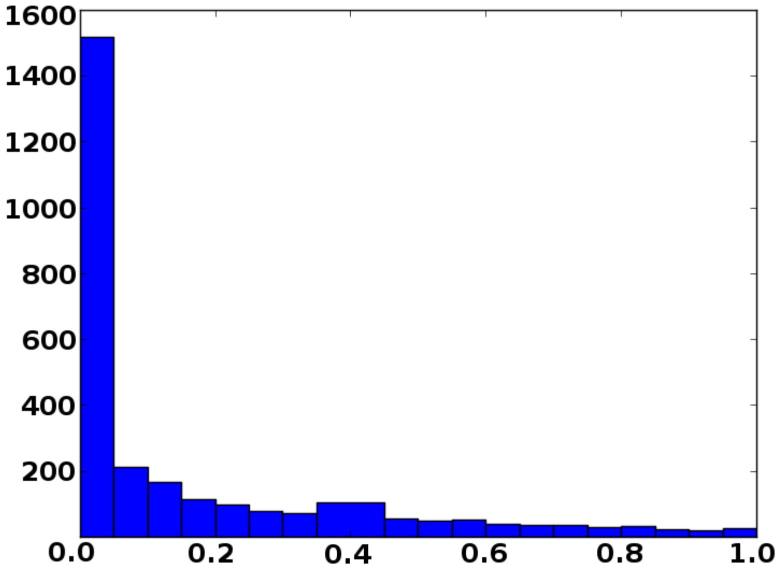
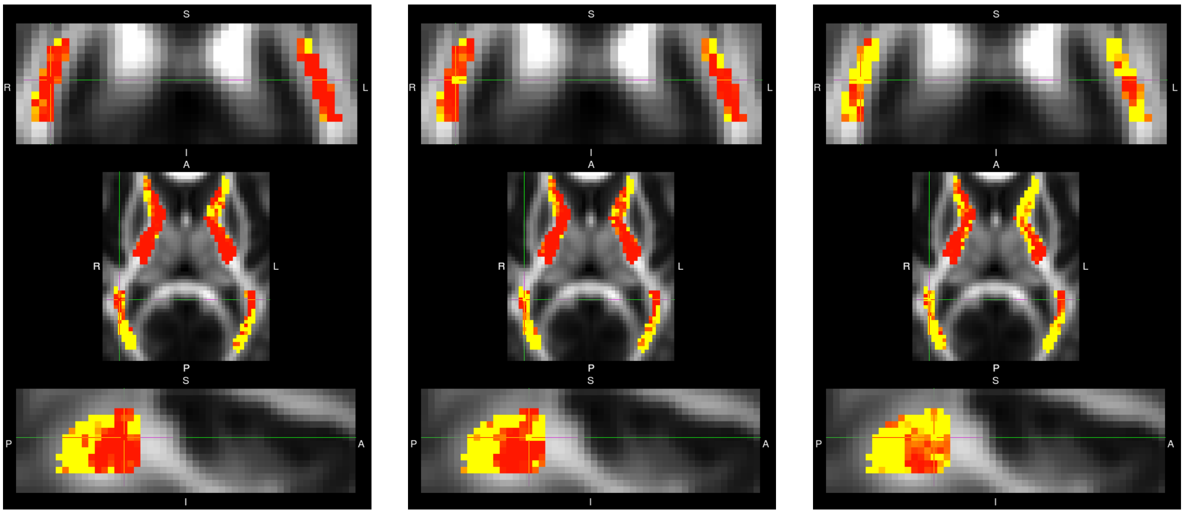


FIGURE 6.4: The histogram of the resulting p-values of the proposed statistical model applied on T2 case using the log-Euclidean distance (eq.2.29) in a ROI with 2741 voxels (bin size = 0.05). The peak in low p-values signifies that the selected region is pathologically affected by NMO, although it produced less biomarkers than the T4 and T2 fODF cases.

Comparisons between T4 and T2 statistics

Supplementary to the histograms of figures 6.1-6.4, figure 6.5 depicts the obtained biomarkers (p-value ≤ 0.05), noted by red color, in a partial view of the total ROI of 2741 voxels (named as ROI 1). A careful look at those three images concludes that working with T4 fODF is more productive and sensitive than T2 models. Moreover, in the T2 case, it is much more effective to process T2 fODFs instead of T2 coefficients (regarding the number of extracted biomarkers).



(a) (b) (c)

FIGURE 6.5: Plot the obtained biomarkers (p-value ≤ 0.05, highlighted by red color) of a particular region (we will refer to it as ROI 1) on the top of a FA template, in (a) T4 fODF case, (b) T2 fODF case and (c) T2 coefficients case. As it can be seen the T4 fODF case produced more biomarkers than the other cases. In addition, working with T2 fODF is much better than T2 coefficients. Top images correspond to coronal views, middle images correspond to axial views and bottom images correspond to sagittal views.

Additionally, since we are interested in comparing the ranking of the biomarkers between different techniques, we came up with the idea of checking if the set of the top $N = 1272$ biomarkers consists of the same voxels in the three different approaches (*i.e.* T4 fODF, T2 fODF, T2 coefficients). The value of N was chosen by the fact that the top 1272 biomarkers had same HPD intervals for the p-values in the T4 case (aiming to work with smaller HPD intervals will eventually lead to smaller N). Working in this direction, a 3-label colorful map can be produced. Voxels appearing in both testing schemes are categorized by the first case (*i.e.* green label), voxels appear only in the list of method (a) (and not in method (b)) will be assigned to the second case (*i.e.* purple label) and finally voxels appear only in method (b) (and not in method (a)) belong to the third case (*i.e.* blue label).

Figures 6.6, 6.7 and 6.8 represent two slices (ROI 1, ROI 2) of the 3-label colorful maps selected by the total 1272 most significantly different voxels extracted from the ROI of the 2741 voxels, in "T4 fODF *vs* T2 fODF", "T4 fODF *vs* T2 coefficients" and "T2 fODF *vs* T2 coefficients" cases, respectively.

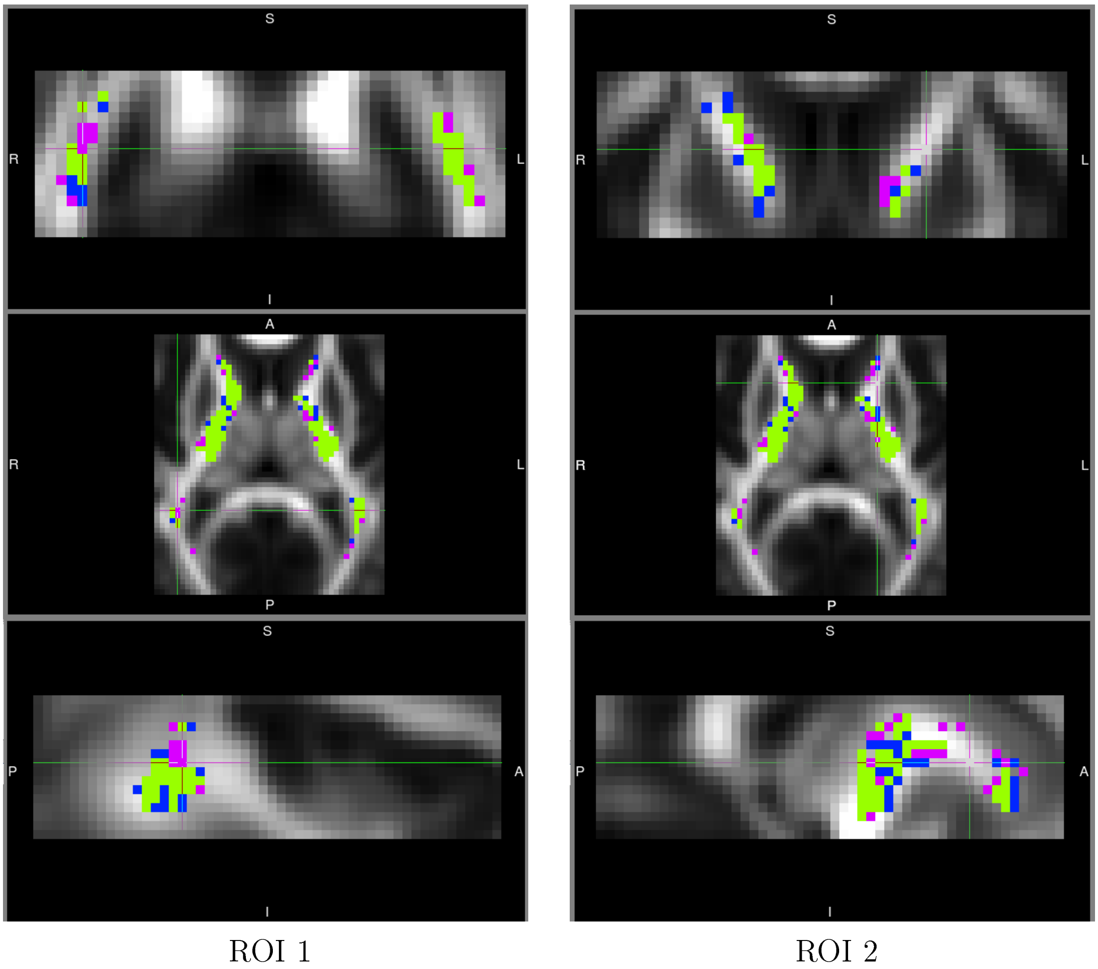


FIGURE 6.6: Comparison of the ranking of T4 fODF statistics with T2 fODF statistics in two different ROIs. Green voxels corresponds to the case of getting both ranks in the top $N = 1272$ dissimilar voxels, purple voxels appeared only in the T4 fODF's top N dissimilar voxels and blue voxels only in the T2 fODF's top N dissimilar voxels. Greyscale images correspond to FA template. Top images correspond to coronal views, middle images correspond to axial views and bottom images correspond to sagittal views.

# of Tested Voxels	**Green (Both cases)**	**Purple (only T4 fODF)**	**Blue (only T2 fODF)**
1500	1044	228	228

TABLE 6.2: Number of voxels with green, purple and blue color of the T4 fODF's *versus* T2 fODF's statistics.

Tables 6.2, 6.3 and 6.4 contain the number of voxels belonging to each of the three labels (green, purple and blue). In the case of "T4 fODF *vs* T2 fODF" 17.9% (= 228/1272) of the top $N = 1272$ voxels are different, while the corresponding percentages of "T4 fODF *vs* T2 coefficients", "T2 fODF *vs* T2 coefficients" are equal to 25.9% (= 330/1272) and 21.6% (= 275/1272), respectively.

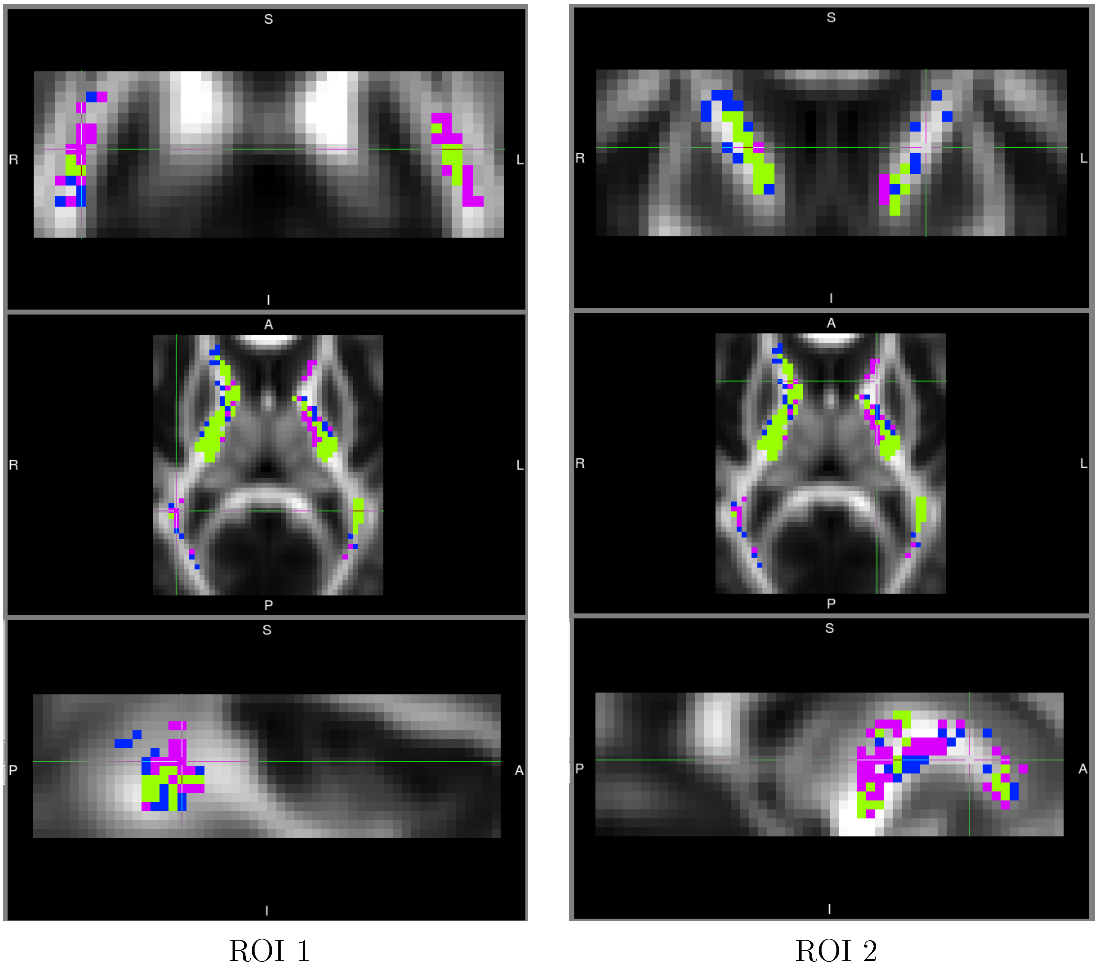


FIGURE 6.7: Comparison of the ranking of T4 fODF statistics with T2 coefficients statistics in two different ROIs. Green voxels corresponds to the case of getting both ranks in the top $N = 1272$ dissimilar voxels, purple voxels appeared only in the T4 fODF's top N dissimilar voxels and blue voxels only in the T2 coefficients' top N dissimilar voxels. Greyscale images correspond to FA template. Top images correspond to coronal views, middle images correspond to axial views and bottom images correspond to sagittal views.

# of Tested Voxels	**Green (Both cases)**	**Purple (only T4 fODF)**	**Blue (only T2 coefficients)**
1602	942	330	330

TABLE 6.3: Number of voxels with green, purple and blue color of the T4 fODF's *versus* T2 coefficients' statistics.

The above percentages of differences between T4 and T2 tensor models triggered us to examine the reasons why those differences happen. A more detailed examination of the obtained sorted lists of p-values indicated us voxels where the decision (Similar/Dissimilar populations) does not agree between T4 and T2 models.

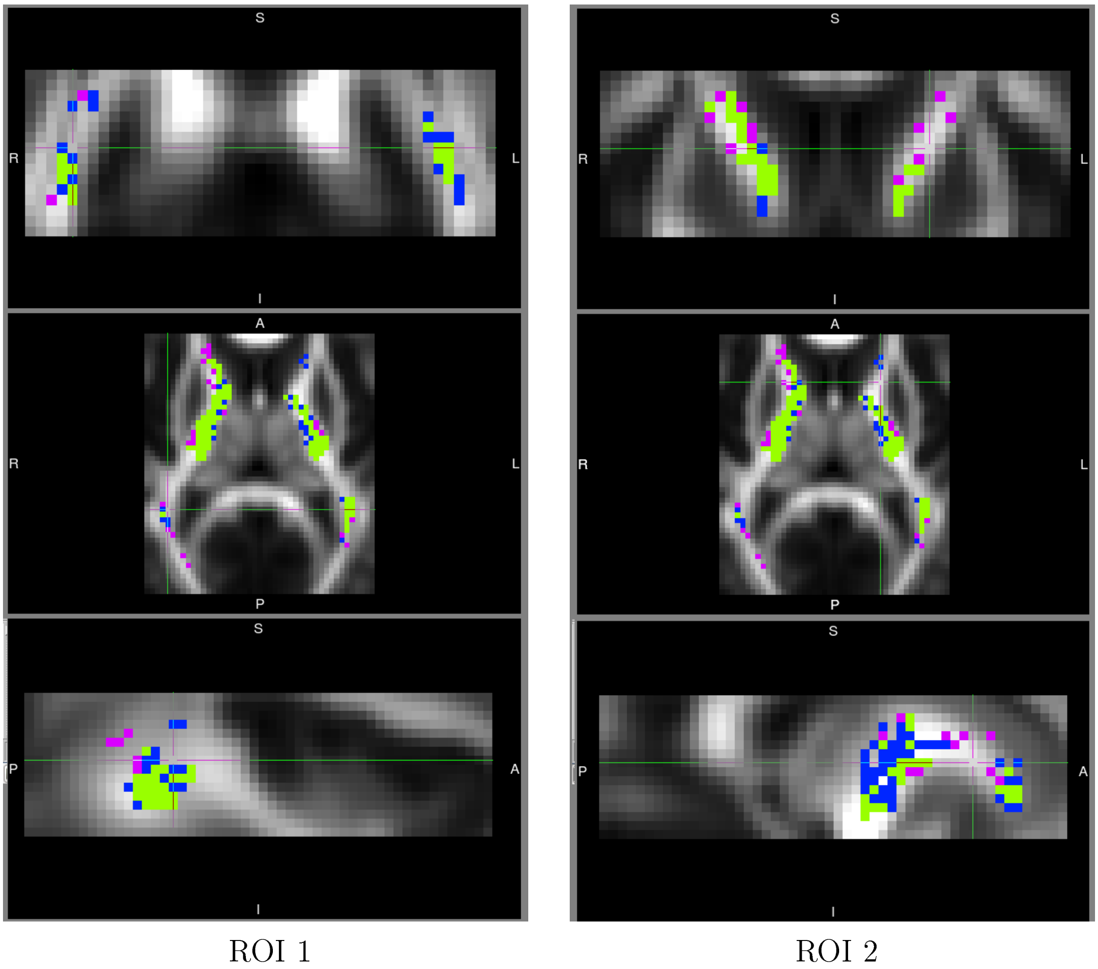


FIGURE 6.8: Comparison of the ranking of T2 fODF statistics with T2 coefficients statistics in 2 different ROIs. Green voxels corresponds to the case of getting both ranks in the top $N = 1272$ dissimilar voxels, purple voxels appeared only in the T2 coefficients' top N dissimilar voxels and blue voxels only in the T2 fODF's top N dissimilar voxels. Greyscale images correspond to FA template. Top images correspond to coronal views, middle images correspond to axial views and bottom images correspond to sagittal views.

# of Tested Voxels	**Green (Both cases)**	**Purple (only T2 coefficients)**	**Blue (only T2 fODF)**
1547	997	275	275

TABLE 6.4: Number of voxels with green, purple and blue color of the T2 fODF's *versus* T2 coefficients' statistics.

Reliability of T2 Statistics

Comparing the final decision (similar/dissimilar) of our statistical test between T4 and T2 cases, we notice that disagreements occurred in some cases. Testing a set of those voxels in disagreement, by building the reduced space with the assistance of an inter-point distance matrix which contains the L_1 differences of the model's residuals (T4 and T2, separately) integrated on the sphere and then applying the proposed statistical model on that reduced space, we came up with the conclusion that the T4 model gave the correct answer, since the information of dissimilarity is contained in the T2 residuals. In other words, the less efficient and less accurate T2 model can influence the conclusion of the statistical test.

6.3 Other Comparisons

In order to better evaluate the precision of the T4 model in comparison to the T2 model, and in addition, to test the proposed statistical model against other approaches, we set the following comparisons on both synthetic and real data.

6.3.1 T2 and T4 fODF models' contributions to populations comparisons - Evaluation on synthetic data

Knowing that T2 models provide less accurate descriptions for crossing fibers, resulting into more close to isotropic or spherical representations, we came up with the construction of the following synthetic test.

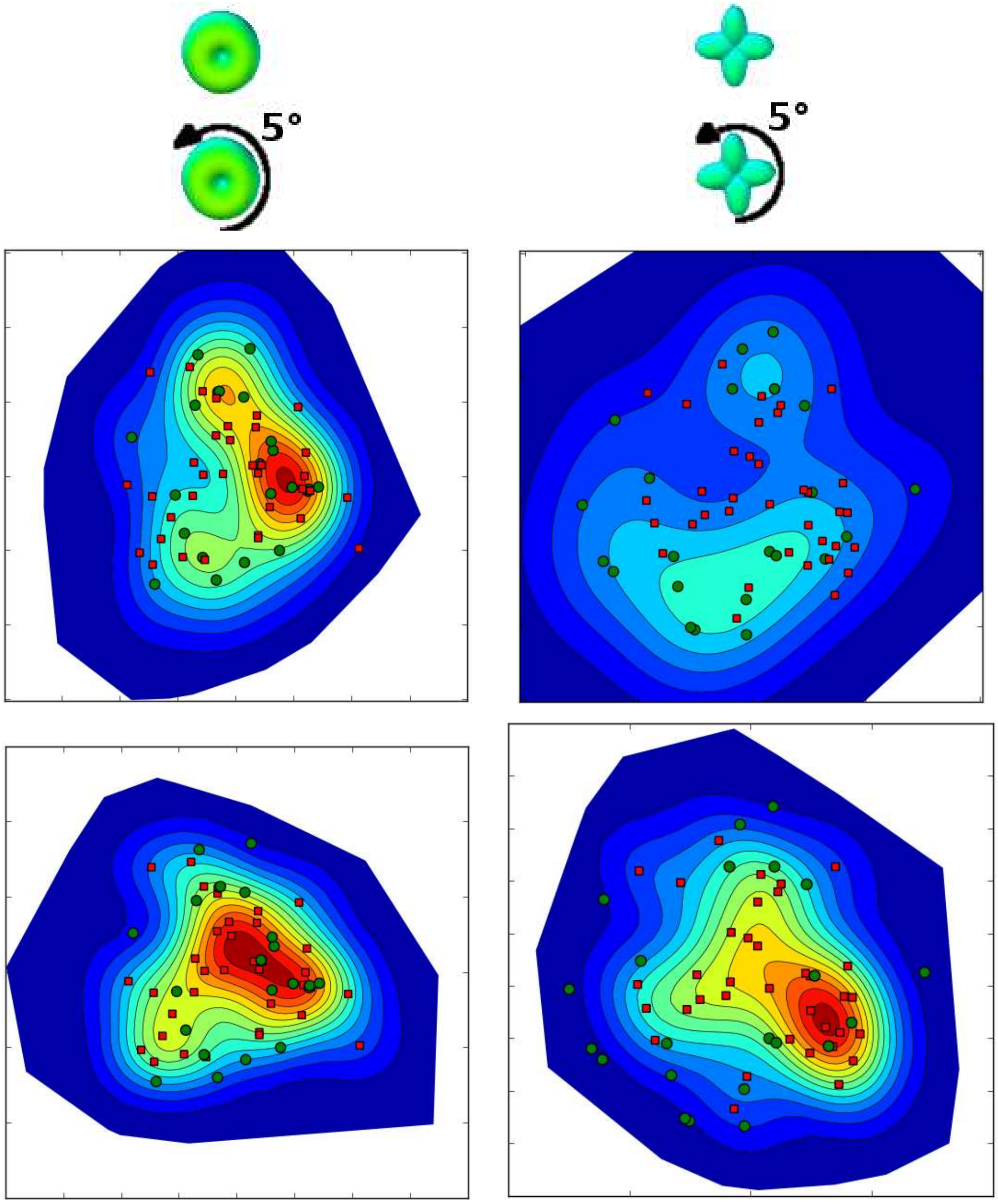
Two tensor templates (one for each population, first two rows of figure 6.9) were constructed, representing two orthogonally crossing fibers of the same diffusion (on left the T2, on right the T4). The whole abnormal tensor template is rotated by 5 degrees in comparison to the normal one (with the aid of [10]). We defined a set of 22 normal and 36 abnormal individuals by adding Gaussian noise to the reference tensor templates.

Application of the proposed statistical model on T4 fODF profiles resulted into characterizing the two populations as dissimilar, while on the other hand T2 fODF profiles concluded to similarity (see figure 6.9 and the corresponding p-values in the left part of table 6.5).

case	**fODF p-value**	**fODF Decision**	**Residual p-value**	**Residual Decision**
T2	$[0.73, 0.80]$	**Similar**	$[0, 0.0046]$	**Dissimilar**
T4	$[0.0058, 0.025]$	Dissimilar	$[0.061, 0.091]$	Similar

Table 6.5: Calculated HPD intervals of p-values on fODF profiles (left part) and on the models' residuals (right part) of the case presented in fig.6.9 and 6.10, along with the characterization of each case. T4 fODF concluded to dissimilar populations, while T2 fODF to similar. The disagreement occurred due to T2 model's residuals.

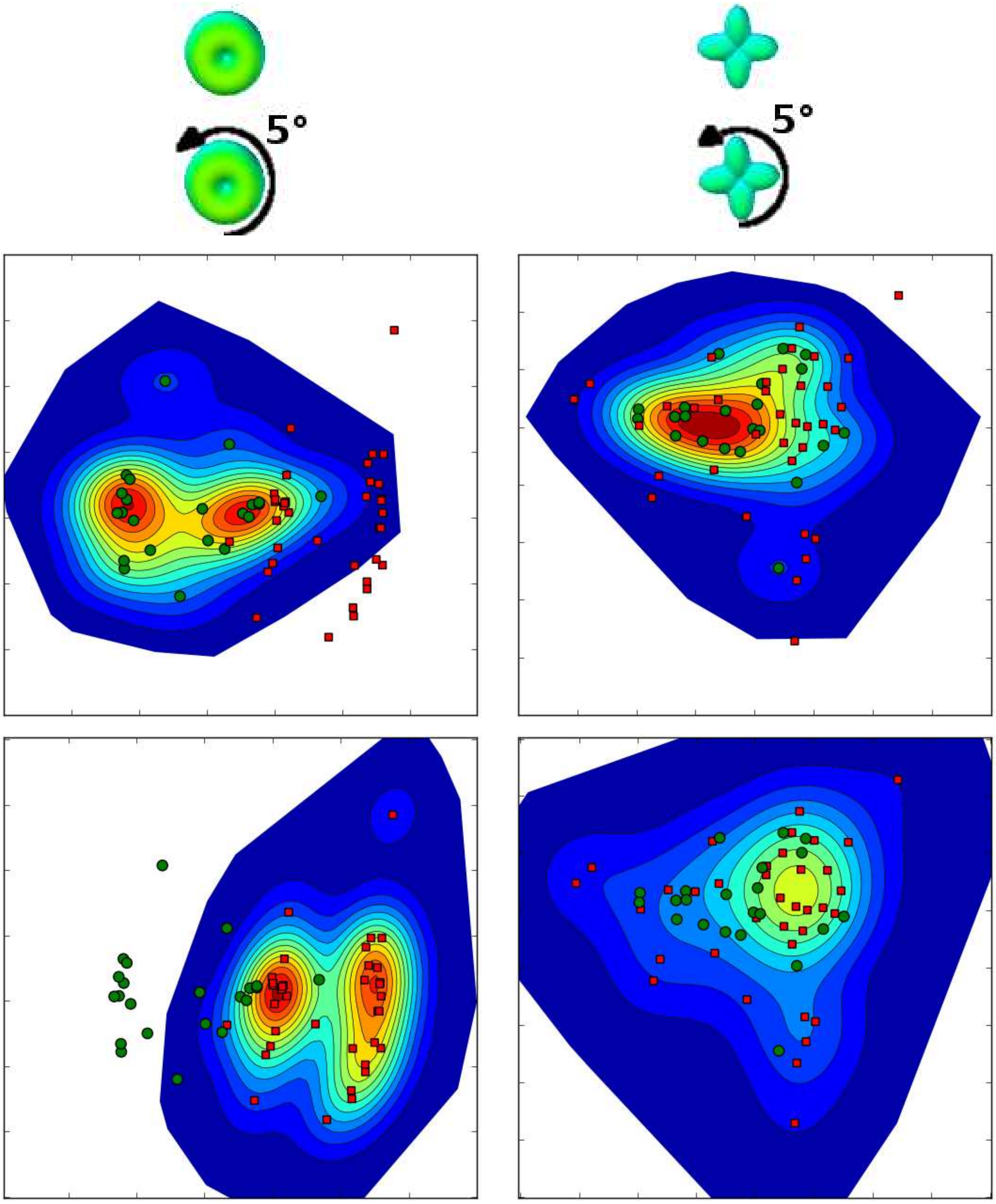
In order to examine why T4s and T2s disagreed, we checked the residuals of both T2 and T4 fODF tensor models. These residuals were initially calculated using the set of gradient directions of the acquisition, a set that generally differs from one patient to another. For that reason we chose to extrapolate the residuals to a common for all patients and more dense



T2 fODF case

T4 fODF case

FIGURE 6.9: Statistical comparisons on the T2 and T4 fODF models. The first two rows give an example of two samples from the normal population (first row) and the abnormal population (second row) for both T2 (on the left) and T4 (on the right) fODF models. Abnormal tensors are rotated by 5 degrees in comparison to the normal tensors. The 22 green points correspond to the normal population and the 36 red squares correspond to the abnormal population. All individuals are constructed by adding noise to their corresponding template. The third row depicts the PDF distribution of the normal points, while the fourth row shows the PDF of the abnormal points. We can notice that normal and abnormal distributions are dissimilar in the T4 fODF case and quite similar in T2 fODF case. Populations' PDFs use the same color scale in each case (column). Red color: high density, blue color: low density.



T2's RESIDUAL case

T4's RESIDUAL case

FIGURE 6.10: Statistical comparisons on T2 and T4 residuals of the fODF models using the L_1 norm integrated on the sphere. The first two rows give an example of two samples from the normal population (first row) and the abnormal population (second row) for both T2 (on the left) and T4 (on the right) fODF models. Abnormal tensors are rotated by 5 degrees in comparison to the normal tensors. The 22 green points correspond to the normal population and the 36 red squares correspond to the abnormal population. All individuals are constructed by adding noise to their corresponding template. The third row depicts the PDF distribution of the normal points, while the fourth row shows the PDF of the abnormal points. As we can see, normal and abnormal distributions are more similar in the T4 case than in T2, meaning that the residual of the T2 model is different between normal and abnormal individuals. In other words, the T2 model cannot capture all the information Populations' PDFs use the same color scale per each case (column). Red color: high density, blue color: low density.

sampling set of directions on the sphere. In this way, we can compare all possible coupled residual profiles by integrating their L_1 differences on the sphere, resulting into obtaining an inter-point distance matrix useful to estimate our new reduced space related to the model's residuals. Application of the proposed statistical approach on that space signified that T2 residuals can influence the final characterization of the test, since populations comparisons on T2 residuals resulted to dissimilarity. On the other hand, populations comparison on T4 residuals concluded to similarity, meaning that all individual T4s had the same kind of residual in both populations. In other words, T2 residuals contain patterns of information capable to affect the statistical analysis, whereas T4 residuals do not contain information.

Furthermore, figure 6.10 represents the PDF of each population (normal and abnormal) for both T2 and T4 cases (built on the residuals), while the right part of table 6.5 contains the estimated HPD intervals of the p-values.

In the next two sections the proposed statistical method is firstly compared to the Hotelling T^2 test on synthetic data, and secondly to one method based on permutations for high dimensional real data (*i.e.* permutations in the inter-point distance matrix).

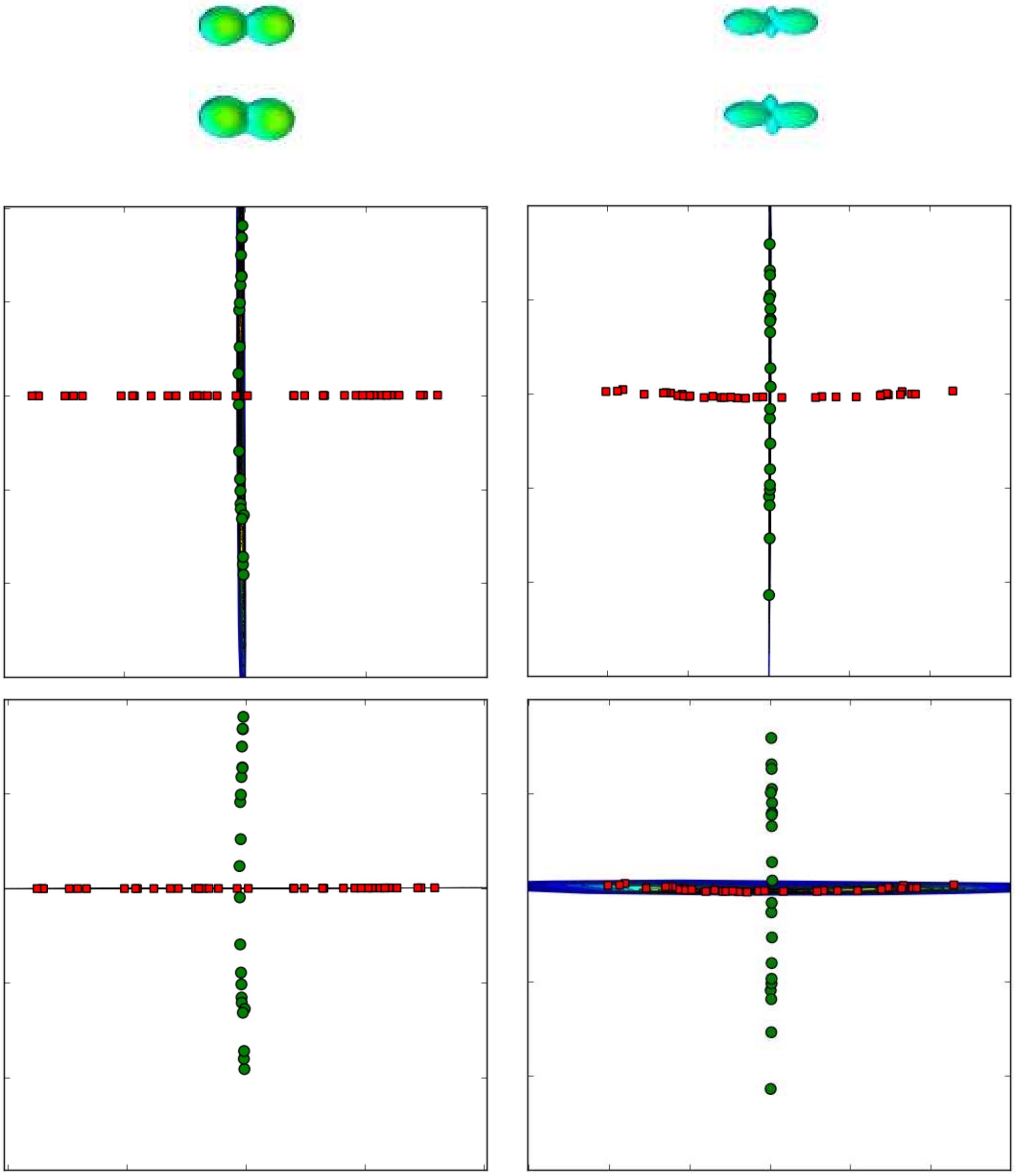
6.3.2 PDF analysis VS population's mean analysis in the reduced space - Evaluation on synthetic data

The purpose of the following comparison is to highlight the superiority of statistical methods based on comparing the whole distribution of the groups, instead of comparing the groups' mean, as achieved by the Hotelling T^2 test.

A normal and an abnormal tensor template are constructed, describing two orthogonally crossing fibers with different diffusion for each fiber. The fiber with the lowest diffusion (*i.e.* vertical) of the abnormal template is rotated by 5 degrees and has a scale different than the corresponding fiber in the normal tensor (the data were synthesized with the method of [10] using the provided Matlab code). We defined 22 normal and 36 abnormal tensors by adding uniform noise to each corresponding tensor template. In this case, the reduced space is determined by comparing the fODFs of the tensors, similarly to the common case.

Figure 6.11 shows the PDFs of each normal and abnormal population in both T2 and T4 cases. In addition, table 6.6 contains the HPD intervals of the estimated p-values for the proposed statistical test (both T2 and T4), along with the resulting p-values of the Hotelling T^2 test. We can notice that, in both T2 and T4 cases, the proposed statistical test characterized the populations as dissimilar, while on the contrary Hotelling T^2 test failed in both cases, concluding to similarity.

Furthermore, table 6.7 contains the obtained p-values for the previous synthetic case of figure 6.9. As previously shown, T2 models failed to recognize the populations as dissimilar using the proposed statistical test. In addition, the Hotelling test failed in the T2 case too. On the other hand, T4 models allowed both statistical tests to find the correct answer.



T2 case

T4 case

Figure 6.11: The first two rows contain the representation of T2 and T4 fODF models in the normal group (first row) and in the abnormal group (second row). The abnormal individuals differ from the normal ones, by changing the angle and the scale of the small vertical fiber. The 22 green points correspond to the normal population and the 36 red squares correspond to the abnormal population. The first row gives an example of two samples from the normal population and the abnormal population for both T2 (on the left) and T4 (on the right) fODF models. Normal and abnormal individuals are constructed by adding uniform noise to each group template (presented in the first two rows). The third row depicts the PDF corresponding to the normal group, while the fourth row shows the PDF of the abnormal group. It is easy to conclude that both T2 and T4 fODFs can discriminate the difference between the two populations, but as mentioned in table 6.6, the Hotelling T^2 test fails to detect the dissimilarity.

case	Proposed p-value	Proposed Decision	Hotelling's p-value	Hotelling's Decision
T2	$[0, 0.0046]$	**Dissimilar**	0.90	**Similar**
T4	$[0, 0.0046]$	**Dissimilar**	0.83	**Similar**

TABLE 6.6: Calculated p-values (HPD intervals for the proposed method) of the case presented in fig. 6.11, by comparing the proposed statistical method against the Hotelling T^2 test on T2 and T4 fODF profiles.

case	Proposed p-value	Proposed Decision	Hotelling's p-value	Hotelling's Decision
T2 fODF	$[0.73, 0.80]$	**Similar**	0.76	**Similar**
T4 fODF	$[0.0058, 0.025]$	Dissimilar	0.0016	Dissimilar

TABLE 6.7: HPD intervals of p-values on fODF profiles using the proposed statistical test (left part) and the Hotelling test (right part) for the case presented in fig. 6.9, along with the decision for each case. T4 fODF concluded to dissimilar, while T2 fODFs to similar populations (for both tests). The disagreement occurred due to structured T2 residuals.

6.3.3 PDF analysis in the reduced space VS inter-point distance matrix analysis in high dimensional space - Evaluation on real NMO data

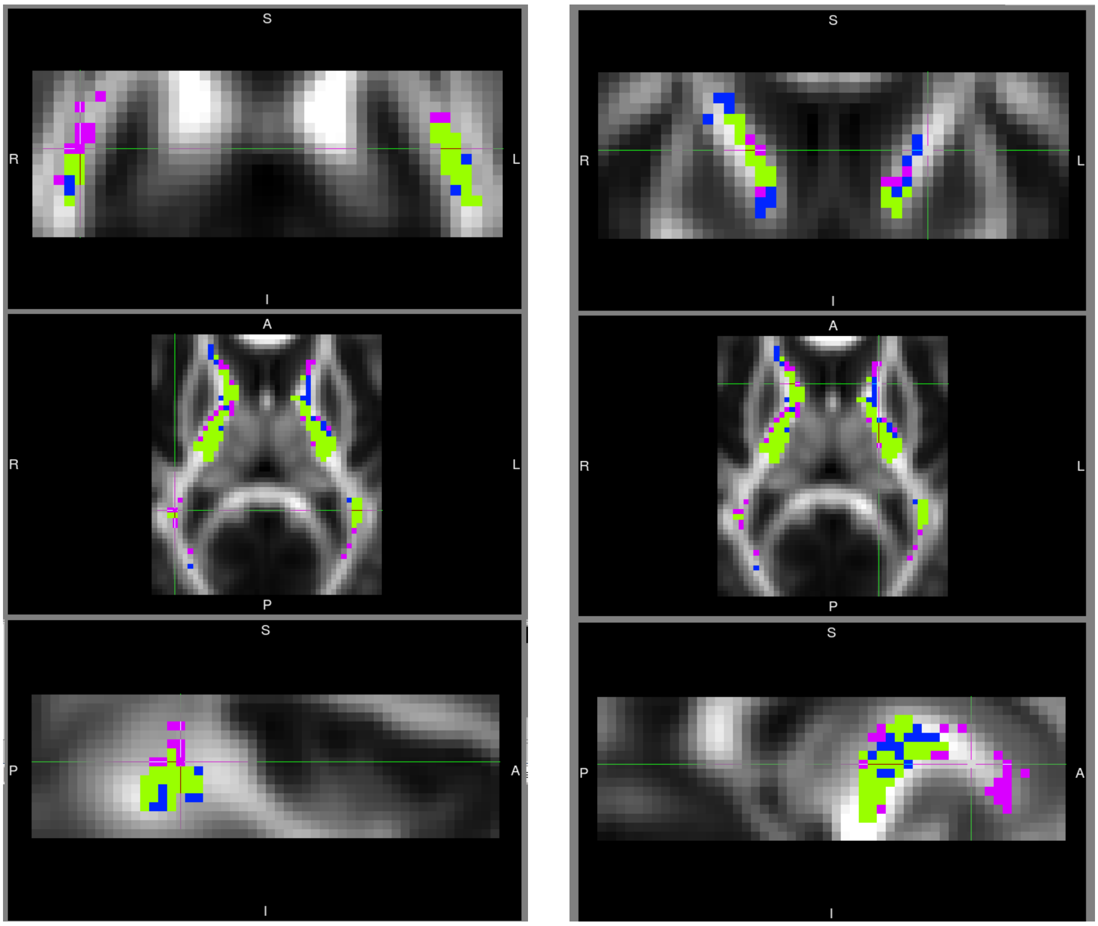
As presented in chapter 4, section 4.3.2, it is possible to define statistics in high dimensional space assisted by an inter-point distance matrix [23]. Moreover, setting a permutation test on the row/column elements of the inter-point distance matrix and measuring at each iteration the statistic proposed in [23] and mentioned in eq. 4.1, allows us to estimate a p-value and its credibility interval.

Similarly to the T2-T4 comparisons on real data, we would like to check the consistency of the proposed statistical test with this particular permutation test on the inter-point distance matrix based on T4 fODF models. Once again we would like to compare the top $N = 1272$ sorted list of voxels for both approaches. Figure 6.12 depicts voxels in the same two ROIs colored by the green-purple-blue coded scheme, where green corresponds to the voxels appearing both in top N voxels, purple produced only by the T4 fODF in the reduced space by the proposed statistical test and blue appear only in the top N voxels resulted by the permutation tests on the distance matrix.

# of Tested Voxels	Green (Both cases)	Purple (only T4 fODF)	Blue (only T4 Matrix shuffles)
1531	1013	259	259

TABLE 6.8: Count of voxels in green, purple and blue color of the T4 fODF's *versus* T4 matrix permutations' statistics.

Table 6.8 contains the exact numbers of voxels in each color case. The percentage of differences in ranking is 20.36% (= 259/1272) in the top $N = 1272$ most significant voxels.



ROI 1 ROI 2

FIGURE 6.12: Comparison of the ranking of T4 fODF statistics with statistics based on permutations on the inter-point distance matrix of the T4 models in two different ROIs. Green voxels corresponds to the case of getting both ranks in the top $N = 1272$ dissimilar voxels, purple voxels appeared only in the T4 fODF's top N dissimilar voxels and blue voxels only in the T4 matrix permutations' top N dissimilar voxels. Greyscale images correspond to FA template. Top images correspond to coronal views, middle images correspond to axial views and bottom images correspond to sagittal views.

Figure 6.13 illustrates the histograms of the resulting p-values (*i.e.* upper bound of the HPD interval of each p-value), for the case of permutation testing in the inter-point distance matrix, in the whole pathological ROI of the 2741 voxels, for T4 fODF, T2 fODF and T2 coefficients cases. A first comparison between the subfigures of figure 6.13 and figures 6.1, 6.3 and 6.4 reveals the higher sensitivity of the proposed statistical approach against permutation testing in the inter-point distance matrix, in a specific pathological area, for all the three cases (T4 fODFs and T2 fODFs using the proposed tensor metric and T2 coefficients using the log-Euclidean distance). Secondly, much more biomarkers were extracted by the proposed statistical method, than the inter-point distance statistical analysis.

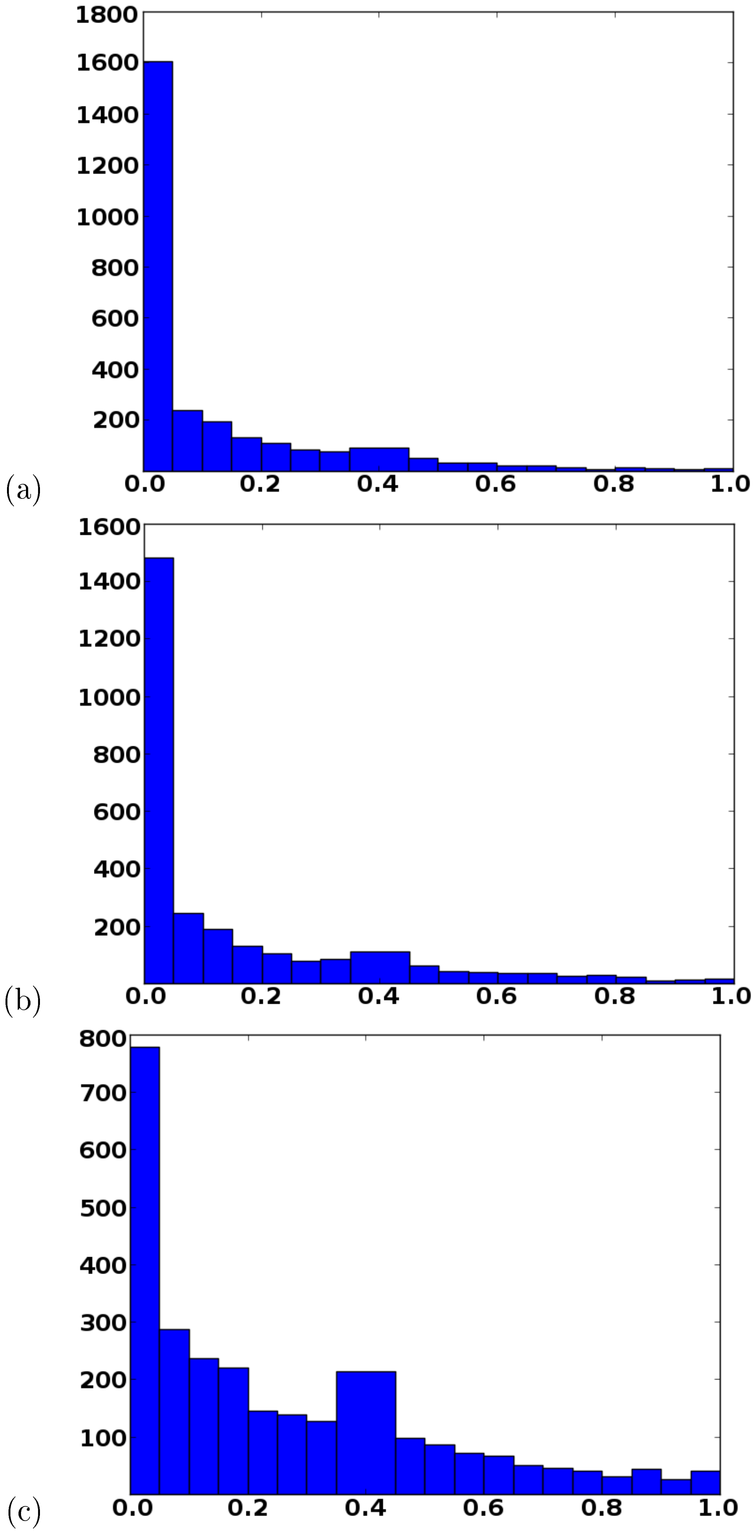


FIGURE 6.13: Histograms of the upper bounds of the p-values' HPD intervals using permutation testing in the inter-point distance matrix for the same pathological ROI with 2741 voxels (bin size = 0.05). (a) T4 fODF case, (b) T2 fODF case, both using the proposed distance (eq. 2.30) and (c) T2 coefficients using log-Euclidean distance (eq. 2.28).

6.3.4 PDF analysis in the reduced space VS RF classification analysis in different feature spaces - Evaluation on real NMO data

It is known that RFs are capable to handle efficiently high dimensional data. In our case, RF classifiers are used. Therefore, we started working with different kind of high dimensional real data and gradually reduced the dimension, by using dimensionality reduction techniques, such as Isomap [154], in order to keep the coherency with the previously mentioned methods.

Given the same set of 58 samples (22 normal and 36 related to NMO datasets), we tested different kind of RF parameterizations (especially for the number of trees and the maximum tree depths) in each feature space. We will present the best parameterization for each case.

Introducing neighboring information in the fODF space

Working in the fODF space of the T4s, we need to discretize the fODF function by sampling it on the unit hemisphere, in order to define a feature vector per patient. Moreover, adding neighboring information at each voxel, for example $5 \times 5 \times 5$ fODF patches, we come up with a set of 58 samples in the dimension of $5 \times 5 \times 5 \times N$, where in our case $N = 242$ fODF samples. Finally the sample's features dimension is 30250.

The RF with 2500 decision trees and maximum tree depth $D = 4$ was the best parameterization for this case. HPD intervals of the GE and the p-value of the proposed method were calculated in a given ROI with 2742 voxels. Figure 6.14 compares the middle values of the HPD intervals between GE and the resulting p-value of the proposed method (each point in the figure corresponds to a voxel in the ROI).

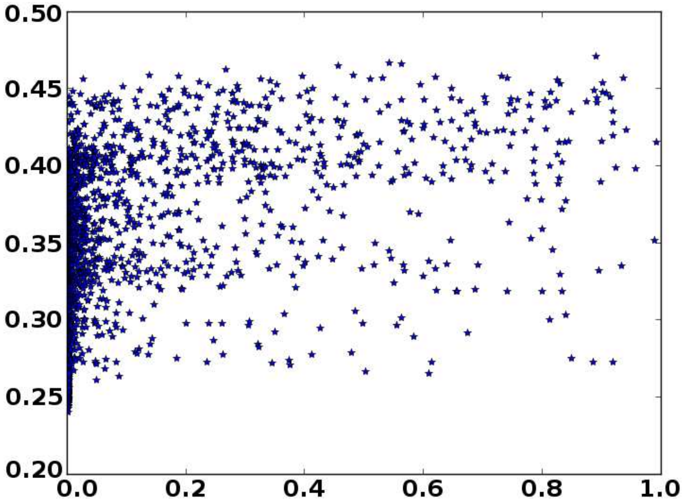


Figure 6.14: Comparing the middle values of the HPD intervals between the p-value of the proposed statistical method (horizontal axis) and the GE of the RF classifier in the $5 \times 5 \times 5 \times 242$ fODF space (vertical axis) in a given ROI of 2742 voxels in the brain. Working in this space did not fulfil our expectations for a dense and increasing form.

Since we would expect to observe a fit forming an increasing function (so that both methods will agree to the same decision *e.g.* similar=high value) and did not happen, we thought to change the feature dimension and instead of working with 242 fODF values in each voxel, to work with the T4 space ($15D$) in order to obtain possibly the best RF results.

To continue, we should note that RF results are assumed incorrect. Strictly speaking RF results should be verified using medical expertise. Since it is not done, we assume that the proposed method provides us with robust results, because it benefits from certain attributes, such as the geodesic distances between fODF profiles, the reduced working space in comparison to high dimensional space with probably no structured and sparse populations as in the RF case.

Introducing neighboring information in the T4 space

Neighboring information in the T4 space of the 15 tensor coefficients, yields a working space of dimension $5 \times 5 \times 5 \times 15 = 1875$. An optimized RF classifier with $T = 2500$ and $D = 4$ provided us the results of figure 6.15. Unfortunately, it did not improve the results significantly, since many deviations still remain.

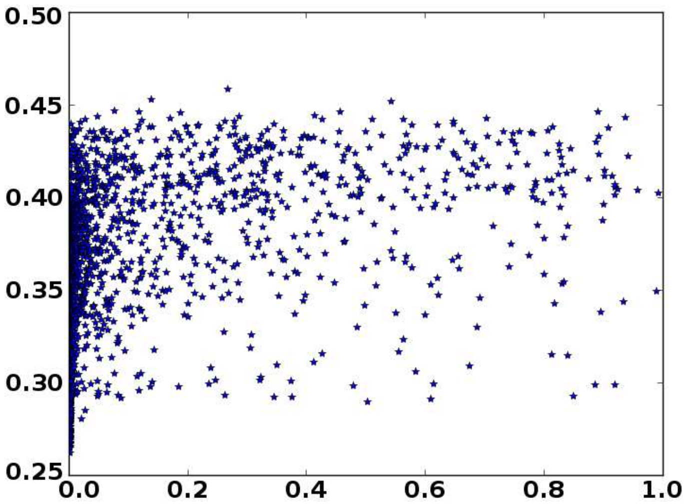


FIGURE 6.15: Comparing the middle values of the HPD intervals between the p-value of the proposed statistical method (horizontal axis) and the GE of the RF classifier in the $5 \times 5 \times 5 \times 15$ T4 space (vertical axis) in a given ROI of 2742 voxels in the brain.

At this point we should remember that the $15D$ space of T4s can be sparsely filled, due to lack of data, or even due to the fact that T4s lie on a submanifold in $15D$. Taking into consideration geodesic distances between T4s could allow us to achieve more accurate solutions. Therefore, we thought to assist the RF classifier by identifying that T4 submanifold, not necessarily $2D$ as before, but we can start working from a higher dimensional space, for example $5D$ (or even higher).

5D reduced space using Isomap

The very high dimension of our two previous cases, along with the medium quality of the obtained results, led us to the reduction of the dimension of the data, similarly to the proposed statistical method, in order to work in more densely filled spaces. Initially, we chose to transform our data to the $5D$ space, where an optimized RF classifier with $T = 500$ and $D = 4$ gave results much more consistent with the ones given by our method. Figure 6.16 presents the corresponding comparison for this set of experiments.

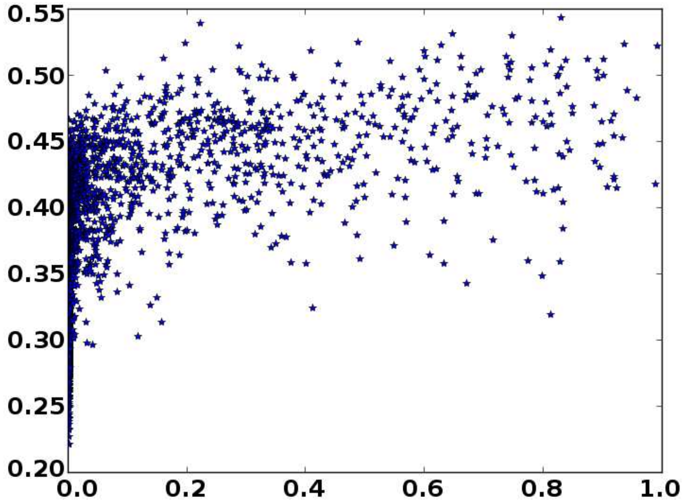


FIGURE 6.16: Comparing the middle values of the HPD intervals between the p-value of the proposed statistical method (horizontal axis) and the GE of the RF classifier in the $5D$ reduced space (vertical axis) in a given ROI of 2742 voxels in the brain.

Thereby, we decided to reduce the dimension more and we decided to work with $2D$ data, the same dimension as the proposed statistical model.

2D reduced space using Isomap

Fitting a RF with $T = 500$ and $D = 4$ in $2D$ data gave the best coherency between the four tested cases, as shown in figure 6.17.

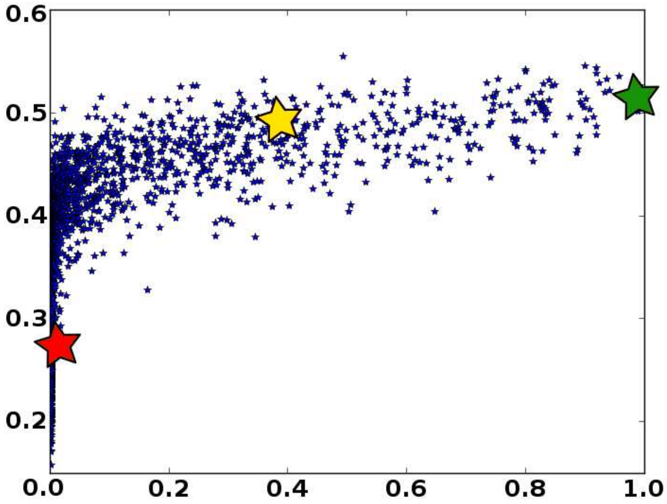


FIGURE 6.17: Comparing the middle values of the HPD intervals between the p-value of the proposed statistical method (horizontal axis) and the GE of the RF classifier in the $2D$ reduced space (vertical axis) in a given ROI of 2742 voxels in the brain.

Three characteristic cases (very similar, similar and dissimilar voxels, represented by green, yellow and red stars in fig. 6.17) are isolated and further studied in figure 6.19, where we can see the green points representing the normal population and the red points displaying the abnormal one. The background color illustrates the RF classification according to the given sets of normal and abnormal individuals. Areas with higher probability to belong to the normal group than the abnormal one are depicted with green, in contrast to the red-coded abnormal group. Moreover, areas with high uncertainty are coded with brown-level colors.

Sorting the p-values (and the generalization errors) of all voxels in a ROI reveals the most important biomarkers of that ROI. Connection between p-value's ranking and RF generalization error's ranking concludes that RF are coherent with the proposed statistical approach (see figure 6.18, data are spread around the $y = x$ line).

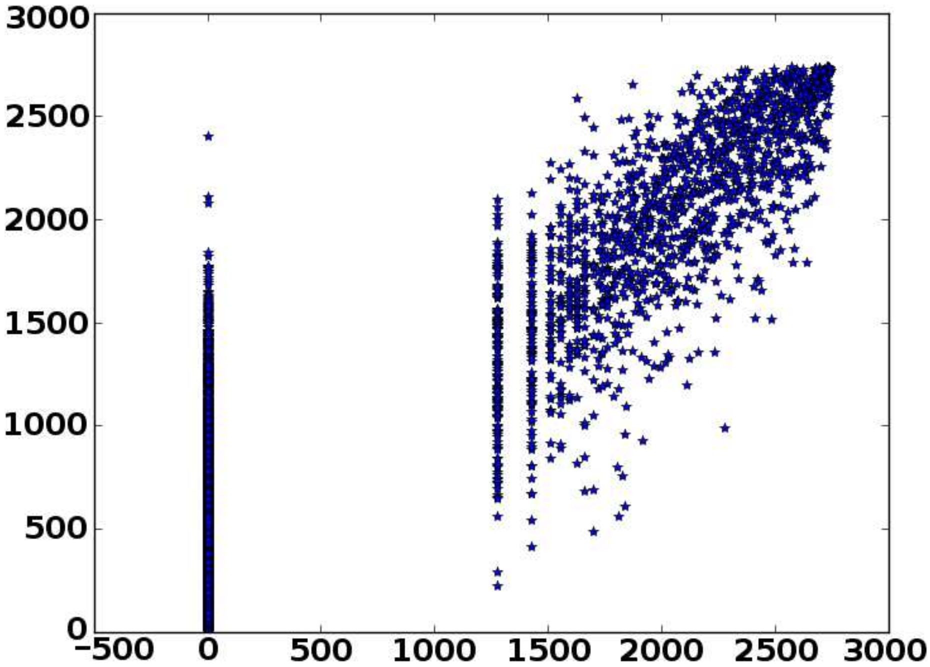


Figure 6.18: Comparing the ranking of the middle values of the HPD intervals between the p-value of the proposed statistical method (horizontal axis) and the GE of the RF classifier in the $2D$ reduced space (vertical axis) in a given ROI of 2742 voxels in the brain. The points follow in general the $y = x$ line.

Our initial expectation was that RF models could perform better in the initial high dimensional space. Probably, given the complexity of the structure of T4/fODF spaces, in practice, it was shown that reducing the dimension is important. RF in 2D reduced space, gave results coherent with the proposed method, although the crucial task was performed by Isomap and the calculation of the reduced space.

Finally, we can say that the proposed method produced more biomarkers than the RF models.

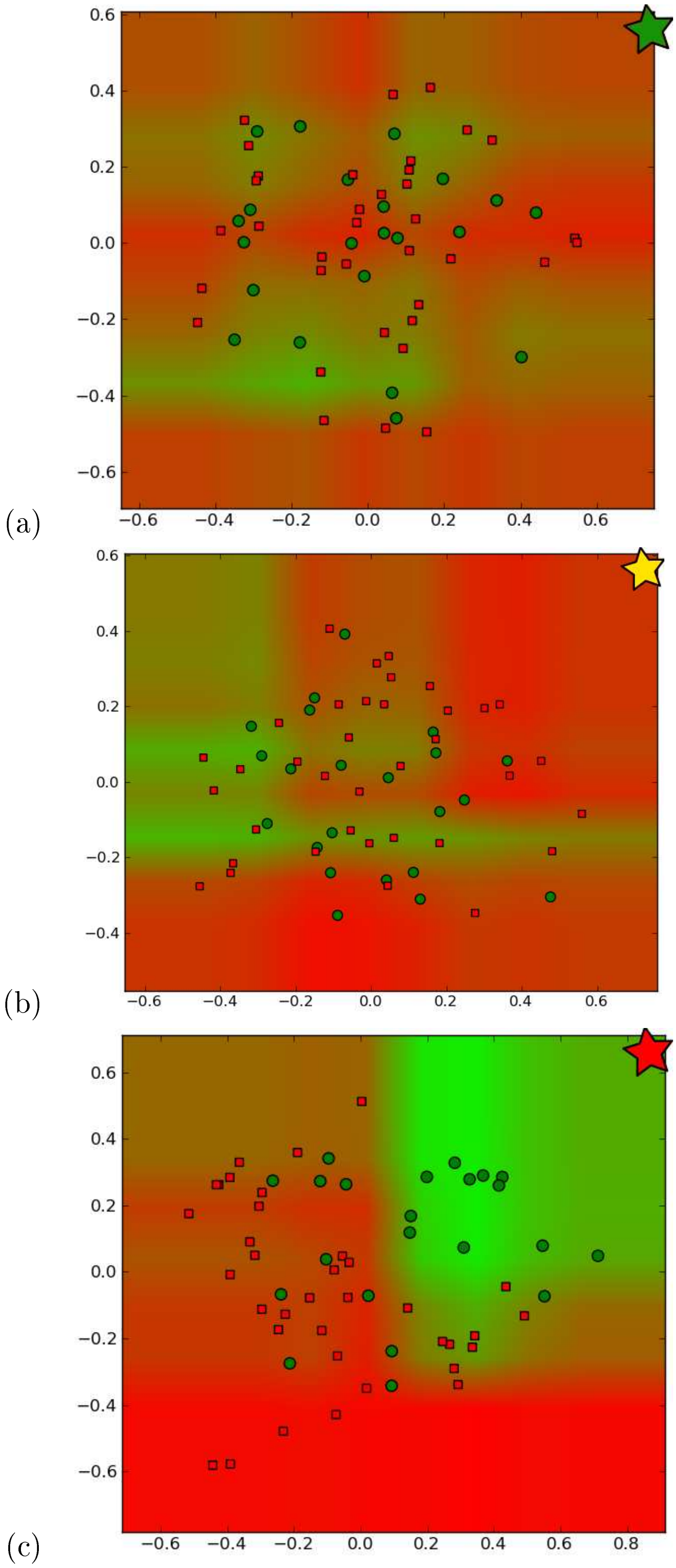


FIGURE 6.19: Visualization of the RF ($T = 500, D = 4$) classifications for three characteristic cases (extracted from fig. 6.17) and comparison with the resulting p-values (with 1000 label shufflings) of the proposed statistical method. (a) Similarity: GE = $[0.50, 0.53]$, p-value = $[0.988, 0.999]$, (b) Similarity: GE = $[0.48, 0.50]$, p-value = $[0.35, 0.43]$ and (c) Dissimilarity: GE = $[0.28, 0.30]$, p-value = $[0.0, 0.0046]$.

6.4 Partial Conclusion

Experimental results of the proposed method were presented in this chapter. Several evaluative schemes, on synthetic and on real NMO data, showed the coherence of the proposed method with medical knowledge. Moreover, the superiority of the T4 tensor model against the T2 model was shown. Results obtained by the methods described in section 4.3 are consistent with those derived from the proposed method, or are even worse.

As part of future work, it will be interesting to apply the proposed statistical model in ROIs with no direct relation with pathological areas, with ultimate view to discovery new biomarkers, since inflammatory diseases, such as NMO, can potentially spread all over the brain.

The next chapter contains an alternative proposed statistical analysis, suitable for individual *versus* normal population comparisons. Application to LIS disease will be presented.

Chapter 7

Individual *VS* Normal Population: Method and Application to LIS disease

In cases where the variability of the abnormal population cannot be totally captured due to the lack of enough pathological data, it is not pertinent and even not safe to rely on population *vs* population approaches, including the proposed method in chapter 5. The existence of empty (unlabeled) areas in the space, due to the lack of (abnormal) points, close to the mass of the normal population will result in data being probably misclassified as normal, in the absence of knowing completely the variability of the abnormal group. Moreover, under certain circumstances, it is much more robust to evaluate the state of every patient separately, for example, in patient follow-up. As a consequence, each incoming abnormal dataset should be tested individually *versus* the normal population (in most of the cases is well-defined by a large dataset), which will permit us to follow the state of the patient across several in time scans.

In this chapter, a variation of the method proposed in chapter 5 is described, aiming to calculate voxelwise statistics in the case of sparse populations. Experimental results for Locked-in syndrome (LIS) are achieved by collecting and post-processing the voxelwise statistics in specific regions of interest related to the motor system, which are characteristically affected by LIS. Both fODF and diffusion (ADC) profiles produced by different T4 models were examined. Finally, comparisons between the proposed scheme and classical approaches are included.

7.1 Proposed Statistical Model

Although many elementary steps such as the DW-MRI normalization, the use of fourth order tensors, the computation of the inter-point distance matrix, based on the same tensor metric (eq. 2.30) including neighboring information (*i.e.* $3 \times 3 \times 3$ best patches) and the idea of determining the reduced space with Isomap retain attached to the core of the new method of this chapter, a few parts differ from our previous scheme.

To begin with, instead of fitting and transforming all individual datasets (normal and abnormal) in the reduced space at once, meaning that the distances referring to the abnormal individuals will affect the position of the normal points (as is the case for the population *vs* population problem), at this time, the reduced space for the normal points is constructed by taking into consideration the part of the inter-point distance matrix only between the normal datasets. In other words, the reference model of the reduced space will not depend on any abnormal point. This is done to avoid estimating a shrunken normal group due to the presence of large distances related to the abnormal points.

Once the normal population is determined in the reduced space, each abnormal individual will be treated as an independent incoming datum, that will be transformed to a new point in the reduced space by fitting it to the reference model, corresponding to the normal population, by introducing its relative part of the inter-point distance matrix (*i.e.* distance vector referring to abnormal i *vs* all normal individuals). It is important to note that transforming the abnormal individuals in the reduced space can be implemented in parallel, without altering either the position of the normal points, or the previously tested abnormal points. In addition, following the analysis of Isomap's reconstruction error with respect to several values of the reduced dimension that was presented in section 5.2, we will keep working in $2D$.

7.1.1 Statistic of Interest and Determination of HPD Interval per p-value

At this point, we have the normal points in the reduced space, for a given voxel. Similarly to the previously proposed method, we choose to fit a Gaussian kernel at each normal point, thus representing the normal population as a Gaussian Mixture Model (GMM) (eq. 7.1) with the aid of kernel density estimation (KDE) [76]:

$$p\left(\boldsymbol{x}\right) = \frac{1}{I}\sum_{i=1}^{I} \mathcal{N}\left(\boldsymbol{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}\right). \tag{7.1}$$

The covariance $\boldsymbol{\Sigma}$ is identical for all kernels and is determined according to Scott's rule [142].

For each incoming transformed abnormal point $\boldsymbol{y}$, we consider its PDF value $p\left(\boldsymbol{y}\right)$ given the distribution of the normal population, as our statistic of interest (in contrast to the discrepancy, that was measured between two PDFs and used in the previous method). In this way, we can estimate the p-value $\nu^{\star}$, referring to the probability of getting a PDF equal or lower than $p\left(\boldsymbol{y}\right)$ under the Null Hypothesis that the abnormal point $\boldsymbol{y}$ belongs to the normal population.

Theoretically, $\nu^{\star} = \int_X p\left(\boldsymbol{x}\right) d\boldsymbol{x}$, $X = \left\{\boldsymbol{x} \mid p\left(\boldsymbol{x}\right) \leqslant p\left(\boldsymbol{y}\right)\right\}$, but in practice this integral cannot be computed analytically. For this reason, we chose to estimate the p-value $\nu^{\star}$ *via* Monte Carlo simulations and the generation of K random samples from $p\left(\boldsymbol{x}\right)$ (*e.g.* $K = 5000$, $\{x_1, \ldots, x_K\}$ samples, $x_k \sim p\left(\boldsymbol{x}\right)$).

HPD Interval Estimation for each p-value

Due to the fact that we are not satisfied with a pointwise estimator $\widehat{\nu}$ of the p-value $\nu^{\star}$ at each voxel (eq. 7.3), we wish to determine the solution's precision by extracting a HPD interval for each p-value, in a similar manner as was done in the populations comparison problem (*e.g.* section 5.4).

Each comparison $p(\boldsymbol{x}) \leqslant p(\boldsymbol{y})$ will result in a binary value $q(\boldsymbol{x})$ equal to 1 when the condition is true and 0 otherwise. We will denote $q_k = q(x_k)$.

$$P(q_k) = \begin{cases} \nu & \text{, if } q_k = 1, \\ 1-\nu & \text{, if } q_k = 0. \end{cases} \tag{7.2}$$

$$\nu^{\star} = \int_X p(\boldsymbol{x})\, d\boldsymbol{x} = \int q(\boldsymbol{x}) p(\boldsymbol{x})\, d\boldsymbol{x} \simeq \frac{1}{K}\sum_{k=1}^{K} q_k = \widehat{\nu}. \tag{7.3}$$

Assuming a uniform prior for $\nu^{\star}$, it is possible to calculate the posterior $p(\nu \mid q_1, \ldots, q_K)$ which follows a $Beta(\alpha+1, \beta+1)$ distribution, with α equals to the number of 1's and β the number of 0's in $\{q_k\}, 1 \leqslant k \leqslant K$ sequence:

$$\left.\begin{aligned} &P(\nu \mid q_1, \ldots, q_K) = \tfrac{P(q_1,\ldots,q_K \mid \nu)\ P(\nu)}{P(q_1,\ldots,q_K)}, \\ &P(q_1, \ldots, q_K \mid \nu) \propto \nu^{\alpha}(1-\nu)^{\beta} \end{aligned}\right\} \Rightarrow P(\nu \mid q_1, \ldots, q_K) \propto \nu^{\alpha}(1-v)^{\beta}$$
$$\sim Beta(\alpha+1, \beta+1) \tag{7.4}$$

To continue, we can estimate the interval of the underlying p-value ν, as the HPD interval (see fig. 4.2) referring to the 99% of the *a posteriori* mass $p(\nu \mid q_1, \ldots, q_K)$. Once again, increasing the number of samples K can effectively reduce the length of the estimated interval.

7.2 Experimental Results

Evaluation of the proposed method on real data was achieved using a set of 22 normal DW-MRI data, describing the normal population, and 4 abnormals referring to two separate scans for each of the two LIS patients in our data repository. DW-MRI data represent HARDI data consisting of 30 non colinear gradient directions (scanned twice), with a b-value equal to 1000 s/mm^2, a resolution of $1.8 \times 1.8 \times 3.5$ mm^3 and an image size of $128 \times 128 \times 41$.

Although the proposed statistical test can be performed voxelwise across the whole volume of the brain's WM, for the purposes of this thesis we will focus on specific regions of interest (ROIs). These ROIs compose the motor system. The ROIs are the pontine crossing tract, the left and right corticospinal tracts, the left and right medial lemniscus, found in the bottom part of the brain (in a lateral view, close to the spinal cord) and finally the left and right posterior limb of the internal capsule and the left and right superior corona radiata, located in the middle and upper parts of the brain. The definitions of these ROIs are available in the JHU-ICBM-labels template of FSL [89]. Patients with LIS are conscious, but unable to move or to communicate (*i.e.* quadriplegia) except using eye movements in some cases. LIS syndrome produces anatomical lesions at the ventral part of the pons, which induce interruptions of WM tracts, especially the corticospinal tract. As a consequence, the motor system is generally considered as a keypoint system that contains lesions due to LIS.

In the next sections, we will discuss the performance of the proposed statistical method for the individual *versus* normal population problem, by using fODF and diffusion (ADC) T4 models to represent our DW-MRI data. In this study, we particularly focus on T4 models, since T4s achieve higher accuracy in describing the diffusion properties and fiber structure than the corresponding T2s (see fig. 7.1 which highlights the better representation of the diffusion T4s/DT4s against diffusion T2s/DT2s in specific ROIs of the motor system).

In addition, ways to improve the performance will be discussed *via* variations of the proposed tensor metric. Finally, comparisons of the proposed method against classical approaches based on standard statistics on FA and MD scalar images will be presented.

7.2.1 Results based on fODF T4s and on DT4s

To begin with, our goal is to measure the percentage of lesions (*i.e.* the amount of voxels related to p-value $\leqslant 0.05$) in each ROI. Initially working, by default, with fODF T4 models, we observed that the percentages of lesions per ROI (*i.e.* table 7.1 or figure 7.2 for both patients) were not as high as we expected, knowing that these ROIs are associated with LIS.

A possible explanation could be that since fODF profiles are scaled functions paying more attention to the diffusion's orientation properties than the diffusivity values, the statistical test did not detect many lesions depending on the orientation of the DW-MRI data (*i.e.* geometry of the fiber structure), meaning that fibers' orientation would possibly maintain its normality in high levels.

In addition, the lower than expected percentages of lesions based on fODF T4 models prompted us for a more detailed study, which was achieved by estimating DT4s [11]. Observing the results included between parentheses in table 7.1, the obtained percentages of lesions using DT4s are significantly higher than the corresponding fODF cases. In other words, it is impressive to note that LIS datasets contain much more lesions affecting the diffusivity's properties, such as the magnitude of diffusion, for example related to the number of fibers passing through each voxel, rather than lesions affecting the geometrical properties (*e.g.* orientation of diffusion).

Furthermore, observation of table 7.1 (or figure 7.2) permits us to give several useful directions of thinking to the physicians, concerning the patient follow up procedure. It is remarkable that the percentages of lesions in the top five ROIs presented in the table, which are located close to the spinal cord, are higher than the percentages of lesions appearing in the middle and upper parts of the brain, locations where tracts started from the spinal

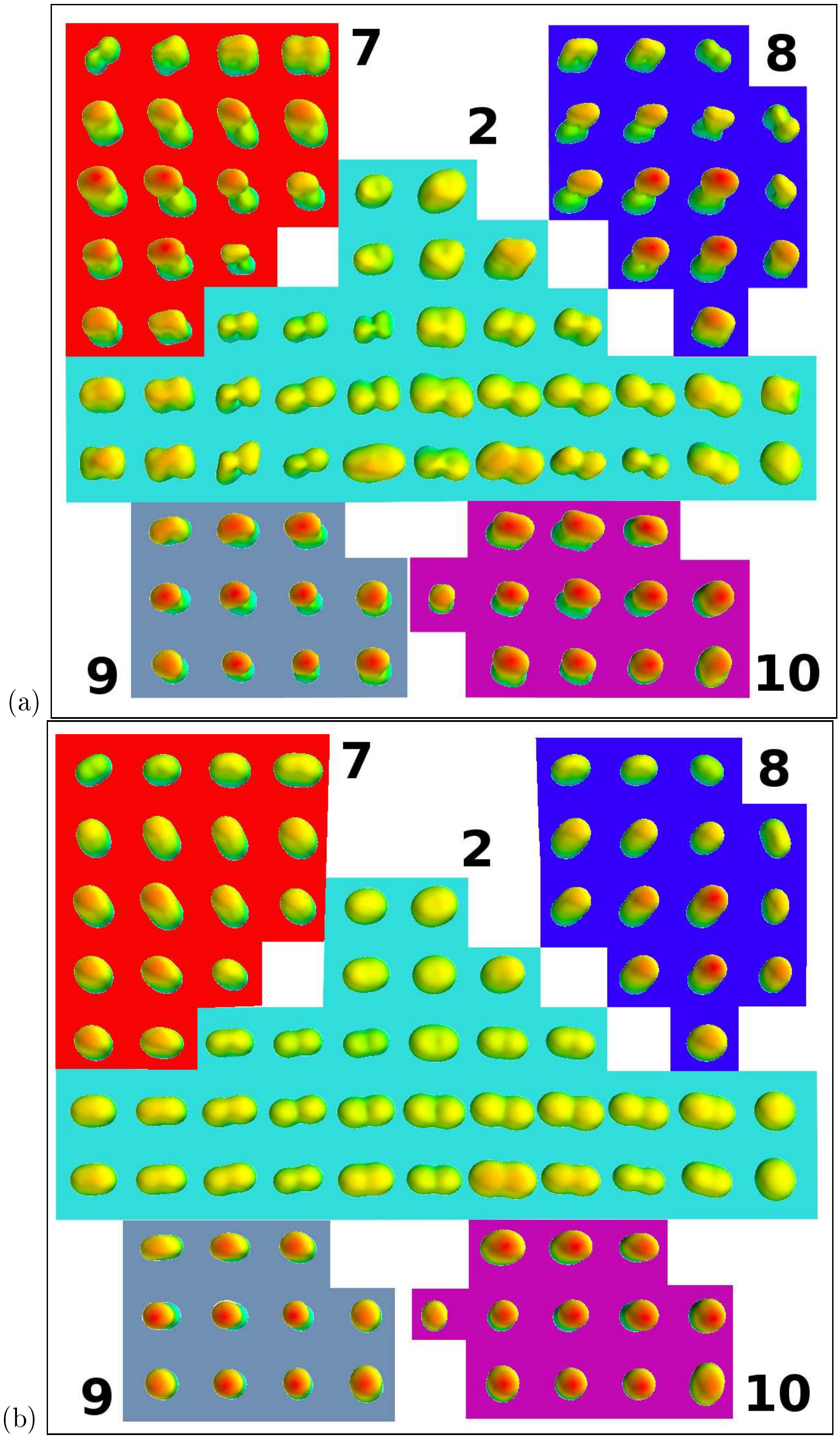


Figure 7.1: Visualization of the embedded (a) DT4 and (b) DT2 models in five patches of specific ROIs of the motor system. ROIs' labels correspond to JHU-ICBM-labels-2mm template of FSL [89]. 2: pontine crossing tract, 7: right and 8: left corticospinal tract, 9: right and 10: left medial lemniscus. It is noticeable that DT4s are more accurate models than DT2s.

Name of ROI	**PATIENT 1 Scan 1 fODF T4 (DT4)**	**PATIENT 1 Scan 2 fODF T4 (DT4)**
Pontine crossing tract	**7.1% (23.5%)**	**21.86% (62.84%)**
Corticospinal tract R	**0.57% (10.23%)**	**6.25% (28.41%)**
Corticospinal tract L	**3.37% (14.04%)**	**11.24% (30.9%)**
Medial lemniscus R	0.0% (0.0%)	**6.98% (17.44%)**
Medial lemniscus L	**6.02% (16.87%)**	**4.82% (20.48%)**
Post. limb of internal capsule R	2.99% (6.59%)	**3.59% (18.76%)**
Post. limb of internal capsule L	5.24% (5.45%)	5.45% (13.21%)
Superior corona radiata R	**1.63% (8.48%)**	**2.72% (23.15%)**
Superior corona radiata L	7.36% (9.63%)	**11.04% (18.72%)**

Name of ROI	**PATIENT 2 Scan 1 fODF T4 (DT4)**	**PATIENT 2 Scan 2 fODF T4 (DT4)**
Pontine crossing tract	3.83% (8.74%)	**2.73% (26.78%)**
Corticospinal tract R	**6.25% (18.18%)**	**4.55% (19.89%)**
Corticospinal tract L	**12.36% (23.03%)**	**4.49% (19.1%)**
Medial lemniscus R	**5.81% (17.44%)**	**9.3% (25.58%)**
Medial lemniscus L	**7.23% (21.69%)**	**12.05% (24.1%)**
Post. limb of internal capsule R	0.4% (5.99%)	0.2% (3.59%)
Post. limb of internal capsule L	2.52% (10.27%)	**2.1% (11.53%)**
Superior corona radiata R	3.37% (3.91%)	1.96% (4.35%)
Superior corona radiata L	1.41% (5.09%)	1.41% (2.6%)

Table 7.1: LIS Patient 1 (top) and 2 (bottom) follow-up for 9 ROIs (from JHU-ICBM-labels template of FSL [89]) related to the motor system. Percentage of lesions (p-value $\leqslant 0.05$) per ROI for both scans using T4 fODF profiles and T4 diffusion profiles (between parentheses) are presented in the table. It is obvious that the percentages of lesions are higher in the diffusion than in fODF profiles.

cord are passing through or end at that level (*e.g.* four ROIs in the bottom of the table) for both patients. The lesions detected in the top five ROIs are coherent with the medical expectations, since these ROIs are the first keypoint areas to detect lesions related to LIS. Appearance of lesions in the four last ROIs (middle and upper parts of the brain) may be caused by Wallerian degeneration. Moreover, someone could say that lesions in the last set of ROIs react differently depending on the patient. For example, patient 2 exhibits less lesions in the middle-upper part of the brain than patient 1.

Another interesting point is the increasing percentages of lesions between the two scans, for both patients in most of the ROIs. Although the clinical status of the patients did not change remarkably between scans, since both patients were totally paralysed from the first time scan, the increasing percentages can be seen as the expected evolution of the corrupted tracts.

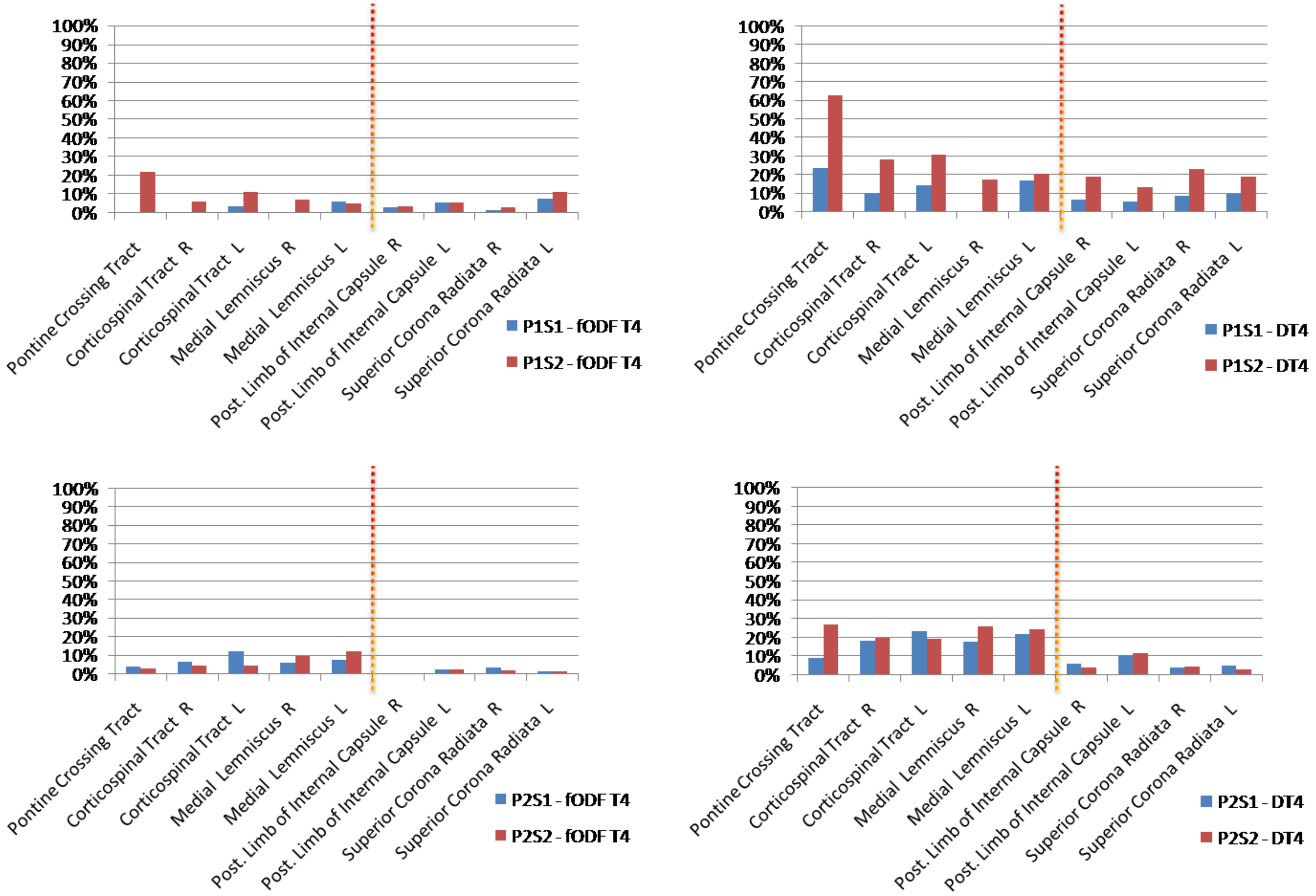


FIGURE 7.2: Plotting the percentages of lesions detected using the proposed method on fODF T4s and diffusion T4s (as presented in table 7.1). The labels are coded as "PiSj-*data*" referring to "Patient i Scan j on specific *data*". The vertical dotted lines separate the two groups of ROIs (ROIs in the bottom part of the brain on the left and ROIs in the middle and upper parts of the brain on the right).

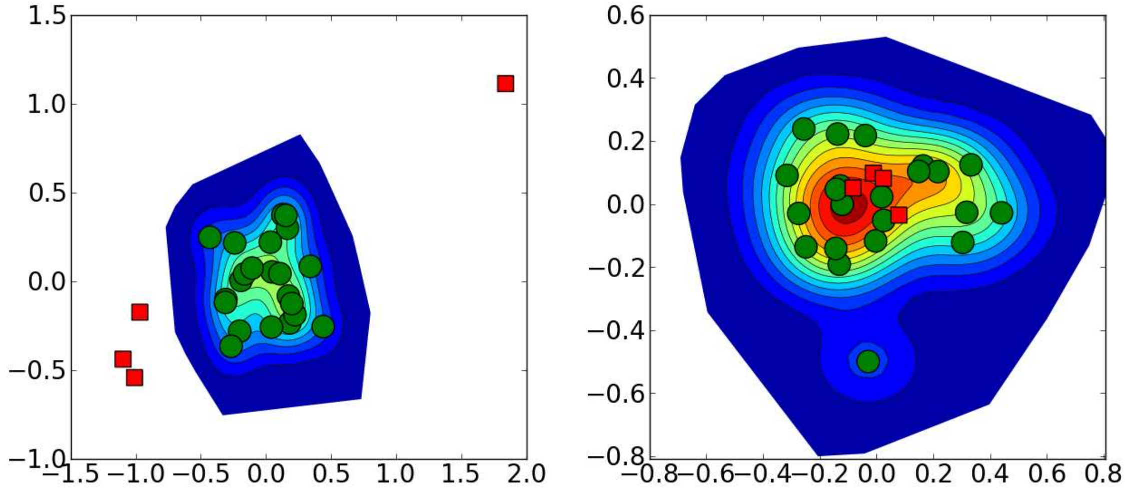


FIGURE 7.3: Two examples of reduced space configurations using DT4s. The green points describe the atlas (*i.e.* reference data), the background represents the PDF (blue: low values; yellow to red: high values). The four red squares describe incoming abnormal data. 1^{st} voxel of interest (left subfigure): all incoming data are identified as abnormal ; 2^{nd} voxel (right subfigure): all incoming data are identified as normal.

To continue, figure 7.3 depicts the reduced space of the normal points (*i.e.* green points) along with their population PDF (*i.e.* colorful background, where blue corresponds to low PDF values and red to high) and the four transformed abnormal points (*i.e.* red squares) of two interesting voxels extracted through processing DT4s. In the left figure, we can notice that all abnormal points are punished, located outside the core of the normal population, due to their highly abnormal DT4 properties. On the other hand, in the right figure, all abnormals are considered to be healthy, equivalent to the normal points, since they are located in the mass of the normal population. Of course, other configurations can be extracted too, especially when the state of the patient's health is altered due to recrudescence of the pathology. In the last case, some red squares could be located in the mass of normality (*i.e.* healthy state) and some other outside the periphery of it (*i.e.* abnormal state).

Due to the fact that DT4s are more sensitive, managing to distinguish higher percentages of lesions than the fODF T4 models within areas related to the LIS disease, it will be better to build our statistical analysis model on DT4 data for this particular disease.

In the next sections, comparisons of the proposed statistical approach adapted to DT4 and DT2 models are presented. Furthermore, classical statistical analysis of FA and MD images will be described.

7.2.2 Results based on DT2s

Diffusion T2s (DT2s) are widely used to describe and analyse DWI data. Despite their popularity, DT2s are limited since they model only a single principal direction of diffusion, offering poor representations for complex crossing fibers and potentially inaccurate statistical analysis.

Table 7.2 contains the percentages of lesions detected using the proposed statistical analysis of DT2 models, at each ROI in the motor system of the brain and figure 7.4 visualizes them. Cases where DT2 detected more lesions exceed the DT4 cases. The absence of ground truth, in order to specify the false positive and false negative rates, hardens the task of deriving conclusions.

Name of ROI	**PATIENT 1** **Scan 1** **DT4 (DT2)**	**PATIENT 1** **Scan 2** **DT4 (DT2)**
Pontine crossing tract	23.5% (35.52%)	62.84% (70.49%)
Corticospinal tract R	10.23% (17.61%)	28.41% (34.09%)
Corticospinal tract L	**14.04% 13.48%)**	30.9% (35.96%)
Medial lemniscus R	0.0% (2.33%)	17.44% (23.26%)
Medial lemniscus L	**16.87% (10.84%)**	**20.48% (12.05%)**
Posterior limb of internal capsule R	6.59% (12.57%)	18.76% (30.54%)
Posterior limb of internal capsule L	5.45% (8.18%)	13.21% (22.43%)
Superior corona radiata R	8.48% (19.13%)	23.15% (40.98%)
Superior corona radiata L	9.63% (21.1%)	18.72% (31.39%)

Name of ROI	**PATIENT 2** **Scan 1** **DT4 (DT2)**	**PATIENT 2** **Scan 2** **DT4 (DT2)**
Pontine crossing tract	8.74% (11.48%)	26.78% (42.08%)
Corticospinal tract R	18.18% (22.16%)	19.89% (22.73%)
Corticospinal tract L	**23.03% (21.35%)**	19.1% (24.16%)
Medial lemniscus R	**17.44% (13.95%)**	25.58% (33.72%)
Medial lemniscus L	21.69% (22.89%)	24.1% (36.14%)
Posterior limb of internal capsule R	5.99% (6.59%)	3.59% (4.79%)
Posterior limb of internal capsule L	10.27% (16.98%)	11.53% (16.35%)
Superior corona radiata R	3.91% (8.7%)	4.35% (8.8%)
Superior corona radiata L	5.09% (7.9%)	2.6% (5.74%)

TABLE 7.2: Comparison between DT4 and DT2 (obtained using FSL) statistical analyses. Table shows the percentage of lesions (p-value $\leqslant 0.05$) per ROI for both LIS patients - both scans. Between parentheses the percentage of lesions (p-value $\leqslant 0.05$) derived from the DT2 analysis. Highlighted percentages correspond to cases where DT4 detected more lesions than DT2.

7.2.3 Classical statistical analysis of FA and MD images

In the literature (see *e.g.* [92]), many classic statistical approaches have been proposed to analyze scalar diffusion images, such as FA or MD images, derived from T2 models. Using scalar images is convenient for simple statistical calculations, for example FA/MD histogram analysis per voxel/ROI, voxelwise or ROI-based calculations of z-scores (also known as standard scores) etc.

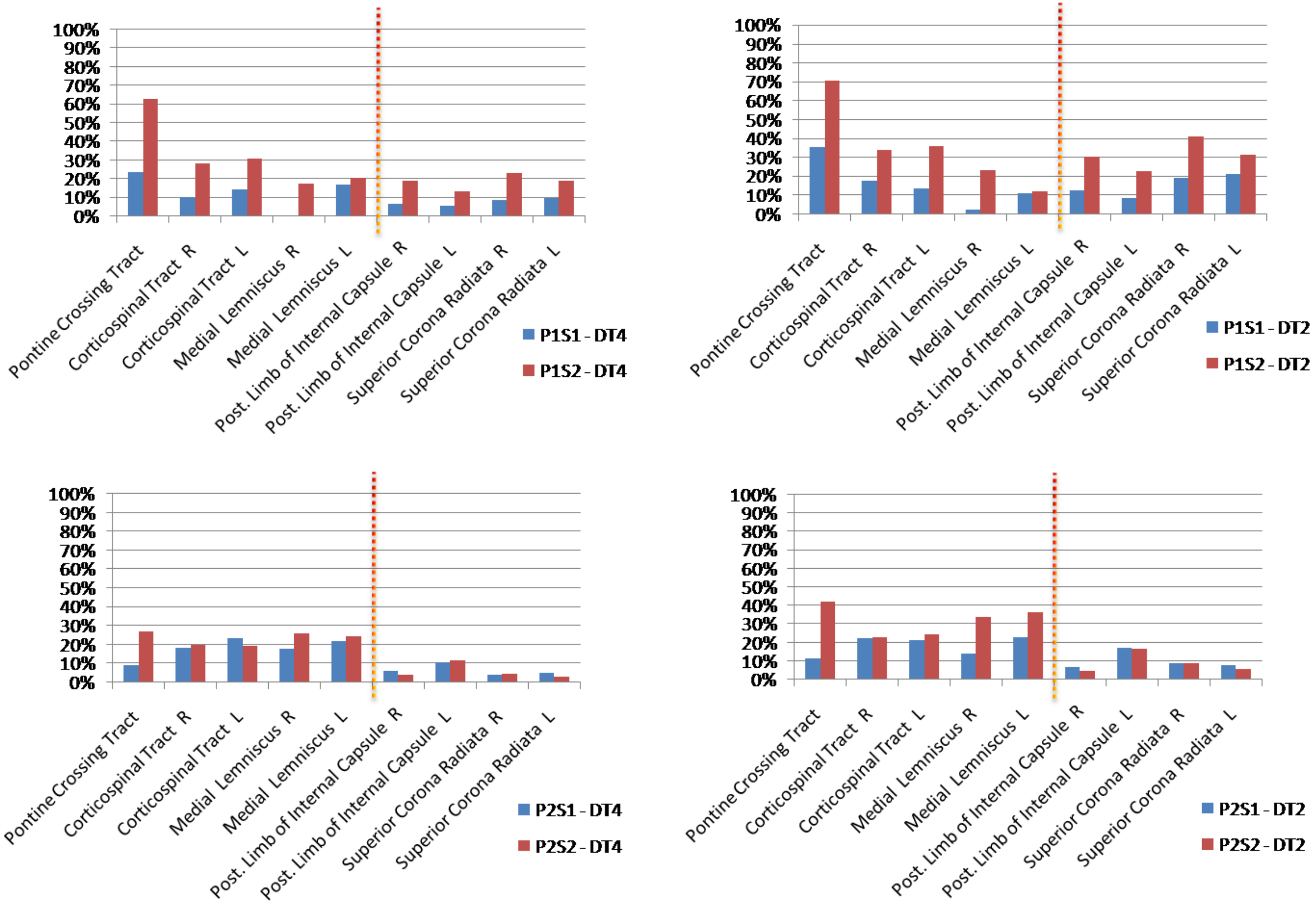


FIGURE 7.4: Plotting the percentages of lesions detected using the proposed method on diffusion T4s and diffusion T2s (as presented in table 7.2). The labels are coded as "P*i*S*j*-*data*" referring to "Patient *i* Scan *j* on specific *data*". The vertical dotted lines separate the two groups of ROIs (ROIs in the bottom part of the brain on the left and ROIs in the middle and upper parts of the brain on the right).

For the purposes of this study, we chose to estimate voxelwise z-scores per patient, based on normal population's mean and standard deviation of FA/MD values of all healthy individuals at each voxel. The z-scores will be post-processed in order to determine the percentages of lesions (*i.e.* voxels with |z-score| $\geqslant 1.96$) in each ROI (using similar process to the proposed method) of each patient. The threshold of 1.96 is equivalent to a p-value of 0.05 in a two-tailed hypothesis.

DT4s p-values *versus* FA z-scores

The statistical analysis of FA images by calculating z-scores for both patients and both data acquisitions is presented in table 7.3 (between parentheses) along with the obtained results of the proposed method on DT4 models. It is noticeable (in both table 7.3 and figure 7.5) that the percentages of lesions based on FA analysis are higher than the corresponding percentages derived from the proposed method.

Name of ROI	**PATIENT 1 Scan 1 DT4 (FA)**	**PATIENT 1 Scan 2 DT4 (FA)**
Pontine crossing tract	23.5% (79.78%)	62.84% (91.8%)
Corticospinal tract R	10.23% (66.48%)	28.41% (74.43%)
Corticospinal tract L	14.04% (61.8%)	30.9% (65.73%)
Medial lemniscus R	0.0% (61.63%)	17.44% (67.44%)
Medial lemniscus L	16.87% (63.86%)	20.48% (63.86%)
Post. limb of internal capsule R	6.59% (32.53%)	18.76% (55.49%)
Post. limb of internal capsule L	5.45% (30.19%)	13.21% (31.03%)
Superior corona radiata R	8.48% (24.35%)	23.15% (29.24%)
Superior corona radiata L	9.63% (17.75%)	18.72% (24.03%)

Name of ROI	**PATIENT 2 Scan 1 DT4 (FA)**	**PATIENT 2 Scan 2 DT4 (FA)**
Pontine crossing tract	8.74% (75.96%)	26.78% (81.97%)
Corticospinal tract R	18.18% (48.86%)	19.89% (54.55%)
Corticospinal tract L	23.03% (37.64%)	19.1% (37.64%)
Medial lemniscus R	17.44% (73.26%)	25.58% (75.58%)
Medial lemniscus L	21.69% (59.04%)	24.1% (71.08%)
Post. limb of internal capsule R	5.99% (42.51%)	3.59% (32.93%)
Post. limb of internal capsule L	10.27% (25.58%)	11.53% (23.27%)
Superior corona radiata R	3.91% (11.3%)	4.35% (11.41%)
Superior corona radiata L	5.09% (31.71%)	2.6% (25%)

Table 7.3: Comparison between DT4 and FA image statistical analyses. Table shows the percentage of lesions (p-value $\leqslant 0.05$) per ROI for both LIS patients - both scans. Between parentheses the percentage of |z-score| $\geqslant 1.96$ based on FA analysis is included.

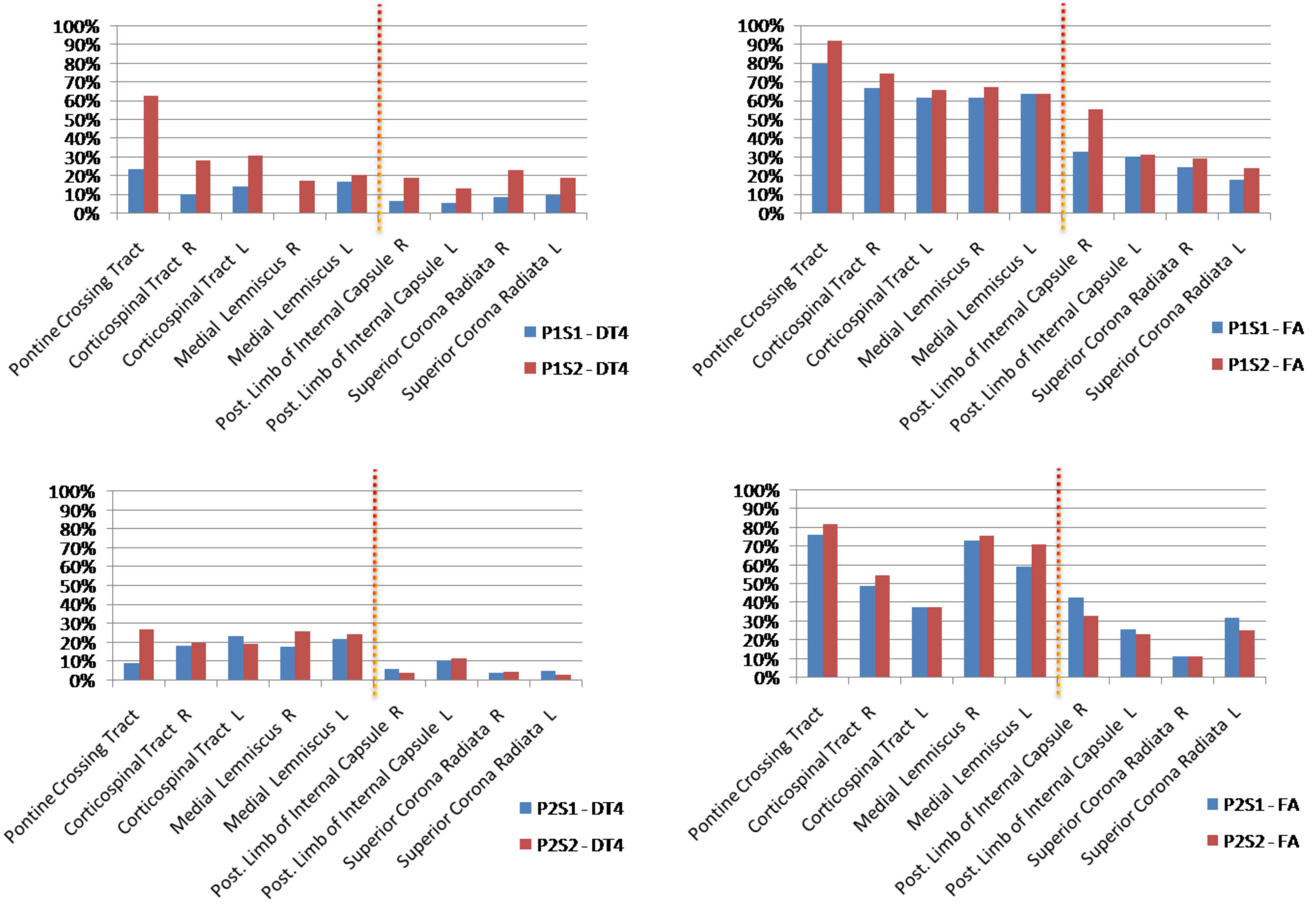


Figure 7.5: Plotting the percentages of lesions detected using the proposed method on diffusion T4s and z-scores on FA images (as presented in table 7.3). The labels are coded as "P*i*S*j*-*data*" referring to "Patient *i* Scan *j* on specific *data*". The vertical dotted lines separate the two groups of ROIs (ROIs in the bottom part of the brain on the left and ROIs in the middle and upper parts of the brain on the right).

Figure 7.6 depicts the evolution of patient 1's FA images through the two scans for three typically affected ROIs (pontine crossing tract and right and left corticospinal tracts), in comparison to a healthy FA template. Due to the absence of any ground truth solution it is hard to say which method is better than the other (it is impossible to measure the false positive and false negative rates per method).

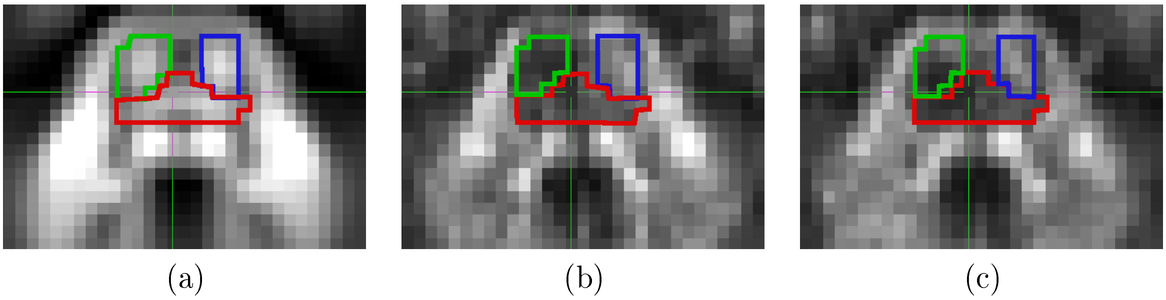

(a) (b) (c)

FIGURE 7.6: FA's axial slices showing the disease's evolution of LIS patient 1 in three ROIs. (a) JHU-FA template, (b) Patient1-scan1 and (c) Patient1-scan2. Red ROI: Pontine crossing tract, green ROI: Corticospinal tract R and blue ROI: Corticospinal tract L. (b)-(c) contain lower FA values (*i.e.* darker colors) in comparison to control image (a).

In favour of the proposed method, we could clarify that T4 diffusion profiles do not collect the same type of information as FA images. Probably, DT4s should be ideally compared to MD images which measure the mean diffusivity across the three main directions, aligned to the three eigenvectors resulting from the spectral analysis of the T2 matrix (see eq. 2.10).

DT4s p-values *versus* MD z-scores

Table 7.4 contains between parentheses the statistical analysis of MD images by calculating z-scores, next to the obtained results of the proposed method on DT4s and figure 7.7 visualizes them. Generally speaking, MD's percentages in many ROIs are lower than the corresponding percentages obtained in the FA case (of course there are some exceptions, such as the last two rows of the tables, referring to Superior corona radiata R and L), signifying that even two scalar measurements derived from the same T2 models can produce different statistics, pointing that the absence of a ground truth solution, once again, makes the evaluation process hard for safe conclusions.

In the next section, an evaluation of the proposed statistical approach on fODF T4 data will be presented, by measuring its performance on a leave-one (normal datum)-out scheme. Furthermore, a set of variations of the proposed distance (eq. 2.30) will be also evaluated.

7.2.4 Leave-one-out Evaluation Scheme in the fODF T4 Case

The low obtained percentages of lesions, for example in the fODF T4 case, increased our interest to measure the ability of the proposed statistical test (based on the proposed tensor metric of eq. 2.30) to correctly classify every unseen normal individual as a healthy person.

Name of ROI	**PATIENT 1 Scan 1 DT4 (MD)**	**PATIENT 1 Scan 2 DT4 (MD)**
Pontine crossing tract	23.5% (54.65%)	62.84% (92.9%)
Corticospinal tract R	10.23% (56.82%)	28.41% (71.59%)
Corticospinal tract L	14.04% (41.57%)	30.9% (62.36%)
Medial lemniscus R	0.0% (19.77%)	17.44% (46.51%)
Medial lemniscus L	16.87% (44.58%)	20.48% (51.81%)
Post. limb of internal capsule R	6.59% (34.73%)	18.76% (63.07%)
Post. limb of internal capsule L	5.45% (37.74%)	13.21% (44.23%)
Superior corona radiata R	8.48% (71.85%)	23.15% (91.41%)
Superior corona radiata L	9.63% (66.88%)	18.72% (81.49%)

Name of ROI	**PATIENT 2 Scan 1 DT4 (MD)**	**PATIENT 2 Scan 2 DT4 (MD)**
Pontine crossing tract	8.74% (27.32%)	26.78% (56.83%)
Corticospinal tract R	18.18% (46.02%)	19.89% (51.14%)
Corticospinal tract L	23.03% (46.63%)	19.1% (51.12%)
Medial lemniscus R	17.44% (43.02%)	25.58% (61.63%)
Medial lemniscus L	21.69% (56.63%)	24.1% (77.11%)
Post. limb of internal capsule R	5.99% (35.93%)	3.59% (26.55%)
Post. limb of internal capsule L	10.27% (33.54%)	11.53% (35.22%)
Superior corona radiata R	3.91% (38.48%)	4.35% (50.54%)
Superior corona radiata L	5.09% (47.73%)	2.6% (48.05%)

Table 7.4: Comparison between DT4s and MD image statistical analyses. Table shows the percentage of lesions ($p\text{-value} \leqslant 0.05$) per ROI for both LIS patients - both scans. Between parentheses the percentage of $|\text{z-score}| \geqslant 1.96$ based on MD analysis is included.

This is done with a serial leave one normal dataset out of the training procedure, during the estimation of the reduced space of the normal population (*via* Isomap).

At this point, we should clarify that our new working scope is the whole brain, instead of a single voxel, since we are interested in comparing a whole normal brain *versus* the normal population. As a result, a single inter-point distance matrix will be constructed and each of the matrix's elements will take into account the sum of all (or M largest) voxelwise distances throughout the whole volume of the brain's WM.

Each normal datum left out of the training step as well as the four abnormal datasets will be statistically compared to the current normal population. An example illustrating the correct classification case and an incorrect classification are presented in figure 7.8.

Initially, we chose to consider all voxels. In this case, the reduced space was unfortunately disturbed by the small distances related to similar voxels in the sum. This led to very spread normal populations where the abnormal points were also included in the mass of the normal population. It was impossible to distinguish them as pathological cases.

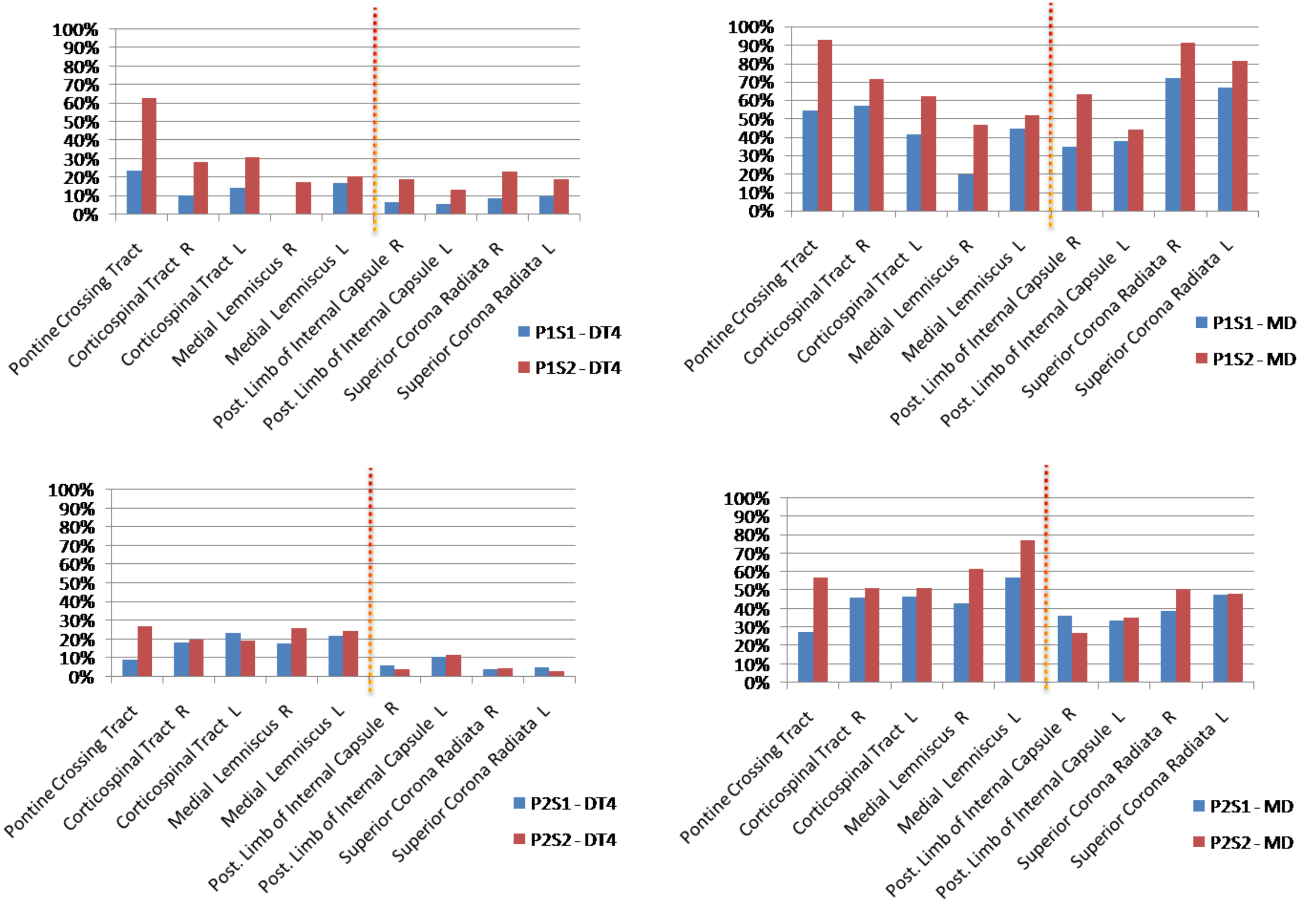


FIGURE 7.7: Plotting the percentages of lesions detected using the proposed method on diffusion T4s and z-scores on MD images (as presented in table 7.4). The labels are coded as "PiSj-*data*" referring to "Patient i Scan j on specific *data*". The vertical dotted lines separate the two groups of ROIs (ROIs in the bottom part of the brain on the left and ROIs in the middle and upper parts of the brain on the right).

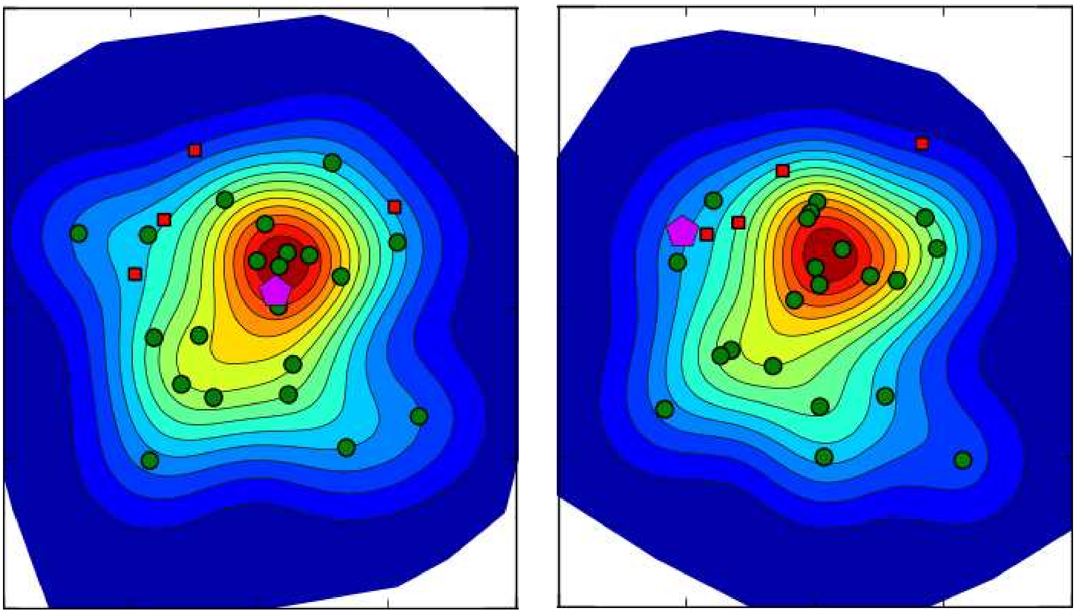


Figure 7.8: Visualization of the leave-one-out evaluation. Left figure corresponds to the correct classification, since the unseen normal (*i.e.* purple pentagon) has larger p-value than the 4 abnormal (*i.e.* red squares) individuals and it is located in the core of the normal population which contains the green points. On the right, the wrong classification is presented, since the normal point (purple one) has lower p-value than 3 out of 4 abnormal points. The colorful background corresponds to the PDF of the normal population.

Alternatively, we thought to sum only the M largest voxelwise distances, corresponding to the most significantly different voxels in the brain for each given couple of individuals. Several values for $M \in \{10, 50, 100, 500, 1000, 2000, 4000\}$ were examined and the number of normal datasets which were correctly classified as healthy people were counted. An unseen normal datum is considered as correctly classified if its p-value is larger than all the four p-values related to the four abnormal datasets. The best performance is for $M = 1000$, where $14/22 = 63.6\%$ normal individuals were correctly classified. The corresponding p-values of the $M = 1000$ test can be found in table 7.5. We should mention that the majority of the p-values (not for normal points which was expected, but for most of the abnormals) are greater than 0.05, meaning that the evaluation scheme did not work very well. Probably, the low percentages of lesions detected globally in the brain using fODF profiles is one reason. Moreover, a careful study in order to estimate an alternative abnormality threshold is required, in order to determine the correct classifications of the pathological brains (*i.e.* as abnormals).

Variations of the Proposed Tensor Metric tested for the LIS disease

Thinking of possible ways to improve the performance of the method, we turned our attention to the definition of the proposed tensor metric that we selected to compare the fODF profiles. If someone carefully observes the proposed tensor metric in eq. 2.30, she will notice that two interesting degrees of freedom can be derived, for example considering the parameters

k, p in equation 7.5 ($N = 242$ is kept fixed).

$$dist\,(d_1, d_2) \simeq \left(\sum_{i=1}^{N} \left| \log \frac{d_1(\theta_i, \phi_i)}{d_2(\theta_i, \phi_i)} \right|^k \sin(\theta_i)\,\Delta\theta\Delta\phi \right)^p . \tag{7.5}$$

Therefore, investigating the performance of the proposed method, in the previously described leave-one-out assessment, using a variety of parameterizations in eq. 7.5 is worth testing.

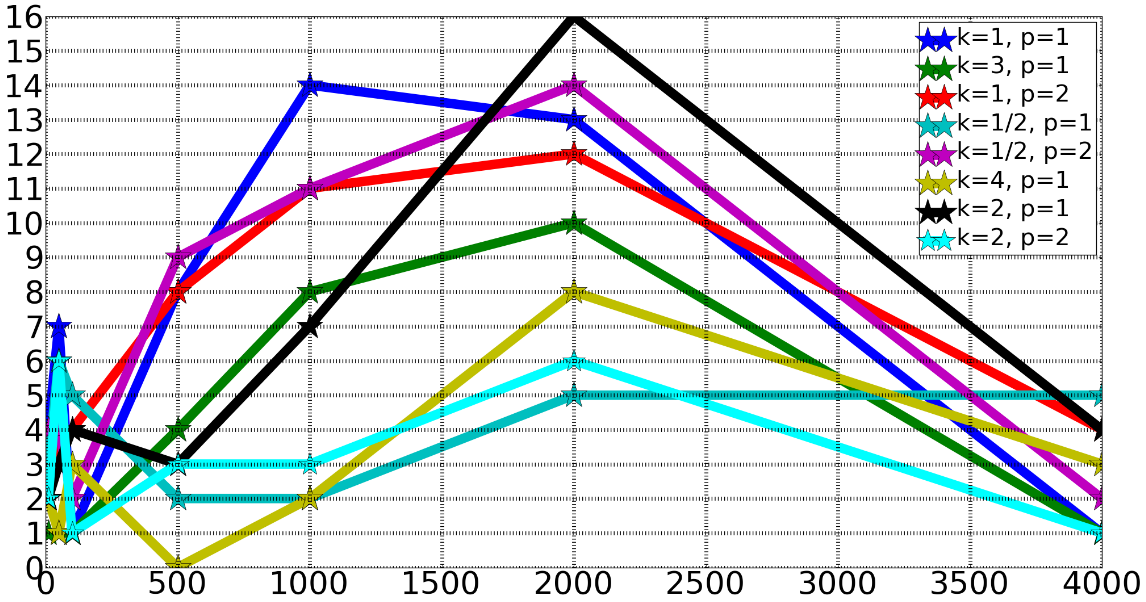


FIGURE 7.9: Performance of several variations of the proposed tensor metric (eq. 2.30) in the leave-one-out evaluation scheme. The vertical axis of the figure corresponds to the number of correctly classified unseen normal data, while the horizontal axis contains the number of the maximum M voxelwise distances included in the inter-point distance matrix, for $M \in \{10, 50, 100, 500, 1000, 2000, 4000\}$. The initial version of the proposed tensor metric ($k = 1$, $p = 1$) achieved a score of 14/22 correct classifications, while an observation in the figure will conclude that for $k = 2$ and $p = 1$ in eq. 7.5 and $M = 2000$ top maximum voxelwise distances, outperforms with score $= 16/22$ correct classifications.

Figure 7.9 depicts the performance of different sets of k, p values in eq. 7.5 for a given number $M \in \{10, 50, 100, 500, 1000, 2000, 4000\}$. The best performance was achieved by the $k = 2$, $p = 1$ parameterization, introducing $M = 2000$ largest voxelwise distances in the inter-point distance matrix, concluded into $16/22 = 72.7\%$ of correctly classified unseen normal datasets, in comparison to the initial formulation of the tensor metric whose performance was equal to $14/22 = 63.6\%$.

7.3 Partial Conclusion

Individual *versus* normal population comparisons has the potentiality to assist the physicians through patient follow-up procedures. In this chapter we proposed a statistical approach to provide a solution to this problem.

Statistical analysis of certain ROIs sensitive to LIS concluded that areas close to the spinal cord (such as the pontine crossing tract, left and right corticospinal tracts and left and right medial lemniscus) contain higher percentages of lesions, in comparison to areas in the middle and upper parts of the brain, connected with the spinal cord (*e.g.* left and right posterior limb of internal capsule and left and right superior corona radiata). Furthermore, patients react differently in the second case.

The evaluation process of the experimental results signified that it is hard to make safe conclusions about which method performs better than the other, in the absence of ground truth solution. A thoughtful clue is that higher order tensor models are more detailed and as a consequence more capable to capture the disease's specificity, due to the complexity of the model, than naive T2 models, or scalar measures, such as FA/MD images. Statistical analysis in synthetic cases could be probably useful to evaluate the performance of the tested methods.

As part of future work, it would be interesting to measure the percentages of lesions per ROI using the new variation of the proposed tensor metric that was found to outperform the selected definition of the tensor metric in subsection 7.2.4.

Left-out Normal point k	**p-value Normal k**	**p-value Patient 1 / Scan 1**	**p-value Patient 1 / Scan 2**	**p-value Patient 2 / Scan 1**	**p-value Patient 2 / Scan 2**
1	0.2386	0.4823	**0.0173**	0.4738	0.4738
2	0.9264	0.4331	**0.0138**	0.4497	0.4536
3	0.2112	0.3835	**0.0197**	0.4503	0.453
4	0.9296	0.8238	**0.0145**	0.7118	0.7056
5	0.1681	0.9464	**0.0073**	0.2827	0.2938
6	0.9321	0.0705	**0.007**	0.3272	0.6324
7	0.9935	0.9544	**0.03**	0.3205	0.5632
8	0.7851	0.4616	**0.0095**	0.4505	0.4518
9	0.5729	0.5152	**0.0126**	0.4193	0.4136
10	0.3166	0.5387	**0.0219**	0.5088	0.5996
11	0.9438	0.8047	**0.018**	0.7714	0.7611
12	0.8935	0.574	0.0665	0.5636	0.5638
13	0.1034	0.6	**0.0065**	0.3946	0.3902
14	0.7133	0.8212	0.0525	0.6835	0.6769
15	0.9995	0.528	**0.011**	0.5001	0.5033
16	0.952	0.9213	**0.016**	0.4644	0.4679
17	0.6195	0.4967	**0.0095**	0.4215	0.585
18	0.8497	0.9975	**0.0063**	0.7159	0.7049
19	0.4596	0.4386	**0.0353**	0.5598	0.5614
20	0.9572	0.4123	**0.0404**	0.4685	0.4718
21	0.637	0.1848	**0.0208**	0.2023	0.2076
22	0.9233	0.9129	**0.0115**	0.3958	0.4829

Table 7.5: Estimating p-values in the case of the best performance ($M = 1000 \rightarrow 14/22$ correctly classified normal points) in the leave-one-out evaluation scheme. The grey level rows correspond to the correctly classified normal points (*i.e.* normal's p-value is greater than all four abnormals). Initially, working in the same direction as previously, a p-value lower than 0.05 would signify the point as pathologically affected. It is noticeable that all normal points are related to greater than 0.05 p-values (*i.e.* not affected by LIS, as it was expected). Moreover, the majority of the abnormal p-values are greater than 0.05 (except from Patient 1 / Scan 2). This might be a clue to choose a more suitable abnormality threshold than 0.05.

Chapter 8

Conclusion and Perspectives

In this final chapter of the dissertation, several remarkable issues which should be kept in mind since they have been concluded through studying the problems of biomarker extraction and patient follow-up, completed by the development of the proposed methods, are discussed. Furthermore, many directions as part of future work are highlighted and presented.

8.1 Discussion

Comparing data from different subjects obliges us to normalize the data in a common reference space. In the case of DW-MRI data, or tensor images, due to the specificity of the data, a simple spatial registration is insufficient, lacking of a mandatory step, known as reorientation, in order to align the registered data to the new underlying fiber orientation. At this point, we should mention that registering tensor images (through spatial registration of each tensor coefficient separately and finally collecting all registered coefficients in one volume) is much more exposed to distortion than registering the raw DW-MRI data. Imagine that a distortion caused to some of the tensor coefficients (on account of registration errors) will have greater impact on altering the diffusivity or fODF profiles across several directions, than absorbing noise in a few directions in the DW-MRI dataset because of DW-MRI registration. Although in the beginning of this thesis we started working with T4 normalizations, thereafter a quite-promising method for non-linear DWI normalization was proposed in 2013. In addition, the reorientation of a crossing T4, using T2 decompositions and reorientations, in order to apply a transformation affecting two peaks of diffusion (*e.g.* principal directions of diffusion) to get very close to each other, will potentially result into losing mistakenly one of the peaks (altering totally the underlying fiber structure, as it is derived from our study in chapter 3). For these reasons, we chose to normalize the raw DW-MRI data.

Our selection to represent the DW-MRI data with HOTs, such as the T4 model, allowed us to increase the robustness and sensitivity of the proposed statistical models by describing the data with more accurate models than T2s, especially in cases of crossing fibers. Additionally, we should not forget that a more accurate model will eventually lead us feasibly to earlier diagnosis.

Following the suggestion of Verma *et al.* [165] to perform statistical analysis in a reduced space seems crucial and reasonable for many reasons. First of all, the definition of the inter-point distance matrix using a proper distance, such as the proposed tensor metric (eq. 2.30), permitted us to introduce not only information about the diffusion, but also about the orientation of the diffusion through the integration on the unit sphere. Secondly, Isomap particularly, assisted us to find the non linear tensor's subspace, by adding geodesic properties to the estimation of the reduced space through the embedded graph theory. Moreover, our suggestion to deal with any registration error left at this point by finding the best matched patches (for each coupled combination of our data), contributed to eliminate as much as possible any potential registration error and produced more sensitive models, since we observed that smoothing the measurements (*e.g.* fODF profiles), which is an alternative popular technique followed by many approaches, can lead to wrong conclusions due to over-smoothing effects and important information lost. Additionally, statistical analysis based on Random Forest Classifiers, which in general are assumed to be powerful tools for high dimensional data, shown less efficiency than expected, due to the high complexity of the tensor models, and they can be overcome by RFs benefiting from a dimensionality reduction step, in advance.

Another interesting topic worth mentioning concerns the ability of the proposed statistical approaches to analyze the levels of abnormality in the pathological data, independently of the size of the abnormal population, given a well-grown normal population (which is feasible in general). To achieve our goals, the statistical analysis is divided in two approaches, one for the case where the number of the treated abnormal datasets is prolific to build an abnormal population and the other case where it is not possible to capture the variability of the abnormal population.

In the case where the pathological population, associated with a certain disease, can be built with an abundant number of patients (*e.g.* application to NMO disease presented in chapter 6), we proposed to perform voxelwise populations comparisons which offers us the potentiality to construct an atlas of abnormality that will characterize the affection of a disease of our interest in every part of the brain. Modelling each population with the aid of GMM in the reduced space, followed by the definition of a permutation test, based on a plethora of label shufflings of the points, that could approximate the distribution of the measured distance between GMMs PDFs (*i.e.* statistic of interest), allowed us to estimate a p-value per voxel, and particularly a HPD interval for each p-value, concluding if the distance related to the true labeling of the points is an extreme value given the distribution of the distances that is produced randomly *via* label shufflings. At this point we should emphasize that many statistical approaches in the literature are reduced to estimate a single p-value, which is an approximation, without justifying the confidence of their estimation by calculating the interval that each p-value is enclosed in.

On the other hand, the case of lacking enough patients, an impracticable situation to construct robustly the abnormal population, lead us to the formulation of the "individual *versus* normal population" problem (*e.g.* application to LIS disease presented in chapter 7). In this particular case, voxelwise statistics will be estimated by fitting a GMM only to the normal population, while each abnormal datum will be examined by measuring its PDF value (*i.e.* statistic of interest) given the estimated distribution of the normal population. Moreover, the corresponding p-value's HPD interval can be calculated relying on Monte Carlo simulations. Generating randomly samples from the normal (*i.e.* healthy) GMM, assists us to compare each PDF related to the abnormal points to the samples' PDFs. Thereby, an outgrowth of the proposed statistical analysis can be the patient follow-up, throughout several examinations.

Finally, we should emphasize that the diagnosis of a new incoming datum (*i.e.* patient) can be performed, either using the extracted biomarkers (calculated through populations comparisons), or by running individual statistical comparisons *versus* the normal population. Afterwards, the new patient is classified to the normal or to the abnormal population, without needing to re-define any population again.

8.2 Future Work

In the end of every constructive research, suggested directions for future work should be indicated. Therefore, following our acquired knowledge through this thesis on DW-MRI data processing and statistical analysis, we come up with many zestful points.

To begin with, in this thesis, data normalization was achieved either by registering the DW-MRI data followed by the reorientation of the embedded b-vectors (limited to apply only the rotation part of the estimated non-linear transformation), or through serial registrations of every T4 coefficient, thereafter, resynthesis of the tensor models by collecting all registered tensor coefficients into one volume and finally reorientation of the registered T4 models with the aid of the methods presented in chapter 3 using the whole transformation (*e.g.* SD+PPD, or HD+PPD) or the rotation part (*e.g.* FS). In both mentioned ways to normalize our data, we focused on completing the process on the same type of data, but in fact, it is possible and worthy to be tested to register the DW-MRI data and then to fit T4 models on the registered DW-MRI in order to reorientate eventually the T4 models (by using methods from chapter 3).

The next points are referred to the statistical model. The evaluation of the proposed tensor metric along with its variations presented in subsection 7.2.4 concluded to the existence of a particular variation which manage to outperform our initial definition. As a consequence, it will be interesting to estimate the reduced space using the best variation of the tensor metric. Maybe the discrepancy between the control and the pathological points can be increased, resulted into more sensitive statistical analysis.

Thirdly, we proposed to fit one Gaussian kernel on each point in the reduced space, in order to describe every population with a flexible and more representative GMM model. Another fruitful approach, a little bit more complicated than our initial thought, but can potentially avoid any occasional overfitting problem, could be to cluster neighboring points into similar groups, supposing that these points are produced by the same single Gaussian kernel. In this case, the corresponding Gaussian kernel could be defined by the following mean $\boldsymbol{\mu}_i = \frac{1}{J}\sum_{j=1}^{J} x_j$ (if x_j are the J points included in the same cluster) and it will be related to a covariance matrix equals to J times the covariance of a single point in the same population.

Fourthly, our study in this thesis was concentrated in examining ROIs proposed in the literature as pathologically affected by the NMO and LIS diseases. In the case of an inflammatory disease, such as, NMO, multiple sclerosis, Alzheimer etc., investigating ROIs outside the already known related areas, could lead to outstanding and innovative results, whether new areas can be extracted as pathologically connected to the disease of our interest. Moreover, it can be useful through the whole procedure of disease staging and patient follow-up.

In the fifth point, the application of the proposed methods on fiber tracts and connectomes, instead of voxels, can be also fruitful. In this case, the inter-point distance matrix should be defined including distances between fiber tracts or connectomes (depending the approach).

Furthermore, the assessment to detect the statistical significance of differences in the level of every direction in the diffusion/fODF profiles, related to the most significantly different voxels in the brain, is indicated *via* the proposed statistical models. In other words, it is possible to detect which directions in the diffusion/fODF profiles contributed in characterizing the voxel as a biomarker.

To conclude, evaluating the abilities of the proposed statistical approaches to perform early diagnosis remain to be tested, under the orientated supervision provided from the neurologists.

Appendix A

Multivariate Two-sample Hotelling T^2 Test

Hotelling T^2 test is the generalization of the Student's t-test [85]. Multivariate two-sample Hotelling T^2 test compares two populations **X, Y** by assuming that both populations follow Normal distributions with different means, but the same covariance matrix. Let us consider N i.i.d. data assigned to population $\mathbf{X}$ denoted as $\{X_1, X_2, \dots, X_N\}$, $X_i \in \mathbb{R}^p$, $\forall i = 1, \dots, N$ and M i.i.d. data belong to the second population $\mathbf{Y}$, $\{Y_1, Y_2, \dots, Y_M\}$, $Y_j \in \mathbb{R}^p$, $\forall j = 1, \dots, M$.

The means μ_X, μ_Y of the two populations correspond to vectors of size $p \times 1$ and are equal to:

$$\mu_X = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad \mu_Y = \frac{1}{M}\sum_{j=1}^{M} Y_j, \tag{A.1}$$

while the sample covariance matrices S_X, S_Y are equal to:

$$S_X = \frac{1}{N-1}\sum_{i=1}^{N} (X_i - \mu_X)(X_i - \mu_X)^T, \quad S_Y = \frac{1}{M-1}\sum_{j=1}^{M} (Y_j - \mu_Y)(Y_j - \mu_Y)^T. \tag{A.2}$$

Due to the assumption that both populations have equal covariance matrix $S_p^\star$, the covariance matrices of the samples S_X, S_Y can help us to estimate $S_p^\star$ considering $\widehat{S_p}$:

$$\widehat{S_p} = \frac{(N-1)S_X + (M-1)S_Y}{N+M-2}. \tag{A.3}$$

The testing Null Hypothesis is considered as $H_0 : \mu_X = \mu_Y$, meaning that the two populations are equal if their means are equal.

The T^2 statistic signifies the differences in the populations by comparing their means and is calculated as:

$$T^2 = (\mu_X - \mu_Y)^T \left\{ S_p \left(\frac{1}{N} + \frac{1}{M} \right) \right\}^{-1} (\mu_X - \mu_Y) \,. \tag{A.4}$$

At this point the T^2 statistic is transformed to F-statistic using the following expression:

$$F_{stat} = \frac{N+M-p-1}{p(N+M-2)} T^2 \sim F_{p,N+M-p-1}, \tag{A.5}$$

where the PDF of the F-distribution is given by

$$f\left(x; d_1, d_2\right) = \frac{\sqrt{\frac{(d_1 x)^{d_1} \; d_2^{d_2}}{(d_1 x + d_2)^{d_1+d_2}}}}{x \;\, Beta\left(\frac{d_1}{2}, \frac{d_2}{2}\right)} \tag{A.6}$$

As a result the corresponding p-value is equal to $1 - CDF_{F_{p,N+M-p-1}}(F_{stat})$.

Bibliography

[1] Agosta, F., Galantucci, S., Svetel, M., Lukic, M. J., Copetti, M., Davidovic, K., Tomic, A., Spinelli, E. G., Kostic, V. S., and Filippi, M. (2014). Clinical, cognitive, and behavioural correlates of white matter damage in progressive supranuclear palsy. *Journal of Neurology*, 261(5):913–924.

[2] Alexander, D. C., Pierpaoli, C., Basser, P. J., and Gee, J. C. (2001). Spatial transformations of diffusion tensor magnetic resonance images. *Medical Imaging, IEEE Transactions on*, 20(11):1131–1139.

[3] Alexander, D. C., Zikic, D., Zhang, J., Zhang, H., and Criminisi, A. (2014). Image Quality Transfer via Random Forest Regression: Applications in Diffusion MRI. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Pt 3)*, volume 8675 of *Lecture Notes in Computer Science*, pages 225–232. Springer.

[4] Anderson, A. W. (2005). Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(5):1194–1206.

[5] Arsigny, V. (2006). *Processing Data in Lie Groups: An Algebraic Approach. Application to Non-Linear Registration and Diffusion Tensor MRI.* Ph.D. Thesis, INRIA, Sophia-Antipolis, France.

[6] Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2006). Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421.

[7] Ashburner, J. and Friston, K. J. (2000). Voxel-based morphometry–the methods. *NeuroImage*, 11(6 Pt 1):805–821.

[8] Ashburner, J. and Friston, K. J. (2001). Why voxel-based morphometry should be used. *NeuroImage*, 14(6):1238–1243.

[9] Baird, A. E. and Warach, S. (1998). Magnetic resonance imaging of acute stroke. *Journal of Cerebral Blood Flow and Metabolism*, 18(6):583–609.

[10] Barmpoutis, A., Bing, J., and Vemuri, B. C. (2009a). Adaptive kernels for multi-fiber reconstruction. In *Proceedings of the Information Processing in Medical Imaging conference (IPMI)*, volume 5636 of *Lecture Notes in Computer Science*, pages 338–349. Springer.

[11] Barmpoutis, A., Hwang, M. S., Howland, D., Forder, J. R., and Vemuri, B. C. (2009b). Regularized positive-definite fourth-order tensor field estimation from DW-MRI. *NeuroImage*, 45:153–162.

[12] Barmpoutis, A., Jian, B., Vemuri, B. C., and Shepherd, T. M. (2007a). Symmetric positive 4th order tensors & their estimation from diffusion weighted MRI. In *Proceedings of the Information Processing in Medical Imaging conference (IPMI)*, volume 4584 of *Lecture Notes in Computer Science*, pages 308–319. Springer.

[13] Barmpoutis, A. and Vemuri, B. C. (2009). Groupwise registration and atlas construction of 4th-order tensor fields using the R+ Riemannian metric. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 5761 of *Lecture Notes in Computer Science*, pages 640–647. Springer.

[14] Barmpoutis, A. and Vemuri, B. C. (2010). A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 1385–1388.

[15] Barmpoutis, A., Vemuri, B. C., and Forder, J. R. (2007b). Registration of high angular resolution diffusion MRI images using 4th order tensors. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 4791 of *Lecture Notes in Computer Science*, pages 908–915. Springer.

[16] Basser, P. J. (1995). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8:333–344.

[17] Basser, P. J. (1998). Fiber-tractography via diffusion tensor MRI (DT-MRI). In *Proceedings of the 6th Annual Meeting of International Society for Magnetic Resonance in Medicine (ISMRM)*, page 1226.

[18] Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR Diffusion Tensor Spectroscopy and Imaging. *Biophysical Journal*, 66:259–267.

[19] Basser, P. J., Mattiello, J., Turner, R., and Le Bihan, D. (1993). Diffusion tensor echo-planar imaging of human brain. In *Proceedings of the 1st Annual Meeting of International Society for Magnetic Resonance in Medicine (ISMRM)*, pages 584–562.

[20] Basser, P. J. and Pajevic, S. (2007). Spectral decomposition of a 4th-order covariance tensor: Applications to diffusion tensor MRI. *Signal Processing*, 87(2):220–236.

[21] Basser, P. J. and Pierpaoli, C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance*, 111(3):209–219.

[22] Betten, J. (1982). Integrity basis for a second-order and a fourth-order tensor. *International Journal of Mathematics and Mathematical Sciences*, 5(1):87–96.

[23] Biswas, M. and Ghosh, A. K. (2014). A nonparametric two-sample test applicable to high dimensional data. *Journal of Multivariate Analysis*, 123(C):160–171.

[24] Bloy, L. (2012). *Development of High Angular Resolution Diffusion Imaging Analysis Paradigms for the Investigation of Neuropathology.* Ph.D. Thesis, University of Pennsylvania, USA.

[25] Bloy, L., Ingalhalikar, M., Eavani, H., Roberts, T. P., Schultz, R. T., and Verma, R. (2011). HARDI based pattern classifiers for the identification of white matter pathologies. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Pt 2)*, volume 6892 of *Lecture Notes in Computer Science*, pages 234–241. Springer.

[26] Bloy, L., Ingalhalikar, M., Eavani, H., Schultz, R. T., Roberts, T. P., and Verma, R. (2012). White matter atlas generation using HARDI based automated parcellation. *NeuroImage*, 59(4):4055–4063.

[27] Boisgontier, H., Noblet, V., Renard, F., Heitz, F., Rumbach, L., and Armspach, J. P. (2009). Statistical detection of longitudinal changes between apparent diffusion coefficient images: Application to multiple sclerosis. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 5761 of *Lecture Notes in Computer Science*, pages 959–966. Springer.

[28] Bouchon, A., Noblet, V., Heitz, F., Lamy, J., Blanc, F., and Armspach, J. P. (2014). General linear models for group studies in diffusion tensor imaging. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 1339–1342.

[29] Bouts, M. J. R. J. y., Tiebosch, I. A. C. W., van der Toorn, A., Viergever, M. A., Wu, O., and Dijkhuizen, R. M. (2013). Early identification of potentially salvageable tissue with MRI-based predictive algorithms after experimental ischemic stroke. *Journal of Cerebral Blood Flow and Metabolism*, 33(7):1075–1082.

[30] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

[31] Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees.* Chapman & Hall, New York.

[32] Buchel, C., Raedler, T., Sommer, M., Sach, M., Weiller, C., and Koch, M. A. (2004). White matter asymmetry in the human brain: a diffusion tensor MRI study. *Cerebral Cortex*, 14(9):945–951.

[33] Caruyer, E. and Verma, R. (2015). On facilitating the use of HARDI in population studies by creating rotation-invariant markers. *Medical Image Analysis (in press)*, 20(1):87–96.

[34] Catani, M., Howard, R. J., Pajevic, S., and Jones, D. K. (2002). Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage*, 17(1):592–617.

[35] Catani, M. and Thiebaut de Schotten, M. (2012). *Atlas of human brain connections.* Oxford University Press, Oxford.

[36] Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J., and Evans, A. C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cerebral Cortex*, 18(10):2374–2381.

[37] Chyzhyk, D., Dacosta-Aguayo, R., Mataro, M., and Grana, M. (2015). An active learning approach for stroke lesion segmentation on multimodal MRI data. *Neurocomputing*, 150(A, SI):26–36.

[38] Commowick, O., Fillard, P., Clatz, O., and Warfield, S. K. (2008). Detection of DTI white matter abnormalities in multiple sclerosis patients. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Pt 1)*, volume 5241 of *Lecture Notes in Computer Science*, pages 975–982. Springer.

[39] Commowick, O., Maarouf, A., Ferré, J. C., Ranjeva, J. P., Edan, G., and Barillot, C. (2015). Diffusion MRI Abnormalities Detection with Orientation Distribution Functions: a Multiple Sclerosis Longitudinal Study. *Medical Image Analysis*, page (in press).

[40] Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1990). *Introduction to Algorithms.* Cambridge, MA: MIT Press.

[41] Criminisi, A., Shotton, J., and Konukoglu, E. (2011). Decision forests for classification, regression, density estimation, manifold learning and semi–supervised learning. Technical Report MSR-TR-2011-114, Microsoft Technical Report, Microsoft Research.

[42] Criminisi, A., Shotton, J., and Konukoglu, E. (2012). Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends in Computer Graphics and Vision*, 7(2–3):81–227.

[43] Daurignac, E., Toga, A. W., Jones, D. K., Aronen, H. J., Hommer, D. W., Jernigan, T. L., Krystal, J. H., and Mathalon, D. H. (2005). Applications of morphometric and diffusion tensor magnetic resonance imaging to the study of brain abnormalities in the alcoholism spectrum. *Alcoholism - Clinical and Experimental Research*, 29(1):159–166.

[44] Dell'acqua, F., Rizzo, G., Scifo, P., Clarke, R. A., Scotti, G., and Fazio, F. (2007). A Model-Based deconvolution approach to solve fiber crossing in Diffusion-Weighted MR imaging. *Biomedical Engineering, IEEE Transactions on*, 54(3):462–472.

[45] Descoteaux, M. (2008). *IRM de Diffusion à Haute Résolution Angulaire: Estimation Locale, Segmentation et Suivi de Fibres.* Ph.D. Thesis (in French), University of Nice-Sophia Antipolis, France.

[46] Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2006). Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. *Magnetic Resonance in Medicine*, 56(2):395–410.

[47] Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510.

[48] Du, J., Goh, A., Kushnarev, S., and Qiu, A. (2014). Geodesic regression on orientation distribution functions with its application to an aging study. *NeuroImage*, 87:416–426.

[49] Du, J., Goh, A., and Qiu, A. (2012). Diffeomorphic metric mapping of high angular resolution diffusion imaging based on Riemannian structure of orientation distribution functions. *Medical Imaging, IEEE Transactions on*, 31(5):1021–1033.

[50] Du, J., Goh, A., and Qiu, A. (2013). Bayesian atlas estimation from high angular resolution diffusion imaging (HARDI). In *Geometric Science of Information*, volume 8085 of *Lecture Notes in Computer Science*, pages 149–157. Springer.

[51] Duarte-Carvajalino, J., Sapiro, G., Harel, N., and Lenglet, C. (2013). A framework for linear and non-linear registration of diffusion-weighted MRIs using angular interpolation. *Frontiers in Neuroscience*, 7(41).

[52] Fellah, S., Caudal, D., De Paula, A. M., Dory-Lautrec, P., Figarella-Branger, D., Chinot, O., Metellus, P., Cozzone, P. J., Confort-Gouny, S., Ghattas, B., Callot, V., and Girard, N. (2013). Multimodal MR Imaging (Diffusion, Perfusion, and Spectroscopy): Is It Possible to Distinguish Oligodendroglial Tumor Grade and 1p/19q Codeletion in the Pretherapeutic Diagnosis? *American Journal of Neuroradiology (AJNR)*, 34(7):1326–1333.

[53] Filippi, M., Cercignani, M., Inglese, M., Horsfield, M. A., and Comi, G. (2001). Diffusion tensor magnetic resonance imaging in multiple sclerosis. *Neurology*, 56(3):304–311.

[54] Filippi, M., Horsfield, M. A., Rovaris, M., Yousry, T., Rocca, M. A., Baratti, C., Bressi, S., and Comi, G. (1998). Intraobserver and interobserver variability in schemes for estimating volume of brain lesions on MR images in multiple sclerosis. *American Journal of Neuroradiology (AJNR)*, 19(2):239–244.

[55] Fink, D. G. and Christiansen, D. (1989). *Electronics engineers' handbook*. McGraw-Hill.

[56] Fletcher, P. T. and Joshi, S. (2004). Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*, volume 3117 of *Lecture Notes in Computer Science*, pages 87–98. Springer.

[57] Fletcher, P. T., Tao, R., Jeong, W. K., and Whitaker, R. T. (2007). A volumetric approach to quantifying region-to-region white matter connectivity in Diffusion Tensor MRI. In *Proceedings of the Information Processing in Medical Imaging conference (IPMI)*, volume 4584 of *Lecture Notes in Computer Science*, pages 346–358. Springer.

[58] Foong, J., Symms, M. R., Barker, G. J., Maier, M., Miller, D. H., and Ron, M. A. (2002). Investigating regional white matter in schizophrenia using diffusion tensor imaging. *Neuroreport*, 13(3):333–336.

[59] Friston, K. (1995). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2:56–78.

[60] Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. (1993). Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism*, 13:5–14.

[61] Fuster, A., van de Sande, J., Astola, L., Poupon, C., Velterop, J., and ter Haar Romeny, B. M. (2011). Fourth-order tensor invariants in high angular resolution diffusion imaging. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Workshop on Computational Diffusion MRI)*, volume 6891 of *Lecture Notes in Computer Science*, pages 54–63. Springer.

[62] Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., and Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics*, 8(8).

[63] Ghosh, A., Descoteaux, M., and Deriche, R. (2008). Riemannian framework for estimating symmetric positive definite 4th order diffusion tensors. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Pt 1)*, volume 5241 of *Lecture Notes in Computer Science*, pages 858–865. Springer.

[64] Ghosh, A., Moakher, M., and Deriche, R. (2009). Ternary quartic approach for positive 4th order diffusion tensors revisited. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 618–621.

[65] Ghosh, A., Papadopoulo, T., and Deriche, R. (2012). Biomarkers for HARDI: 2nd & 4th order tensor invariants. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 26–29.

[66] Goh, A. (2012). *Estimation and Processing of Orientation Distribution Functions for High Angular Resolution Diffusion Images*. Ph.D. Thesis, Johns Hopkins University, Ballimore, Maryland, USA.

[67] Goh, A., Lenglet, C., Thompson, P. M., and Vidal, R. (2009). A nonparametric Riemannian framework for processing high angular resolution diffusion images (HARDI). In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2496–2503.

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

[69] Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., and Beaulieu, C. (2009). Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using In Vivo Diffusion Tensor Imaging Tractography. *Cerebral Cortex*, 19(3):524–536.

[70] Gong, G., Jiang, T., Zhu, C., Zang, Y., Wang, F., Xie, S., Xiao, J., and Guo, X. (2005). Asymmetry analysis of cingulum based on scale-invariant parameterization by diffusion tensor imaging. *Human Brain Mapping*, 24(2):92–98.

[71] Grigis, A. (2012). *Approches statistiques pour la détection de changements en IRM de diffusion. Application au suivi longitudinal de pathologies neuro-dégénératives.* Ph.D. Thesis (in French), University of Strasbourg, France.

[72] Grigis, A., Noblet, V., Heitz, F., Blanc, F., de Sèze, J., Kremer, S., Rumbach, L., and Armspach, J. P. (2012). Longitudinal change detection in diffusion MRI using multivariate statistical testing on tensors. *NeuroImage*, 60(4):2206–2221.

[73] Grigis, A., Renard, F., Noblet, V., Heinrich, C., Heitz, F., and Armspach, J. P. (2011). A new high order tensor decomposition: Application to reorientation. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 258–261.

[74] Hagmann, P. (2005). *From diffusion MRI to brain connectomics.* Ph.D. Thesis, EPFL, Lausanne, Switzerland.

[75] Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., and Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, 6(7):e159.

[76] Hastie, T., Tibshirani, R., and Friedman, J. (2011). *The elements of statistical learning.* Springer, second edition.

[77] He, X. and Niyogi, P. (2003). Locality preserving projections. In *Advances in Neural Information Processing Systems 16 (NIPS)*, volume 16, pages 153–160.

[78] He, X., Yan, S., Hu, Y., Niyogi, P., and Zhang, H. J. (2005). Face recognition using Laplacianfaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(3):328–340.

[79] Hess, C. P., Mukherjee, P., Han, E. T., Xu, D., and Vigneron, D. B. (2006). Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine*, 56(1):104–117.

[80] Higham, N. J. (1986). Computing the polar decomposition with applications. *SIAM Journal on Scientific and Statistical Computing*, 7:1160–1174.

[81] Hong, S., Kwon, B., Yun, I. D., Lee, S. U., Kim, K., and Kim, J. (2013). Prediction of the potential clinical outcomes for post-resuscitated patients after cardiac arrest. In *Medical Imaging 2013: Computer-aided Diagnosis*, volume 8670 of *Proceedings of SPIE*. SPIE - The international society for optical engineering.

[82] Hong, X., Arlinghaus, L., and Anderson, A. (2009). Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging data. *Magnetic Resonance in Medicine*, 61(6):1520–1527.

[83] Horsfield, M. A. and Jones, D. K. (2002). Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases – a review. *NMR in Biomedicine*, 15(7-8):570–577.

[84] Hosey, T., Williams, G., and Ansorge, R. (2005). Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489.

[85] Hotelling, H. (1931). The generalization of student's ratio. *The Annals of Mathematical Statistics*, 2(3):360–378.

[86] Ingalhalikar, M., Parker, D., Bloy, L., Roberts, T. P., Roberts, T. P., and Verma, R. (2011). Diffusion based abnormality markers of pathology: Towards learned diagnostic prediction of ASD. *NeuroImage*, 57(3):918–927.

[87] Iturria-Medina, Y., Sotero, R. C., Canales-Rodríguez, E. J., Alemán-Gómez, Y., and Melie-García, L. (2008). Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory. *NeuroImage*, 40(3):1064–1076.

[88] Jeantroux, J., Kremer, S., Lin, X. Z., Collongues, N., Chanson, J. B., Bourre, B., Fleury, M., Blanc, F., Dietemann, J. L., and de Seze, J. (2012). Diffusion tensor imaging of normal-appearing white matter in neuromyelitis optica. *Journal of Neuroradiology*, 5(39):295–300.

[89] Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *NeuroImage*, 62(2):782–790.

[90] Jian, B. and Vemuri, B. C. (2007). A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. *Medical Imaging, IEEE Transactions on*, 26(11):1464–1471.

[91] Johansen-Berg, H. and Behrens, T. E. J. (2009). *Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy.* Elsevier Science.

[92] Jones, D. K. (2011). *Diffusion MRI: theory, methods, and applications.* Oxford University Press.

[93] Jones, D. K., Catani, M., Pierpaoli, C., Reeves, S. J. C., Shergill, S. S., O'Sullivan, M., Golesworthy, P., McGuire, P., Horsfield, M. A., Simmons, A., Williams, S. C. R., and Howard, R. J. (2006). Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. *Human Brain Mapping*, 27(3):230–238.

[94] Jones, E., Oliphant, T., Peterson, P., et al. (2001–). SciPy: Open source scientific tools for Python. http://www.scipy.org/.

[95] Kacar, K., Rocca, M. A., Copetti, M., Sala, S., Mesaros, S., Opincal, T. S., Caputo, D., Absinta, M., Drulovic, J., Kostic, V. S., Comi, G., and Filippi, M. (2011). Overcoming the Clinical-MR Imaging Paradox of Multiple Sclerosis: MR Imaging Data Assessed with a Random Forest Approach. *American Journal of Neuroradiology (AJNR)*, 32(11):2098–2102.

[96] Kruskal, J. B. and Wish, M. (1978). *Multidimensional Scaling*. Number 11 in 07. SAGE Publications.

[97] Lazar, M. (2010). Mapping brain anatomical connectivity using white matter tractography. *NMR in Biomedicine*, 23(7):821–835.

[98] Lazar, M. and Alexander, A. L. (2005). Bootstrap white matter tractography (BOOT-TRAC). *NeuroImage*, 24(2):524–32.

[99] Lenglet, C., Rousson, M., and Deriche, R. (2006). DTI segmentation by statistical surface evolution. *Medical Imaging, IEEE Transactions on*, 25(6):685–700.

[100] Lin, F., Yu, C., Jiang, T., Li, K., Li, X., Qin, W., Sun, H., and Chan, P. (2006). Quantitative analysis along the pyramidal tract by length-normalized parameterization based on diffusion tensor tractography: application to patients with relapsing neuromyelitis optica. *NeuroImage*, 33(1):154–160.

[101] Lutsep, H. L., Albers, G. W., DeCrespigny, A., Kamat, G. N., Marks, M. P., and Moseley, M. E. (1997). Clinical utility of diffusion-weighted magnetic resonance imaging in the assessment of ischemic stroke. *Annals of Neurology*, 41(5):574–580.

[102] Malcolm, J. G., Shenton, M. E., and Rathi, Y. (2009). Filtered tractography: Validation on a physical phantom. In *Fiber Cup (in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI))*, pages 220–223.

[103] Mardia, K. V. and Jupp, P. E. (2000). *Directional statistics*. West Sussex, England: Wiley.

[104] Mazziotta, J., Toga, A., and Frackowiak, R. (2000). *Brain Mapping: The Disorders*. Academic Press.

[105] McIlwain, H. and Bachelard, H. S. (1971). *Biochemistry and the central nervous system*. Churchill Livingstone.

[106] Mesaros, S., Rocca, M. A., Kacar, K., Kostic, J., Copetti, M., Stosic-Opincal, T., Preziosa, P., Sala, S., Riccitelli, G., Horsfield, M. A., Drulovic, J., Comi, G., and Filippi, M. (2012). Diffusion tensor MRI tractography and cognitive impairment in multiple sclerosis. *Neurology*, 78(13):969–975.

[107] Mitchell, J. R., Karlik, S. J., Lee, D. H., Eliasziw, M., Rice, G. P., and Fenster, A. (1996). The variability of manual and computer assisted quantification of multiple sclerosis lesion volumes. *Medical Physics*, 23(1):85–97.

[108] Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvado, O., Connelly, A., Campbell, B., Palmer, S., Sharma, G., Christensen, S., and Carey, L. (2014). Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage*, 98:324–335.

[109] Moakher, M. (2005). A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal of Matrix Analysis & Applications*, 26(3):735–747.

[110] Moakher, M. (2008). Fourth-order Cartesian tensors: old and new facts, notions and applications. *The Quarterly Journal of Mechanics and Applied Mathematics*, 61(2):181–203.

[111] Moakher, M. (2009). The algebra of fourth-order tensors with application to diffusion MRI. In *Visualization and Processing of Tensor Fields*, Mathematics and Visualization, pages 57–80. Springer.

[112] Mori, S. (2007). *Introduction to Diffusion Tensor Imaging*. Elsevier Science.

[113] Mori, S., Crain, B. J., Chacko, V. P., and van Zijl, P. C. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, 45(2):265–269.

[114] Mori, S., Oishi, K., and Faria, A. V. (2009). White matter atlases based on diffusion tensor imaging. *Current Opinion in Neurology*, 22(4):362–369.

[115] Mori, S., Wakana, S., van Zijl, P. C. M., and Nagae-Poetscher, L. M. (2005). *MRI Atlas of Human White Matter*. Elsevier Science.

[116] Naylor, M. G., Cardenas, V. A., Tosun, D., Schuff, N., Weiner, M., and Schwartzman, A. (2014). Voxelwise multivariate analysis of multimodality magnetic resonance imaging. *Human Brain Mapping*, 35(3):831–846.

[117] Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15(1):1–25.

[118] Oja, H. and Randles, R. H. (2004). Multivariate nonparametric tests. *Statistical Science*, 19(4):598–605.

[119] Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science and Engineering*, 9(3):10–20.

[120] Osborne, D., Patrangenaru, V., Ellingson, L., Groisser, D., and Schwartzman, A. (2013). Nonparametric two-sample tests on homogeneous Riemannian manifolds, Cholesky decompositions and Diffusion Tensor Image analysis. *Journal of Multivariate Analysis*, 119(C):163–175.

[121] Ozarslan, E. and Mareci, T. H. (2003). Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50(5):955–965.

[122] Ozarslan, E., Vemuri, B. C., and Mareci, T. H. (2005). Generalized scalar measures for diffusion MRI using trace, variance, and entropy. *Magnetic Resonance in Medicine*, 53(4):866–876.

[123] Pagani, E., Filippi, M., Rocca, M. A., and Horsfield, M. A. (2005). A method for obtaining tract-specific diffusion tensor MRI measurements in the presence of disease: application to patients with clinically isolated syndromes suggestive of multiple sclerosis. *NeuroImage*, 26(1):258–265.

[124] Paldino, M. J., Hedges, K., and Zhang, W. (2014). Independent contribution of individual white matter pathways to language function in pediatric epilepsy patients. *NeuroImage Clinical*, 6:327–332.

[125] Park, H. J., Kubicki, M., Shenton, M. E., Guimond, A., McCarley, R. W., Maier, S. E., Kikinis, R., Jolesz, F. A., and Westin, C. F. (2003). Spatial normalization of diffusion tensor MRI using multiple channels. *NeuroImage*, 20(4):1995–2009.

[126] Parker, G. J. and Alexander, D. C. (2003). Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *Proceedings of the Information Processing in Medical Imaging conference (IPMI)*, volume 2732 of *Lecture Notes in Computer Science*, pages 684–695. Springer.

[127] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[128] Pennec, X. (2008). Statistical computing on manifolds: From Riemannian geometry to computational anatomy. In *Emerging Trends in Visual Computing*, pages 347–386.

[129] Pennec, X., Fillard, P., and Ayache, N. (2006). A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66.

[130] Pontabry, J., Rousseau, F., Oubel, E., Studholme, C., Koob, M., and Dietemann, J. L. (2013). Probabilistic tractography using Q-ball imaging and particle filtering: Application to adult and in-utero fetal brain studies. *Medical Image Analysis*, 17(3):297–310.

[131] Preziosa, P., Rocca, M. A., Mesaros, S., Pagani, E., Drulovic, J., Stosic-Opincal, T., Dackovic, J., Copetti, M., Caputo, D., and Filippi, M. (2014). Relationship between Damage to the Cerebellar Peduncles and Clinical Disability in Multiple Sclerosis. *Radiology*, 271(3):822–830.

[132] Qian, C., Masad, I. S., Rosenberg, J. T., Elumalai, M., Brey, W. W., Grant, S. C., and Gor'kov, P. L. (2012). A volume birdcage coil with an adjustable sliding tuner ring for neuroimaging in high field vertical magnets: Ex and in vivo applications at 21.1T. *Journal of Magnetic Resonance*, 221C:110–116.

[133] Reiss, P. T., Stevens, M. H. H., Shehzad, Z., Petkova, E., and Milham, M. P. (2010). On distance-based permutation tests for between-group comparisons. *Biometrics*, 66(2):636–643.

[134] Renard, F. (2011). *Création et utilisation d'atlas en IRM de diffusion. Application à l'étude des troubles de la conscience.* Ph.D. Thesis (in French), University of Strasbourg, France.

[135] Renard, F., Noblet, V., Heinrich, C., and Kremer, S. (2010). Reorientation strategies for high order tensors. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 1185–1188.

[136] Rovaris, M., Rocca, M. A., Sormani, M. P., Comi, G., and Filippi, M. (1999). Reproducibility of brain MRI lesion volume measurements in multiple sclerosis using a local thresholding technique: Effects of formal operator training. *Journal of European Neurology*, 41:226–230.

[137] Schultz, T., Fuster, A., Ghosh, A., Deriche, R., Florack, L., and Lim, L. H. (2014). Higher-order tensors in diffusion imaging. In Westin, C. F., Vilanova, A., and Burgeth, B., editors, *Visualization and Processing of Tensors and Higher Order Descriptors for Multi-Valued Data.* Springer.

[138] Schwartzman, A. (2006). *Random ellipsoids and false discovery rates: statistics for diffusion tensor imaging data.* PhD thesis, Stanford University.

[139] Schwartzman, A., Dougherty, R. F., and Taylor, J. E. (2005). Cross-subject comparison of principal diffusion direction maps. *Magnetic Resonance in Medicine*, 53(6):1423–1431.

[140] Schwartzman, A., Dougherty, R. F., and Taylor, J. E. (2008). False discovery rate analysis of brain diffusion direction maps. *Annals of Applied Statistics*, 2(1):153–175.

[141] Schwartzman, A., Dougherty, R. F., and Taylor, J. E. (2010). Group comparison of eigenvalues and eigenvectors of diffusion tensors. *Journal of the American Statistical Association*, 105(490):588–599.

[142] Scott, D. (1992). *Multivariate density estimation: theory, practice, and visualization.* Wiley.

[143] Sfikas, G., Constantinopoulos, C., Likas, A., and Galatsanos, N. (2005). An analytic distance metric for Gaussian mixture models with application in image retrieval. In *15th International Conference on Artificial Neural Networks (IEEE ICANN 2005)*, volume 3697 of *Lecture Notes in Computer Science*, pages 835–840. Springer.

[144] Smith, S., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T., Mackay, C., Watkins, K., Ciccarelli, O., Cader, M. Z., Matthews, P., and Behrens, T. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31(4):1487–1505.

[145] Sonka, M. and Fitzpatrick, J. M. (2000). *Handbook of medical imaging(Volume 2, Medical image processing and analysis).* SPIE - The international society for optical engineering.

[146] Sorensen, A. G., Buonanno, F. S., Gonzalez, R. G., Schwamm, L. H., Lev, M. H., Huang-Hellinger, F. R., Reese, T. G., Weisskoff, R. M., Davis, T. L., Suwanwela, N., Can, U., Moreira, J. A., Copen, W. A., Look, R. B., Finklestein, S. P., Rosen, B. R., and Koroshetz, W. J. (1996). Hyperacute stroke: evaluation with combined multisection diffusion-weighted and hemodynamically weighted echo-planar MR imaging. *Radiology*, 199(2):391–401.

[147] Sporns, O., Honey, C. J., and Kötter, R. (2007). Identification and classification of hubs in brain networks. *PLoS ONE*, 2(10).

[148] Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 1(1).

[149] Stejskal, E. O. and Tanner, J. E. (1965). Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42(1):288–292.

[150] Stough, J. V., Glaister, J., Ye, C., Ying, S. H., Prince, J. L., and Carass, A. (2014). Automatic Method for Thalamus Parcellation Using Multi-modal Feature Classification. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Pt 3)*, volume 8675 of *Lecture Notes in Computer Science*, pages 169–176. Springer.

[151] Surajit, R. (2003). *Distance-Based Model Selection with Application to the Analysis of Gene Expression Data.* Ms Thesis, Department of Statistics, Pennsylvania State University.

[152] Tao, X. and Miller, J. (2006). A method for registering diffusion weighted magnetic resonance images. In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 4191 of *Lecture Notes in Computer Science*, pages 594–602. Springer.

[153] Tarantola, A. (2006). *Elements for Physics: Quantities, Qualities, and Intrinsic Theories.* Springer.

[154] Tenenbaum, J. B., Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323.

[155] Toga, A. and Mazziotta, J. (1996). *Brain Mapping: The Methods.* Academic Press.

[156] Toga, A. and Mazziotta, J. (2000). *Brain Mapping: The Systems.* Academic Press.

[157] Tournier, J. D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472.

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

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

[160] Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., and Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582.

[161] Tuch, D. S., Reese, T. G., Wiegell, M. R., and Wedeen, V. J. (2003). Diffusion MRI of complex neural architecture. *Neuron*, 40(5):885–95.

[162] Vaughan, T., DelaBarre, L., Snyder, C., Tian, J., Akgun, C., Shrivastava, D., Liu, W., Olson, C., Adriany, G., Strupp, J., Andersen, P., Gopinath, A., van de Moortele, P. F., Garwood, M., and Ugurbil, K. (2006). 9.4T human MRI: Preliminary results. *Magnetic Resonance in Medicine*, 56(6):1274–1282.

[163] Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., and et al. (2001). The sequence of the human genome. *Science*, 291(5507):1304–1351.

[164] Verma, R. and Davatzikos, C. (2006). Manifold based analysis of diffusion tensor images using isomaps. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 790–793.

[165] Verma, R., Khurd, P., and Davatzikos, C. (2007). On analyzing diffusion tensor images by identifying manifold structure using isomaps. *Medical Imaging, IEEE Transactions on*, 26(6):772–778.

[166] Vogelstein, J. (2013). Advancing information superiority through applied neuroscience. Technical Report 4, Johns Hopkins APL Technical Digest.

[167] Warach, S., Gaa, J., Siewert, B., Wielopolski, P., and Edelman, R. R. (1995). Acute human stroke studied by whole brain echo planar diffusion-weighted magnetic resonance imaging. *Annals of Neurology*, 37(2):231–241.

[168] Watanabe, M., Aoki, S., Masutani, Y., Abe, O., Hayashi, N., Masumoto, T., Mori, H., Kabasawa, H., and Ohtomo, K. (2006). Flexible ex vivo phantoms for validation of diffusion tensor tractography on a clinical scanner. *Radiation Medicine*, 24(9):605–609.

[169] Watanabe, M., Masutani, Y., Aoki, S., Mori, H., Abe, O., Hayashi, N., Masumoto, T., Yamada, H., Yoshikawa, T., Kabasawa, H., and Ohtomo, K. (2005). In-vitro phantoms for estimating diffusion tensor tractography. *Nihon Igaku Hoshasen Gakkai Zasshi*, 65(4):449–51.

[170] Wedeen, V. J., Hagmann, P., Tseng, W. Y. I., Reese, T. G., and Weisskoff, R. M. (2005). Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386.

[171] Weinberger, K. and Saul, L. (2006). Unsupervised learning of image manifolds by semidefinite programming. *International Journal of Computer Vision*, 70(1):77–90.

[172] Weldeselassie, Y., Barmpoutis, A., and Atkins, S. (2012). Symmetric positive-definite Cartesian tensor fiber orientation distributions (CT-FOD). *Medical Image Analysis*, 16(6):1121–1129.

[173] Weldeselassie, Y. T., Barmpoutis, A., and Atkins, S. (2010). Symmetric positive-definite Cartesian tensor orientation distribution functions (CT-ODF). In *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 6361 of *Lecture Notes in Computer Science*, pages 582–589. Springer.

[174] Whitcher, B., Wisco, J. J., Hadjikhani, N., and Tuch, D. S. (2007). Statistical group comparison of diffusion tensors via multivariate hypothesis testing. *Magnetic Resonance in Medicine*, 57(6):1065–1074.

[175] Wycoco, V., Shroff, M., Sudhakar, S., and Lee, W. (2013). White matter anatomy: what the radiologist needs to know. *Neuroimaging Clinics of North America*, 23(2):197–216.

[176] Yassine, I. and McGraw, T. (2009). 4th order diffusion tensor interpolation with divergence and curl constrained Bézier patches. In *Proceedings of ISBI: IEEE International Symposium on Biomedical Imaging*, pages 634–637.

[177] Zar, J. H. (2007). *Biostatistical Analysis (5th Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[178] Zhu, H., Chen, Y., Ibrahim, J. G., Li, Y., Hall, C., and Lin, W. (2009). Intrinsic regression models for positive-definite matrices with applications to diffusion tensor imaging. *Journal of the American Statistical Association*, 104(487):1203–1212.

Author's Publications

Communications in International Conferences

[1] Gkamas, Th., Renard, R., Heinrich, Ch., and Kremer, S. (2015). A Fourth Order Tensor Statistical Model for Diffusion Weighted MRI - Application to Population Comparison. *In Proceedings of the 4th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, pp. 277-282, DOI: 10.5220/0005252602770282, Lisbon, Portugal, 10-12 January 2015.

OTHER PUBLICATIONS RELATED TO PREVIOUS STUDIES

Articles in International Journals

[2] Chantas, G., Gkamas, Th., and Nikou, C. (2014). Variational-Bayes optical flow. *Journal of Mathematical Imaging and Vision (JMIV)*, 50(3): 199-213.

Communications in International Conferences

[3] Gkamas, Th., Chantas, G., and Nikou, C. (2012). A probabilistic formulation of the optical flow problem. *In Proceedings of the 21st International Conference on Pattern Recognition (ICPR)*, pp. 754-757, Tsukuba, Japan, 11-15 November 2012.

[4] Gkamas, Th., and Nikou, C. (2011). Guiding optical flow estimation using superpixels. *17th International Conference on Digital Signal Processing (DSP)*, DOI: 10.1109/ICDSP. 2011.6004871, Corfu, Greece, 6-8 July 2011.

Theodosios GKAMAS
Modélisation statistique de tenseurs d'ordre supérieur en imagerie par résonance magnétique de diffusion

Résumé

L'IRMd est un moyen non invasif permettant d'étudier *in vivo* la structure des fibres nerveuses du cerveau. Dans cette thèse, nous modélisons des données IRMd à l'aide de tenseurs d'ordre 4 (T4). Les problèmes de comparaison de groupes ou d'individu avec un groupe normal sont abordés, et résolus à l'aide d'analyses statistiques sur les T4s. Les approches utilisent des réductions non linéaires de dimension, et bénéficient des métriques non euclidiennes pour les T4s. Les statistiques sont calculées dans l'espace réduit, et permettent de quantifier la dissimilarité entre le groupe (ou l'individu) d'intérêt et le groupe de référence. Les approches proposées sont appliquées à la neuromyélite optique et aux patients atteints de *locked in syndrome*. Les conclusions tirées sont cohérentes avec les connaissances médicales actuelles.

Mots-clés : IRMd, tenseur d'ordre supérieur, métrique non-euclidienne, réduction de dimension non linéaire, comparaison de groupe ou d'individu *vs* groupe normal, analyse statistique, test de permutation, maladie NMO, LIS syndrome.

Résumé en anglais

DW-MRI is a non-invasive way to study *in vivo* the structure of nerve fibers in the brain. In this thesis, fourth order tensors (T4) were used to model DW-MRI data. In addition, the problems of group comparison or individual against a normal group were discussed and solved using statistical analysis on T4s. The approaches use nonlinear dimensional reductions, assisted by non-Euclidean metrics for T4s. The statistics are calculated in the reduced space and allow us to quantify the dissimilarity between the group (or the individual) of interest and the reference group. The proposed approaches are applied to neuromyelitis optica and patients with locked in syndrome. The derived conclusions are consistent with the current medical knowledge.

Keywords: DW-MRI, high order tensor, non-Euclidean metric, nonlinear dimension reduction, group or individual *vs* normal group comparison, statistical analysis, permutation testing, NMO disease, LIS syndrome.