



Approches statistiques pour la détection de changements en IRM de diffusion : application au suivi longitudinal de pathologies neuro-dégénératives

Antoine Grigis

► To cite this version:

Antoine Grigis. Approches statistiques pour la détection de changements en IRM de diffusion : application au suivi longitudinal de pathologies neuro-dégénératives. Traitement du signal et de l'image [eess.SP]. Université de Strasbourg, 2012. Français. NNT : 2012STRAD018 . tel-00750933

HAL Id: tel-00750933

<https://theses.hal.science/tel-00750933>

Submitted on 12 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE : *Mathématiques, Sciences de l'Information et de l'Ingénieur*
Laboratoire des Sciences de l'Image, de l'Information et de la Télédétection,
UMR 7005, Université de Strasbourg-CNRS
Laboratoire d'Imagerie et de Neurosciences Cognitives,
UMR 7237, Université de Strasbourg-CNRS

THÈSE

présentée par :

Antoine GRIGIS

soutenue le : 25 septembre 2012

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline / Spécialité : Traitement du Signal et des Images

**Approches statistiques pour la détection de
changements en IRM de diffusion.
Application au suivi longitudinal de
pathologies neuro-dégénératives.**

THÈSE dirigée par :

M. HEITZ Fabrice
M. ARMSPACH Jean-Paul

LSIIT, Strasbourg
LINC-IPB, Strasbourg

RAPPORTEURS :

M. MANGIN Jean-François
Mme RUAN Su

LNAO-Neurospin-CEA, Saclay
LITIS, Rouen

AUTRES MEMBRES DU JURY :

M. DOUSSET Vincent
M. NOBLET Vincent

CHU Pellegrin, Bordeaux
LSIIT, Strasbourg

Remerciements

Je tiens dans un premier temps à remercier mes directeurs de thèse, Fabrice Heitz et Jean-Paul Armspach, pour m'avoir confié ce travail de recherche, ainsi que pour leur aide et leurs précieux conseils au cours de ces années. Je remercie également mon encadrant, Vincent Noblet, pour ses conseils et son aide. Un grand merci à Frédéric Blanc pour sa disponibilité et son aide précieuse lors de l'interprétation des résultats.

Je remercie Jean-François Mangin et Su Ruan d'avoir accepté d'être les rapporteurs de ma thèse ainsi que Vincent Dousset pour avoir présidé mon jury de thèse.

Enfin, je n'oublie pas non plus de remercier mes collègues, mes amis et ma famille qui m'ont encouragé et accompagné tout au long de ma thèse.

Table des matières

Remerciements	3
Introduction	11
I Mesure et modélisation de la diffusion des molécules d'eau	15
1 L'acquisition des images	19
1.1 La diffusion moléculaire	19
1.1.1 L'équation de diffusion	19
1.1.2 La probabilité de déplacement des molécules d'eau : le propagateur	21
1.2 L'Imagerie par Résonance Magnétique nucléaire (IRM)	22
1.2.1 L'acquisition des images : le signal de diffusion	22
1.2.2 Représentation : le modèle de diffusion	24
2 Le modèle du tenseur de diffusion sur une base de polynômes homogènes	27
2.1 L'Imagerie du Tenseur de Diffusion (ITD)	28
2.1.1 Formulation générale du modèle	28
2.1.2 Cas particulier : le tenseur d'ordre 2	29
2.1.3 Cas particulier : le tenseur d'ordre 4	35
2.2 L'estimation des paramètres du modèle en ITD	39
2.2.1 La méthode des moindres carrés : modélisation de la diffusivité	39
2.2.2 La méthode des quartiques : modélisation de la diffusivité	40
2.2.3 Modélisation de la probabilité de déplacement : la fonction de distribution des orientations	43
2.2.4 Comparaison des différentes méthodes d'estimation de tenseurs	46
3 Le modèle de diffusion sur une base d'harmoniques sphériques	49
3.1 Décomposition sur une base d'harmoniques sphériques	50
3.1.1 Les Harmoniques Sphériques	50
3.1.2 Formulation générale du modèle	51
3.1.3 La fonction de distribution des orientations	52

3.2	Lien entre les bases de polynômes homogènes et des harmoniques sphériques	56
Bibliographie		61
II Détection de changements en IRM de diffusion		65
4	Statistiques inférentielles	69
4.1	Les tests d'hypothèses	69
4.2	Les deux types d'erreurs possibles	70
4.3	Signification statistique : accepter ou rejeter l'hypothèse nulle	72
4.3.1	Tests d'hypothèses simples	72
4.3.2	Problème des comparaisons multiples	73
5	La détection de changements en imagerie médicale	79
5.1	Etat de l'art sur la détection de changements en imagerie médicale	80
5.1.1	Les problématiques de recalage dans la détection de changements	80
5.1.2	La détection de changements basée sur un recalage des images	82
5.2	Etat de l'art sur la détection de changements en imagerie de diffusion	83
5.2.1	Détection de changements sur des indices scalaires	84
5.2.2	Détection de changements multivariée sur des vecteurs de paramètres	87
5.2.3	Détection de changements multivariée sur la structure matricielle des tenseurs de diffusion	88
6	Tests statistiques adaptés aux tenseurs de diffusion	91
6.1	Loi multinormale pour les matrices symétriques définies positives	92
6.1.1	Distribution multinormale vectorisée	92
6.1.2	Distribution multinormale matricielle pour les matrices symétriques	93
6.1.3	Distribution multinormale matricielle pour les matrices symétriques définies positives	94
6.2	Tests statistiques multivariés sur des populations de tenseurs	96
6.2.1	Le test du rapport de vraisemblance à un échantillon (RV1)	96
6.2.2	Le test du rapport de vraisemblance à deux échantillons (RV2)	101
6.2.3	La distribution empirique sous l'hypothèse nulle	108
Bibliographie		111
III Suivi longitudinal des tenseurs d'ordre 2 : prétraitements et outils statistiques		117
7	Les prétraitements des images	121
7.1	Les artefacts en imagerie de diffusion	121
7.1.1	Les courants de Foucault	122

7.1.2	L'effet de susceptibilité magnétique	122
7.2	Le recalage des images	122
7.3	L'interpolation et la réorientation des images de tenseurs	123
8	Extraction de populations de tenseurs	127
8.1	Prise en compte du voisinage spatial à l'échelle du voxel	127
8.2	Méthode de rééchantillonnage : le bootstrap	128
8.2.1	Le Bootstrap Local (BL)	129
8.2.2	Le Bootstrap Spatial (BS)	132
8.3	Prise en compte d'un voisinage basé sur une modélisation des faisceaux de fibres de la substance blanche	133
8.3.1	Modélisation des fibres de la substance blanche : la tractographie	133
8.3.2	Regroupement des fibres en faisceaux	135
8.3.3	Estimation des paramètres dans une section de faisceau	139
9	Les deux chaînes de traitement proposées pour le suivi longitudinal	145
9.1	Détection de changement de diffusion à l'échelle du voxel	146
9.2	Détection de changements de diffusion sur une modélisation des faisceaux de fibres	147
9.2.1	Statistique locale	147
9.2.2	Statistique globale	148
	Bibliographie	151
IV	Simulations et études de cas cliniques	155
10	Les maladies neurodégénératives étudiées	159
10.1	La sclérose en plaques (SEP)	159
10.1.1	Description de la pathologie	159
10.1.2	Description de la base de patients utilisée	161
10.2	La neuromyéélite optique (NMO)	162
10.2.1	Description de la pathologie	162
10.2.2	Description de la base de patients utilisée	163
11	Evaluation des méthodes de détection de changement de diffusion à l'échelle du voxel	165
11.1	Simulations	166
11.1.1	Distributions sous l'hypothèse nulle	166
11.1.2	Résultats sur données simulées	167
11.1.3	Evaluation du taux de recouvrement	171
11.1.4	Influence de l'anisotropie des images	171
11.2	Application à des données cliniques	172
11.2.1	Etude détaillée de deux patients	172

11.2.2	Base de données NMO	176
12	Evaluation des méthodes de détection de changements de diffusion sur une modélisation des faisceaux de fibres	183
12.1	Résultats sur données simulées	184
12.2	Application à des données cliniques	190
12.2.1	Etude détaillée de deux patients	191
12.2.2	Base de données NMO	192

Bibliographie	197
----------------------	------------

V	Vers un suivi longitudinal sur des modèles d'ordre supérieur	199
----------	---	------------

13	Les prétraitements des images : la réorientation des tenseurs d'ordre 4	203
13.1	Etat de l'art des méthodes de réorientation	204
13.1.1	Réorientation de l'ODF discrétisée	204
13.1.2	Réorientation par décomposition sur une base de diracs	204
13.1.3	Réorientation rigide	205
13.1.4	Réorientation par décomposition spectrale ou par décomposition de Hilbert	205
13.2	Méthode proposée pour la réorientation des tenseurs d'ordre 4	205
13.2.1	Calcul des maxima de l'ODF	206
13.2.2	Décomposition des tenseurs d'ordre supérieur	208
13.2.3	Application à la réorientation	209
13.3	Résultats expérimentaux	210
13.3.1	Utilisation de l'approche proposée	210
13.3.2	Comparaison avec des méthodes de la littérature	212
14	Un test statistique multivarié pour des tenseurs d'ordre 4	215
14.1	Chaîne de traitement	215
14.2	Résultats	217
14.2.1	Simulations	217
14.2.2	Application à des données cliniques	218
15	Détection de changements d'orientations des fibres par une mesure de similarité	223
15.1	Mesure de similarité pour des tenseurs d'ordre 2	224
15.1.1	La matrice de dispersion	224
15.1.2	La similarité entre deux populations	224
15.2	Mesure de similarité pour des tenseurs d'ordre 4	224
15.2.1	Une association intra-population	225
15.2.2	Une association inter-population	225
15.3	Chaîne de traitement	225

15.4 Résultats	226
15.4.1 Simulations	226
15.4.2 Application à des données cliniques	227
Bibliographie	235
Conclusion générale et perspectives	237
Bibliographie	241
Annexes	243

Introduction

Contexte de la thèse

L'IRM de diffusion (IRMd) est une modalité récente d'imagerie médicale qui permet d'apporter des informations nouvelles sur les micro-structures locales des tissus, par rapport aux modalités d'IRM anatomiques plus classiques. Utilisée en particulier en neuroimagerie, cette modalité permet de caractériser *in vivo* l'organisation neuronale et apporte ainsi une information unique sur l'organisation des fibres de la substance blanche. Les IRM de diffusion sont des images multivariées (vectorielles, tensorielles) et leur visualisation et interprétation par un expert humain est donc une tâche délicate et souvent laborieuse. Des outils de traitements sont ainsi nécessaires pour assister le médecin dans l'analyse et l'interprétation de ces données. De surcroît, ces outils permettent une amélioration de la prise en charge des patients en facilitant notamment leur suivi au cours du temps. Dans ce contexte, l'analyse automatique de changements longitudinaux intra-patient a fait l'objet de nombreux travaux de recherche. Notre recherche a ainsi porté sur le développement de nouveaux modèles et tests statistiques permettant la détection automatique de changements sur des séquences temporelles d'images IRMd. Il s'agit d'un problème difficile en raison du caractère très bruité des données disponibles, de leur nature multivariée et de la structure des espaces associés (tenseurs définis positifs, ...). Ces nouvelles méthodes d'analyse ont été évaluées sur données simulées et appliquées, dans le contexte de recherches cliniques, au suivi longitudinal de patients atteints de pathologies neurodégénératives telles que la sclérose en plaques ou la neuromyéélite optique de Devic.

Ce travail de thèse a été mené au sein de deux équipes de recherche : l'équipe « Modèles, Images et Vision »(MIV) du Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT, UMR 7005, Université de Strasbourg-CNRS) et l'équipe « Imagerie In Vivo »(IIV) du Laboratoire d'Imagerie et de Neurosciences Cognitives (LINC, UMR 7237, Université de Strasbourg-CNRS). Cette recherche fait suite à une précédente thèse menée dans les deux équipes sur la même problématique. Les contributions par rapport à ce précédent travail sont une meilleure prise en compte de la nature tensorielle des modèles d'ordre 2 en IRMd (tests statistiques sur des matrices définies positives), une extension vers des modèles d'ordre supérieur et une gestion plus fine des voisinages sur lesquels les tests sont menés (avec en particulier la conception de tests statistiques sur des faisceaux de fibres).

Problématique

L'objet de cette thèse est de développer de nouvelles approches statistiques pour la détection de changements dans des séquences temporelles d'images d'IRMd. Même si l'IRMd est une modalité très prometteuse pour étudier l'architecture des fibres de la substance blanche, elle reste d'exploitation limitée, en particulier en routine clinique. De nombreux problèmes inhérents à cette modalité doivent ainsi être résolus en amont, pour permettre la mise en place en routine clinique de chaînes de traitements adaptées pour la détection longitudinale de changements.

L'IRMd consiste à mesurer, dans chaque direction de l'espace, le mouvement (la diffusion) des molécules d'eau. Cette diffusion (qui peut être anisotrope) est alors caractérisée par une matrice de covariance $\mathbf{D}$ du déplacement des particules (appelée également tenseur de diffusion). Cette matrice de covariance est symétrique et définie positive. L'espace des matrices symétriques, définies positives n'est malheureusement pas un espace vectoriel. Plus précisément, cet espace est un *cône* (c'est-à-dire une partie de l'espace vectoriel des matrices, qui est stable uniquement pour la multiplication par des réels positifs). La matrice nulle n'appartient pas à cet espace. Une perturbation, même faible, introduite lors de l'estimation des éléments de la matrice, peut faire sortir de cet espace. Ceci rend délicat la définition d'estimateurs et de tests statistiques sur les matrices définies positives. Différentes approches peuvent être envisagées : calculs sous contrainte de positivité, changement d'espaces matriciels (vers par exemple des espaces log-euclidiens), voire géométrie différentielle sur des variétés riemanniennes. Avant de calculer les statistiques de détection de changements, une transformation vers un espace log-euclidien, qui autorise les opérations linéaires sur les tenseurs de diffusion, tout en assurant que le résultat reste un tenseur défini positif, a été mis en œuvre dans notre cas.

Un autre problème important est le choix du modèle utilisé pour représenter les données en IRMd. Le classique modèle tensoriel d'ordre 2 n'est en effet qu'un modèle physique approché qui ne permet par exemple pas une prise en compte locale des croisements de fibres. Or, les résolutions d'acquisition actuelles permettent d'étudier des micro-structures (*i.e.*, de 1 à 30 μm pour le diamètre des fibres myélinisées) à des échelles macroscopiques. Ainsi, dans certaines régions, le signal mesuré provient des contributions de plusieurs faisceaux. Il faut alors augmenter le degré de complexité du modèle afin de modéliser au mieux les configurations de croisements de fibres attendues.

Cette thèse s'inscrit également dans une problématique de recherche clinique, visant à apporter de nouvelles informations aux praticiens dans le suivi longitudinal de diverses pathologies neurodégénératives telles que la sclérose en plaques ou la neuromyéélite optique. L'objectif est d'une part la détection précoce de l'apparition des lésions ou de leur évolution, d'autre part la mise en évidence de modifications pathologiques dans la matière blanche d'apparence normale en IRM conventionnelle. L'enjeu de ces travaux est d'apporter aux neurologues, neurochirurgiens, neuro-oncologues et radiothérapeutes de précieux outils diagnostiques et prédictifs. Cet enjeu est d'importance, dans des sociétés vieillissantes, dans lesquelles la

prégnance des maladies neurodégénératives augmente régulièrement.

Contributions

En chaque point d'une acquisition d'IRMd, la distribution des directions de diffusion des molécules d'eau est modélisée par un tenseur de diffusion. Pour détecter des changements, de nouvelles méthodes de détection statistique adaptées aux images du tenseur de diffusion ont été mises en place. Ces méthodes permettent notamment de respecter la structure des espaces associés aux tenseurs de diffusion. Une extension de ces méthodes statistiques est proposée. Elle permet de prendre en compte un modèle d'ordre supérieur qui représente plus fidèlement la diffusion des molécules d'eau dans le cerveau humain. L'interprétation d'un modèle d'ordre supérieur reste toutefois complexe. Une mesure de similarité, directement construite sur les directions des fibres, est alors proposée pour ces modélisations d'ordre supérieur.

Ces méthodes comparent toutes deux populations de tenseurs. Dans le contexte d'un suivi longitudinal, la constitution des populations de tenseurs peut exploiter immédiatement un voisinage spatial de taille fixée par l'utilisateur. Une gestion plus fine des voisinages sur lesquels les tests sont menés est proposée, avec en particulier la génération de populations de tenseurs le long des faisceaux de fibres de la substance blanche ou la prise en compte de la variabilité du signal de diffusion.

D'autres contributions méthodologiques sont suggérées, notamment au niveau des modèles d'ordre supérieur. La complexité de ces modèles nécessite la mise en place de stratégies spécifiques pour l'extraction des directions des fibres. De plus, un modèle d'ordre supérieur est représenté, en chaque voxel de l'image, par une forme géométrique pouvant avoir plusieurs maxima. Lors de l'alignement de deux images à comparer, il faut tenir compte du fait que cette modélisation représente une information directionnelle. Une stratégie de réorientation des modèles d'ordre supérieur, tenant compte des différentes orientations des fibres, est alors présentée.

D'un point de vue applicatif, après une première phase de validation sur données simulées, les méthodes ont été appliquées sur des cas cliniques. L'étude d'une base de données constituée d'exams longitudinaux de 12 patients atteints de sclérose en plaques a montré la capacité de la méthode à détecter des modifications au sein de lésions n'évoluant pas en IRM conventionnelle. L'étude d'une base de données constituée d'exams longitudinaux de 11 patients atteints de neuromyéélite optique a montré qu'il se produisait des changements au sein de la substance blanche d'apparence normale qui pouvaient être corrélés avec l'évolution des capacités cognitives ou motrices des patients. Ces méthodes de détection de changements en IRMd pourraient ainsi offrir de nouveaux biomarqueurs pour le suivi des pathologies neurodégénératives.

Organisation du mémoire de thèse

Le mémoire de thèse est organisé en cinq parties :

- **Partie I, Mesure et modélisation de la diffusion des molécules d'eau :** cette partie se décompose en trois sous chapitres. Le chapitre 1 introduit l'IRM de diffusion. Le chapitre 2 présente comment modéliser le signal de diffusion en utilisant le modèle du tenseur de diffusion. Ce modèle permet de représenter soit le signal de diffusion, soit la diffusivité, soit la probabilité de déplacement des molécules d'eau. Le chapitre 3 propose d'exprimer un tenseur de diffusion sur une base d'harmoniques sphériques afin de faciliter la modélisation de la probabilité de déplacement des molécules d'eau.
- **Partie II, Détection de changements en IRM de diffusion :** cette partie se décompose en trois sous chapitres. Le chapitre 4 introduit les concepts relatifs aux statistiques inférentielles. Le chapitre 5 expose un état de l'art sur la détection de changements en imagerie médicale, en s'intéressant particulièrement aux méthodes basées sur une modélisation statistique et à l'imagerie de diffusion. Le chapitre 6 décrit les tests d'hypothèses utilisés pour le suivi longitudinal en IRM de diffusion.
- **Partie III, Suivi longitudinal des tenseurs d'ordre 2 : prétraitements et outils statistiques :** cette partie se décompose en trois sous chapitres. Le chapitre 7 décrit les prétraitements permettant de corriger les artefacts d'acquisition et d'aligner deux images de tenseurs d'ordre 2. Le chapitre 8 propose une gestion fine des voisinages sur lesquels les tests statistiques sont menés. Le chapitre 9 présente les deux chaînes de traitement opérant sur les tenseurs d'ordre 2.
- **Partie IV, Simulations et études de cas cliniques :** cette partie se décompose en trois sous chapitres. Le chapitre 10 décrit les pathologies étudiées lors de cette thèse. Les chapitres 11 et 12 présentent les résultats de validation obtenus avec les deux chaînes de traitement développées pour les tenseurs d'ordre 2.
- **Partie V, Vers un suivi longitudinal avec des modèles d'ordre supérieur :** cette partie se décompose en trois sous chapitres. Le chapitre 13 décrit les prétraitements spécifiques aux images de tenseurs d'ordre 4. Les chapitres 14 et 15 présentent les deux chaînes de traitement opérant sur les tenseurs d'ordre 4 et les résultats de validation associés.

PARTIE I

Mesure et modélisation de la diffusion des molécules d'eau

Introduction

DANS cette première partie, les différents concepts théoriques nécessaires à la compréhension des travaux de cette thèse dans le domaine de l'imagerie de diffusion sont exposés. Les différentes idées et techniques présentées dans cette partie nous entraîneront de l'acquisition à la modélisation du processus de diffusion des molécules d'eau.

Le premier chapitre introduit l'acquisition des images de diffusion. Il permet de prendre connaissance de l'origine des images traitées, du phénomène physique sous-jacent à mesurer et des méthodes utilisées pour parvenir à cette fin. Dans le deuxième chapitre, une étude détaillée présente le modèle du tenseur de diffusion, les différentes stratégies d'estimation associées, ainsi que les outils permettant une caractérisation synthétique des propriétés de diffusion. Le modèle de diffusion permet de modéliser soit le signal de diffusion, soit la diffusivité, soit la probabilité de déplacement des molécules d'eau. Des contraintes de positivité peuvent être imposées lors de l'estimation du modèle. Les outils mathématiques mis en place dans ce chapitre permettent de modéliser localement le phénomène de diffusion en rendant compte des faisceaux de fibres, ou des croisements de faisceaux de fibres de la substance blanche. Dans le troisième chapitre, le tenseur de diffusion est exprimé sur une base d'harmoniques sphériques. Dans cette représentation, le calcul d'un modèle de la probabilité de déplacement des molécules d'eau, dont les maxima sont alignés sur les fibres, est simplifié. Finalement, dans ce même chapitre, le lien entre ces deux représentations est donné. Dès lors, il sera possible de passer d'une représentation à l'autre en fonction des objectifs ou des applications visés.

L'objectif de cette première partie est de permettre au lecteur de comprendre, voire de reproduire aisément les résultats présentés dans la suite de ce mémoire de thèse. Notons que dans cette thèse, seuls deux modèles de diffusion sont utilisés (les modèles du tenseur de diffusion sur une base de polynômes homogènes et du tenseur de diffusion sur une base d'harmoniques sphériques), sachant qu'il en existe bien d'autres dans la littérature [39, 17, 26, 9, 12, 2, 37, 34] (cette liste n'est pas exhaustive).

Chapitre 1

L'acquisition des images

Sommaire

1.1	La diffusion moléculaire	19
1.1.1	L'équation de diffusion	19
1.1.2	La probabilité de déplacement des molécules d'eau : le propagateur	21
1.2	L'Imagerie par Résonance Magnétique nucléaire (IRM)	22
1.2.1	L'acquisition des images : le signal de diffusion	22
1.2.2	Représentation : le modèle de diffusion	24

EN IRMd, le mouvement d'une molécule d'eau est mesuré dans une direction spécifique (appelée direction de gradient) en effectuant un marquage des molécules d'eau à un temps $t = t_1$ et un deuxième marquage à un temps $t = t_1 + \Delta$. L'atténuation du signal mesurée, induit par le changement de phase des molécules d'eau pendant cet intervalle de temps Δ , caractérise le déplacement des molécules dans la direction d'étude et est reliée au coefficient de diffusion dans cette direction. Cette mesure ne nous informe pas, tout du moins immédiatement, sur la manière dont les molécules se sont déplacées entre les positions initiale et finale. Cette information est contenue dans la probabilité conditionnelle $P(\vec{r}_0, \vec{r}_1, t)$ de trouver une molécule initialement à la position $\vec{r}_0$, à la position $\vec{r}_1$ après un temps t . La probabilité $P(\vec{r}_0, \vec{r}_1, t)$ est obtenue à partir de l'équation de diffusion. Ainsi, ce chapitre commence par un rappel de l'équation de diffusion et de ses origines et montre comment relier $P(\vec{r}_0, \vec{r}_1, t)$ à cette équation (voir section 1.1). Puis, la manière d'estimer $P(\vec{r}_0, \vec{r}_1, t)$ à partir de mesures effectuées dans plusieurs directions de l'espace est présentée (voir section 1.2).

1.1 La diffusion moléculaire

1.1.1 L'équation de diffusion

L'IRM de diffusion (IRMd) et en particulier l'IRM du tenseur de diffusion est une technique qui permet de caractériser *in vivo* l'anisotropie des tissus [6]. Cette technique permet de

quantifier en chaque direction de l'espace la diffusion des molécules d'eau. La diffusion est un transport irréversible de la matière qui est induit thermiquement. Ce mouvement est fonction de la température, de la viscosité du milieu et de la masse des particules. Le phénomène de diffusion explique le flux net (déplacement) des particules des zones de forte concentration vers celles de faible concentration. Voici un exemple simple illustrant la différence entre flux et diffusion. Lorsque l'on ajoute une goutte d'encre dans un verre d'eau, le gradient de concentration entre l'eau et l'encre engendre un flux net qui a tendance à homogénéiser les concentrations du milieu. Maintenant, si l'on remplace la goutte d'encre par une goutte d'eau, la goutte d'eau diffuse autant que l'encre même si ce phénomène n'est pas observable à l'œil. Ici, on ne parle pas de flux puisqu'il n'y a pas de gradient de concentration, mais de diffusion. Au sens large, la diffusion désigne des transferts obéissant aux lois de Fick. En terme de concentration de particules par unité de volume $c(\vec{r}, t)$, le flux d'une particule est donné en un point $\vec{r}$ et pour un instant t par la première loi de diffusion de Fick :

$$\vec{J}(\vec{r}, t) = -D \vec{grad}(c(\vec{r}, t)) \quad (1.1)$$

Le signe moins indique que la direction du flux (dans des milieux isotropes) va des zones de forte concentration vers celles de faible concentration. Quand le processus de diffusion est isotrope, il peut être caractérisé par un unique coefficient de diffusion D (*i.e.*, un scalaire). Pour des processus de diffusion plus complexes (*i.e.*, anisotrope), il est possible de les représenter par un tenseur dans un repère cartésien (x, y, z) . A l'ordre 2, ce tenseur $\mathbf{D}$ est une matrice 3×3 de terme général $d_{\alpha\beta}$, où α et β peuvent prendre chaque direction du repère cartésien (x, y, z) remplacées par $(1, 2, 3)$ dans les notations. Ainsi, on peut réécrire l'équation (1.1) au point $\vec{r} = (x, y, z)$:

$$\begin{pmatrix} J_x(x, y, z, t) \\ J_y(x, y, z, t) \\ J_z(x, y, z, t) \end{pmatrix} = - \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \begin{pmatrix} \frac{\partial c(x, y, z, t)}{\partial x} \\ \frac{\partial c(x, y, z, t)}{\partial y} \\ \frac{\partial c(x, y, z, t)}{\partial z} \end{pmatrix} \quad (1.2)$$

Les éléments diagonaux de $\mathbf{D}$ (*i.e.*, quand $\alpha = \beta$) caractérisent des concentrations de gradient et des flux dans la même direction, alors que les éléments anti-diagonaux (*i.e.*, quand $\alpha \neq \beta$) associent des flux et des concentrations de gradient dans des directions orthogonales. En utilisant l'équation de conservation de la masse, on obtient :

$$\frac{\partial c(\vec{r}, t)}{\partial t} = -div(\vec{J}(\vec{r}, t)) \quad (1.3)$$

Ainsi, la quantité $\frac{\partial c(\vec{r}, t)}{\partial t}$ est égale à la différence entre l'afflux et l'efflux du point situé à la position $\vec{r}$. Finalement, en combinant (1.1) et (1.3), on obtient la deuxième loi de diffusion de Fick :

$$\frac{\partial c(\vec{r}, t)}{\partial t} = div(\mathbf{D} \vec{grad}(c(\vec{r}, t))) \quad (1.4)$$

Dans le cas d'un processus de diffusion isotrope on peut simplifier (1.4) :

$$\frac{\partial c(\vec{r}, t)}{\partial t} = D \nabla^2 c(\vec{r}, t) \quad (1.5)$$

où ∇^2 est l'opérateur Laplacien.

1.1.2 La probabilité de déplacement des molécules d'eau : le propagateur

Dans le cas d'un processus de diffusion sans gradient de concentration, la quantité d'intérêt est la probabilité $P(\vec{r}_1, t)$ de trouver une particule à la position $\vec{r}_1$ à l'instant t . Cette probabilité est donnée par la relation :

$$P(\vec{r}_1, t) = \int \rho(\vec{r}_0) P(\vec{r}_0, \vec{r}_1, t) d\vec{r}_0 \quad (1.6)$$

où $\rho(\vec{r}_0)$ représente la densité de probabilité de présence d'une particule à la position $\vec{r}_0$ et $P(\vec{r}_0, \vec{r}_1, t)$ représente la probabilité d'une particule initialement à la position $\vec{r}_0$ de se déplacer à la position $\vec{r}_1$ à l'instant t . Intégrer sur $\vec{r}_0$ revient à prendre en compte toutes les positions de départ possibles. Puisque $P(\vec{r}_1, t)$ décrit la probabilité de trouver une particule à un emplacement et à un temps donnés, on peut assimiler cette quantité à une concentration. Il est ainsi raisonnable de faire l'hypothèse que cette quantité vérifie (1.4). Il est possible de réécrire (1.4) pour obtenir l'équation de Fokker-Planck :

$$\frac{\partial P(\vec{r}_1, t)}{\partial t} = \text{div} \left(\mathbf{D} \text{grad} (P(\vec{r}_1, t)) \right) \quad (1.7)$$

La probabilité $P(\vec{r}_0, \vec{r}_1, t)$ (appelée également propagateur) est solution de l'équation de Fokker-Planck. Pour un processus de diffusion isotrope en trois dimensions et dans un milieu homogène, le propagateur est donné par [23, 28] :

$$P(\vec{r}_0, \vec{r}_1, t) = (4\pi Dt)^{-3/2} \exp \left(-\frac{(\vec{r}_1 - \vec{r}_0)^2}{4Dt} \right) \quad (1.8)$$

où D est un scalaire. Cette équation formule que la fonction de distribution des molécules d'eau pour une référence des temps arbitraire est gaussienne. De plus, $P(\vec{r}_0, \vec{r}_1, t)$ ne dépend pas de la position initiale $\vec{r}_0$, mais uniquement du déplacement absolu $\vec{r}_1 - \vec{r}_0$ (le vecteur $\vec{r}_1 - \vec{r}_0$ est souvent noté $\vec{R}$). La solution est plus compliquée lorsqu'un processus de diffusion anisotrope est considéré (*i.e.*, le mouvement des particules n'est pas le même dans chaque direction de l'espace). L'extension pour une diffusion anisotrope s'écrit [19] :

$$P(\vec{R}, t) = ((4\pi t)^3 |\mathbf{D}|)^{-1/2} \exp \left(-\frac{\vec{R}^T \mathbf{D}^{-1} \vec{R}}{4t} \right) \quad (1.9)$$

La moyenne quadratique des déplacements est alors donnée par :

$$\langle (\vec{r}_1 - \vec{r}_0)^2 \rangle = \int \int (\vec{r}_1 - \vec{r}_0)^2 \rho(\vec{r}_0) P(\vec{r}_0, \vec{r}_1, t) d\vec{r}_0 d\vec{r}_1 \quad (1.10)$$

Pour une diffusion libre (*i.e.*, isotrope), la probabilité $P(\vec{r}_0, \vec{r}_1, t)$ peut être remplacée par son expression (1.8) et la moyenne quadratique des déplacements est donnée par [23, 28] :

$$\langle (\vec{r}_1 - \vec{r}_0)^2 \rangle = nDt \quad (1.11)$$

où $n = 2, 4$ ou 6 pour un processus de diffusion en une, deux ou trois dimensions. Ainsi, pour une diffusion libre, la moyenne quadratique des déplacements varie linéairement avec le temps. De plus, d'après (1.9), on remarque que $P(\vec{R}, t) = P(-\vec{R}, t)$. Ce résultat est en accord avec la physique d'acquisition puisque la diffusion est un processus radialement symétrique.

1.2 L'Imagerie par Résonance Magnétique nucléaire (IRM)

1.2.1 L'acquisition des images : le signal de diffusion

Les gradients de champ magnétique comme marqueurs spatiaux

Des spins dans un champ magnétique uniforme $\vec{B}_0$ décrivent un mouvement de précession autour de l'axe défini par le champ magnétique avec une vitesse de rotation donnée par la fréquence de Larmor :

$$\omega_0 = \gamma B_0 \quad (1.12)$$

où ω_0 est la fréquence de Larmor ($rad\ s^{-1}$), γ le rapport gyromagnétique ($rad\ T^{-1}s^{-1}$) et B_0 l'amplitude du champ magnétique statique (T). γ est une caractéristique du noyau, et dans le cas de l'hydrogène H^1 , $\gamma/2\pi = 42,58\ MHz\ T^{-1}$. On fait l'hypothèse que ce champ magnétique $\vec{B}_0$ est orienté suivant l'axe z (dans un repère cartésien). Puisque B_0 est homogène dans l'espace, ω_0 est le même dans tout l'échantillon. Si l'on superpose à $\vec{B}_0$ un gradient de champ magnétique $\vec{g}$ ($T\ m^{-1}$) qui varie spatialement, la fréquence de Larmor devient une fonction de la position spatiale $\vec{r}$ dans l'échantillon [28] :

$$\omega(\vec{r}) = \omega_0 + \gamma \vec{G} \cdot \vec{r} \quad (1.13)$$

Ainsi, lorsque l'on applique un champ magnétique d'amplitude connue à un échantillon, la fréquence de Larmor devient un marqueur spatial qui est fonction de la direction d'application du gradient de champ magnétique (ce principe est classiquement utilisé pour la sélection de coupe en IRM). Pour simplifier les calculs, considérons un gradient orienté suivant l'axe z . Dans ce cas, l'amplitude de $\vec{G}$ ne dépend que de sa projection sur l'axe z , $G_z = \vec{G} \cdot \vec{z}$. Cette amplitude est désormais notée G .

Pour simplifier les calculs dans l'introduction proposée dans ce mémoire de thèse, nous nous intéresserons uniquement au cas idéal d'impulsions de type créneau (on fait l'hypothèse que l'amplitude de gradient est constante lors de l'application de l'impulsion). A partir de (1.13), pour un spin se déplaçant par diffusion dans le milieu, le déphasage (*i.e.*, changement de phase) cumulé sur une durée t est donné par :

$$\phi(t) = \gamma B_0 t + \gamma G \int_0^t z(t') dt' \quad (1.14)$$

Le premier terme correspond au déphasage introduit par le champ magnétique permanent et le deuxième terme correspond au déphasage introduit par le gradient de champ magnétique.

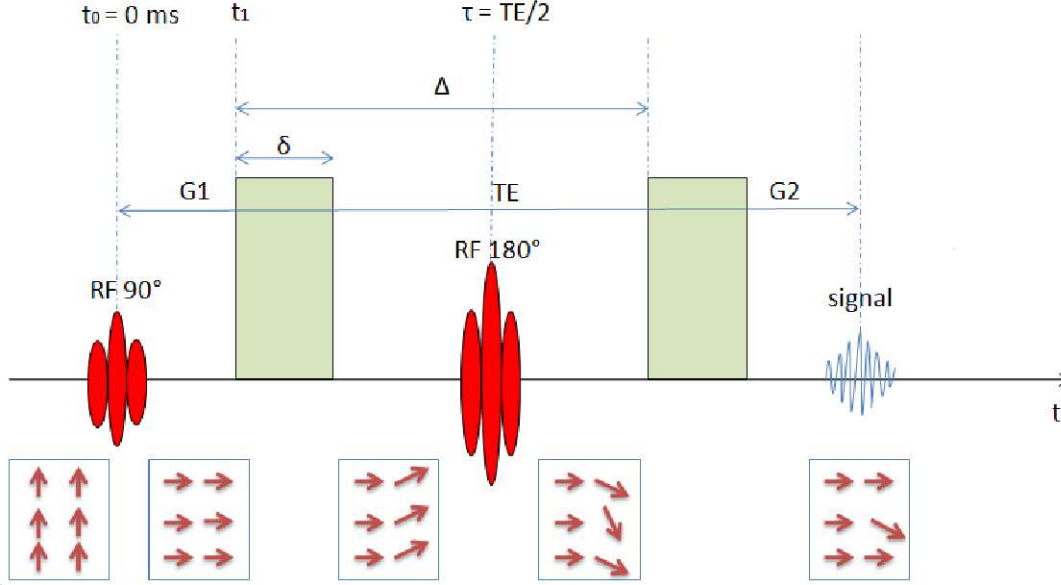


FIG. 1.1 – Séquence d'acquisition de diffusion PGSE. Les flèches rouges représentent la direction de l'aimantation des spins.

Mesurer la diffusion en appliquant des gradients de champ magnétique

A partir de (1.13), il est clair que l'application d'un gradient de champ magnétique peut être utilisé comme marqueur de la position d'un spin (par l'intermédiaire de la fréquence de Larmor). Cette observation est à la base de toute mesure de diffusion. Pour ce faire, l'approche la plus commune est de modifier la séquence d'écho de spin en introduisant deux impulsions de gradient identiques de forme créneau autour de l'impulsion 180° de rephasage [35]. La figure 1.1 décrit cette séquence d'acquisition appelée PGSE pour *Pulse-Gradient Spin-Echo*. Notons respectivement G_1 et G_2 les gradients de champ magnétique dans la direction $\vec{z}$ appliqués pendant une impulsion de type créneau (hypothèse de vitesse de commutation des gradients infinie). Décrivons comment la séquence de la figure 1.1 permet de mesurer le phénomène de diffusion. Soit un ensemble de spins diffusant à l'équilibre thermique. Une impulsion radiofréquence (RF) de 90° est appliquée pour basculer l'aimantation alignée suivant l'axe z dans le plan $x - y$ (le plan perpendiculaire au champ permanent). Ensuite, le gradient G_1 est appliqué au temps t_1 . A l'instant $\tau = TE/2$, le spin i présente un déphasage (1.14) :

$$\phi_i(\tau) = \gamma B_0 \tau + \gamma G \int_{t_1}^{t_1 + \delta} z_i(t) dt \quad (1.15)$$

A l'instant $TE/2$, un pulse RF de 180° permet d'inverser le signe de l'orientation de l'aimantation des spins. La deuxième impulsion G_2 est appliquée au temps $t_1 + \Delta$. Si les spins n'ont pas subi de mouvement de translation suivant l'axe z , l'effet des deux impulsions G_1 et G_2 s'annule et les spins sont rephasés. Dans le cas contraire, le degré de déphasage des spins est proportionnel à leurs déplacements dans la direction de gradient (ici la direction z) pendant l'intervalle de temps Δ . Le déphasage dépend donc du parcours des spins considérés pen-

dant l'intervalle de temps Δ séparant les deux gradients de diffusion. A la fin de la séquence PGSE, au temps TE , le déphasage total du spin i est donné par :

$$\begin{aligned}\phi_i(TE) &= \left\{ \gamma B_0 \tau + \gamma G \int_{t_1}^{t_1+\delta} z_i(t) dt \right\} - \left\{ \gamma B_0 \tau + \gamma G \int_{t_1+\Delta}^{t_1+\Delta+\delta} z_i(t) dt \right\} \\ &= \gamma G \left\{ \int_{t_1}^{t_1+\delta} z_i(t) dt - \int_{t_1+\Delta}^{t_1+\Delta+\delta} z_i(t) dt \right\}\end{aligned}\quad (1.16)$$

Maintenant, au lieu d'intégrer sur tous les spins, l'intégration se fait sur les déphasages. Ainsi, l'atténuation du signal au temps TE est donnée par [28] :

$$S(TE) = S(TE)_{g=0} \int_{-\infty}^{+\infty} P(\phi, TE) \exp(i\phi) d\phi \quad (1.17)$$

où $S(TE)_{g=0}$ est le signal (*i.e.*, moment magnétique résultant) en l'absence de gradient de champ magnétique et $P(\phi, TE)$ est la fonction de distribution des phases à l'instant TE . En appliquant la relation de Moivre dans l'équation (1.17) ($\exp(i\phi) = \cos\phi + i\sin\phi$) et en ne conservant que la partie réelle du signal $S(TE)$ (la partie imaginaire s'annule car $P(\phi, TE)$ est une fonction paire par rapport à la variable ϕ), on obtient :

$$S(TE) = S(TE)_{g=0} \int_{-\infty}^{+\infty} P(\phi, TE) \cos(\phi) d\phi \quad (1.18)$$

Les équations (1.16) et (1.18) permettent de comprendre ce que l'on mesure en IRMd. A partir de (1.16), on remarque que le déphasage introduit par le champ permanent s'annule. En l'absence de diffusion, le déphasage introduit par les deux impulsions G_1 et G_2 est également nul. Ainsi la phase ϕ_i de chaque spin i est nulle et en introduisant $\phi = 0$ dans (1.18) un maximum de signal est enregistré. En présence de diffusion, la fonction de déplacement $z_i(t)$ dépend du temps et le déphasage accumulé par chaque spin lors de l'application des deux impulsions G_1 et G_2 ne s'annule pas. Plus le déplacement introduit par la diffusion (*i.e.*, mouvement aléatoire) dans la direction du gradient est grand, plus le terme de déphasage augmente. Ces déphasages aléatoires issus du phénomène de diffusion sont moyennés sur l'ensemble des spins qui contribuent à la mesure du signal. Ainsi, le signal observé n'est pas déphasé mais atténué. Plus la diffusion est importante, plus l'atténuation du signal au temps TE est importante. De même, plus l'amplitude de gradient est augmentée en présence de diffusion, plus le signal mesuré au temps TE est diminué.

1.2.2 Représentation : le modèle de diffusion

Maintenant, il nous reste à formuler mathématiquement le lien entre l'atténuation du signal et le coefficient de diffusion. Dans le cas d'une diffusion libre, des relations analytiques peuvent être trouvées. Pour une diffusion contrainte, aucune relation analytique ne peut être formulée et il faut avoir recours à des approximations. Dans la littérature, deux approximations sont couramment utilisées. Au premier ordre, l'approximation SGP (*Short Gradient Pulse*) fait l'hypothèse que les gradients de diffusion sont appliqués instantanément (distributions de Dirac) [28]. Ceci est vrai lorsque l'on fait l'hypothèse $\delta \ll \Delta$, permettant

de négliger le processus de diffusion pendant l'application des gradients. Or en pratique $\delta \approx 5 \sim 30 \text{ ms}$ et on s'efforce d'obtenir un $T_E < 100 \text{ ms}$, ainsi il n'est pas possible de négliger le déplacement des molécules d'eau pendant l'application des gradients de diffusion. L'approximation au deuxième ordre GPD (*Gaussian Phase Distribution*) permet de prendre en compte ce phénomène et est présentée dans cette partie.

L'approximation GPD

On a montré que le signal au temps d'écho peut être défini par (1.18) avec le terme de phase ϕ défini par (1.16). On cherche à exprimer $P(\phi, TE)$ à partir de (1.16). On s'intéresse ici au cas où la diffusion est isotrope avant de généraliser, dans la prochaine section, au cas anisotrope. On remarque que $z_i(t)$ est décrit par l'équation de diffusion à une dimension qui est une gaussienne dans le cas d'une diffusion libre (*i.e.*, la version monodimensionnelle de (1.8) obtenue en intégrant selon x et y) :

$$P(0, z, t) = (4\pi Dt)^{-1/2} \exp\left(-\frac{z^2}{4Dt}\right) \quad (1.19)$$

Maintenant, puisque l'intégrale d'une variable suivant une probabilité gaussienne (*i.e.*, $z_i(t)$) suit elle même une distribution gaussienne, on a [28] :

$$P(\phi, TE) = (2\pi\langle\phi^2\rangle)^{-1/2} \exp\left(-\frac{\phi^2}{2\langle\phi^2\rangle}\right) \quad (1.20)$$

où $\langle\phi^2\rangle$ est la moyenne quadratique des changements de phase au temps d'écho. Cette quantité est donnée par [28] :

$$\langle\phi^2\rangle = \gamma^2 G^2 2\delta^2 \left(\Delta - \frac{\delta}{3}\right) D \quad (1.21)$$

Si on évalue (1.18) en utilisant la distribution des phases (1.20), on trouve que l'atténuation du signal au temps d'écho est donnée par :

$$E = \frac{S(TE)}{S(TE)_{g=0}} = \exp\left(-\frac{\langle\phi^2\rangle}{2}\right) \quad (1.22)$$

En remplaçant (1.21) dans (1.22), on obtient finalement :

$$\ln(E) = -\gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right) D \quad (1.23)$$

Généralisation : diffusion anisotrope

La diffusion isotrope présentée dans la partie précédente est un cas particulier de la diffusion anisotrope. Dans le cas d'une diffusion anisotrope, le phénomène de diffusion est représenté par un tenseur cartésien $\mathbf{D}$. Dans un tel système, le signal mesuré au temps d'écho dépend de l'orientation suivant laquelle la mesure est effectuée (*i.e.*, la direction du gradient $\vec{g} = [g_1, g_2, g_3]^T$). Par exemple à l'ordre 2, pour une diffusion anisotrope, le scalaire D dans (1.23) doit être remplacé par $\vec{g}^T \cdot \mathbf{D} \cdot \vec{g}$, où :

$$\vec{g}^T \cdot \mathbf{D} \cdot \vec{g} = \sum_{\alpha} \sum_{\beta} d_{\alpha\beta} g_{\alpha} g_{\beta}, \quad \alpha, \beta = 1, 2, 3 \quad (1.24)$$

Et on obtient (1.23) :

$$\ln(E) = -\gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right) \vec{g}^T \cdot \mathbf{D} \cdot \vec{g} \quad (1.25)$$

Soient N directions de gradient dans l'espace noté $\vec{g}_i$. Notons E_i les atténuations du signal associées ($S_i(TE)/S(TE)_{g=0} = S_i/S_0$) et $b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right)$ ($s \text{ mm}^{-2}$) la b-valeur. On obtient finalement l'équation de Stejskal-Tanner [35] :

$$E_i = \frac{S_i}{S_0} = \exp(-b \vec{g}_i^T \cdot \mathbf{D} \cdot \vec{g}_i) \quad (1.26)$$

Afin de faire le lien entre l'équation (1.26) et (1.23), il suffit de remarquer que $\vec{g}_i^T \mathbf{D} \vec{g}_i$ est un scalaire donnant la mesure de la diffusion dans la direction spatiale $\vec{g}_i$. Ainsi tout est identique, à condition que la mesure s'effectue dans une direction spatiale donnée, d'où les indices $i = 1 \dots N$. Il est important de souligner que l'équation (1.26) permet uniquement de détecter un déplacement dans la direction des gradients.

Chapitre 2

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

Sommaire

2.1	L’Imagerie du Tenseur de Diffusion (ITD)	28
2.1.1	Formulation générale du modèle	28
2.1.2	Cas particulier : le tenseur d’ordre 2	29
2.1.3	Cas particulier : le tenseur d’ordre 4	35
2.2	L’estimation des paramètres du modèle en ITD	39
2.2.1	La méthode des moindres carrés : modélisation de la diffusivité	39
2.2.2	La méthode des quartiques : modélisation de la diffusivité	40
2.2.3	Modélisation de la probabilité de déplacement : la fonction de distribution des orientations	43
2.2.4	Comparaison des différentes méthodes d’estimation de tenseurs	46

EN IRMd, des mesures effectuées dans plusieurs directions de l’espace (appelées directions de gradient) nous renseignent sur le déplacement des molécules d’eau dans le milieu. Le phénomène de diffusion peut-être représenté par un tenseur (qui contient tous les paramètres nécessaires pour représenter le phénomène). Quand on développe l’expression analytique d’un tenseur, on obtient un polynôme homogène (polynôme dont les termes sont des monômes dont le degré est le même) (2.2). Ce chapitre commence par un rappel des différents outils mathématiques nécessaires à la modélisation du signal de diffusion (voir section 2.1). Après une formulation générale du problème de modélisation, les modèles et les outils mathématiques associés au tenseur d’ordre deux et d’ordre quatre sont détaillés. Puis l’équation de Stejskal-Tanner (1.26) (qui établit la relation entre le signal de diffusion S et le tenseur de diffusion $\mathbf{D}$) est utilisée pour estimer les paramètres du modèle, en utilisant différentes stratégies d’estimation (voir section 2.1). Une première méthode d’estimation au sens des moindres carrés est présentée. Elle permet d’estimer analytiquement des tenseurs modélisant la diffusivité, sans contrainte sur la positivité du tenseur estimé. Puis, une seconde méthode d’estimation basée sur une paramétrisation des tenseurs par des polynômes d’ordre

pair est décrite, permettant ainsi de prendre en compte la contrainte de positivité. En utilisant la même paramétrisation, un modèle décrivant le déplacement des molécules d'eau peut également être estimé. L'importance du choix du modèle lors de la modélisation d'un signal provenant de différentes configurations de faisceaux de fibres est également discutée (voir section 2.2). Il est important de différencier les méthodes permettant de modéliser la diffusivité (qui n'est pas nécessairement alignée avec les structures anatomiques sous-jacentes pour des modèles d'ordre supérieur) et les méthodes permettant de représenter le propagateur (qui lui l'est toujours).

2.1 L'Imagerie du Tenseur de Diffusion (ITD)

2.1.1 Formulation générale du modèle

Un tenseur d'ordre l permet de modéliser la dépendance entre les mesures de diffusion et la diffusivité des molécules d'eau [25]. Plus précisément, le tenseur de diffusion est la matrice de covariance du déplacement des molécules d'eau. Rappelons qu'un tenseur d'ordre l est un tableau à l dimensions, qui inclut notamment les scalaires ($l = 0$), les vecteurs ($l = 1$) et les matrices ($l = 2$). Soit la direction de mesure $\vec{g} = [g_1, g_2, g_3]$, l'équation (1.26) peut se généraliser de la façon suivante :

$$S(\vec{g}) = S_0 \exp \left(-b \sum_{i_1=1}^3 \sum_{i_2=1}^3 \dots \sum_{i_l=1}^3 d_{i_1, i_2, \dots, i_l} g_{i_1} g_{i_2} \dots g_{i_l} \right) \quad (2.1)$$

où $d_{i_1, i_2, \dots, i_l}$ représente les composantes du tenseur d'ordre l . Le profil de diffusion représenté par ce tenseur peut alors s'écrire :

$$d(\vec{g}) = \sum_{i_1=1}^3 \sum_{i_2=1}^3 \dots \sum_{i_l=1}^3 d_{i_1, i_2, \dots, i_l} g_{i_1} g_{i_2} \dots g_{i_l} \quad (2.2)$$

Sans imposer aucune restriction sur ce profil, il est possible d'écrire la relation suivante :

$$d(-\vec{g}) = \begin{cases} d(\vec{g}) & \text{si } l \text{ pair} \\ -d(\vec{g}) & \text{si } l \text{ impair} \end{cases} \quad (2.3)$$

Or si on accepte le deuxième sous-cas (l impair), on autorise des valeurs de diffusion négatives, ce qui est en désaccord avec la physique d'acquisition et ne respecte également pas la propriété de symétrie radiale du processus de diffusion. Seuls les tenseurs d'ordre pair sont donc autorisés. Un tenseur d'ordre l en dimension 3 est composé de 3^l éléments. La propriété de symétrie totale associée à un tenseur de diffusion permet de réduire le nombre d'éléments constitutifs du tenseur. La symétrie totale vient du fait qu'un tenseur de diffusion doit relier les composantes d'un même vecteur au scalaire $d(\vec{g})$. Par exemple, lorsque $l = 2$ on peut écrire :

$$d(\vec{g}) = \sum_{i_1=1}^3 \sum_{i_2=1}^3 d_{i_1, i_2} g_{i_1} g_{i_2} = \sum_{i_1=1}^3 \sum_{i_2=1}^3 d_{i_1, i_2} g_{i_2} g_{i_1}$$

$$= \sum_{i_2=1}^3 \sum_{i_1=1}^3 d_{i_2,i_1} g_{i_1} g_{i_2} = \sum_{i_1=1}^3 \sum_{i_2=1}^3 d_{i_2,i_1} g_{i_1} g_{i_2} \quad (2.4)$$

Cette relation implique $d_{i_1,i_2} = d_{i_2,i_1}$ puisqu'elle est vérifiée pour n'importe quel vecteur $\vec{g}$. Une analyse similaire pour un tenseur d'ordre l donne :

$$d_{i_1,i_2,\dots,i_l} = d_{(i_1,i_2,\dots,i_l)} \quad (2.5)$$

où $(i_1, i_2, \dots, i_l)$ représente toutes les permutations possibles d'indices. En trois dimensions un tenseur symétrique a :

$$N_l = \Gamma_3^l = \frac{(l+1)(l+2)}{2} \quad (2.6)$$

composantes distinctes (Γ_3^l représente le nombre de l-combinaisons avec répétition de 3 éléments), où chaque élément d_k ($k \in [1, \dots, N_l]$) est répété μ_k^l fois, avec :

$$\mu_k^l = \binom{l}{n_1^l(k)} \binom{l - n_1^l(k)}{n_2^l(k)} = \frac{l!}{n_1^l(k)! n_2^l(k)! n_3^l(k)!} \quad (2.7)$$

$n_1^l(k)$, $n_2^l(k)$ et $n_3^l(k)$ sont respectivement le nombre d'indice 1, 2 et 3 définissant la composante du k ième élément indépendant du tenseur étudié :

$$d_k = \underbrace{d_1, \dots, 1}_{n_1^l(k)}, \underbrace{2, \dots, 2}_{n_2^l(k)}, \underbrace{3, \dots, 3}_{n_3^l(k)} \quad (2.8)$$

avec $n_1^l(k) + n_2^l(k) + n_3^l(k) = l$. Finalement, il est possible de réécrire l'équation (2.1) :

$$S(\vec{g}) = S_0 \exp \left(-b \sum_{k=1}^{N_l} \mu_k^l d_k (g_1)^{n_1^l(k)} (g_2)^{n_2^l(k)} (g_3)^{n_3^l(k)} \right) \quad (2.9)$$

Dans la suite de cet exposé, nous développerons plus en détail le cas $l = 2$ (tenseur d'ordre 2, le plus utilisé dans la littérature) et le cas $l = 4$ (tenseur d'ordre 4).

2.1.2 Cas particulier : le tenseur d'ordre 2

Le tenseur d'ordre 2 est un modèle couramment utilisé dans la littérature qui repose sur l'hypothèse d'un processus de diffusion gaussien. Lorsque $l = 2$, la matrice symétrique définie positive $\mathbf{D}$ contient $N_l = 6$ éléments distincts. Les termes diagonaux d_{11} , d_{22} et d_{33} ont une multiplicité $\mu = 1$, alors que les termes anti-diagonaux d_{12} , d_{13} et d_{23} ont une multiplicité $\mu = 2$. $\mathbf{D}$ est donc une matrice 3×3 de la forme :

$$\mathbf{D} = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ * & d_{22} & d_{23} \\ * & * & d_{33} \end{pmatrix} \quad (2.10)$$

où les $*$ représentent les éléments symétriques.

Décomposition d'un tenseur d'ordre 2

Les valeurs propres de la matrice symétrique définie positive associée au tenseur d'ordre 2 sont les valeurs de diffusivité données dans les axes du repère de diagonalisation contenus dans la matrice des vecteurs propres. Ainsi, la direction de plus grande diffusion est donnée par le vecteur propre du tenseur associé à la plus grande valeur propre. La processus de diagonalisation est le suivant : on cherche une base (celle formée par les vecteurs propres) dans laquelle $\mathbf{D}$ est diagonale (avec pour éléments diagonaux les valeurs propres). Rappelons que d'après le théorème spectral, toute matrice symétrique à éléments réels est diagonalisable à l'aide d'une matrice orthogonale. Ses valeurs propres sont donc réelles et ses sous-espaces propres sont orthogonaux. Notons $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ ces valeurs propres, et $\vec{e}_1, \vec{e}_2, \vec{e}_3$ les vecteurs propres associés. Une propriété intrinsèque au tenseur $\mathbf{D}$ est qu'il est défini positif, ce qui impose que toutes les valeurs propres soient positives. Soit $\mathbf{U}$ la matrice orthogonale de passage constituée des colonnes de coordonnées $\vec{e}_1, \vec{e}_2, \vec{e}_3$ écrits dans la base canonique. Alors, on peut vérifier que :

$$\begin{aligned} \mathbf{D}\vec{e}_i &= \lambda_i\vec{e}_i \\ \text{ou} \\ (\mathbf{D} - \lambda_i\mathbf{I})\vec{e}_i &= 0 \end{aligned} \quad (2.11)$$

avec $\mathbf{I}$ la matrice identité. Notons que les valeurs propres λ_i apparaissent sur la diagonale de la matrice dans le même ordre que nous avons placé les vecteurs propres pour former $\mathbf{U}$. Les valeurs propres λ_i peuvent être trouvées en résolvant l'équation caractéristique suivante :

$$\begin{vmatrix} d_{11} - \lambda & d_{12} & d_{13} \\ d_{21} & d_{22} - \lambda & d_{23} \\ d_{31} & d_{32} & d_{33} - \lambda \end{vmatrix} = 0 \quad (2.12)$$

En développant le déterminant, on obtient :

$$\lambda^3 - \lambda^2(d_{11} + d_{22} + d_{33}) + \lambda \left(\begin{vmatrix} d_{22} & d_{33} \\ d_{23} & d_{33} \end{vmatrix} + \begin{vmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{vmatrix} + \begin{vmatrix} d_{11} & d_{31} \\ d_{13} & d_{33} \end{vmatrix} \right) - \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix} = 0 \quad (2.13)$$

où λ est un scalaire qui est indépendant du choix du système de coordonnées. On peut ainsi déduire les quantités invariantes par rotation du tenseur $\mathbf{D}$:

$$\begin{aligned} I_1 &= d_{11} + d_{22} + d_{33} \\ I_2 &= \begin{vmatrix} d_{22} & d_{33} \\ d_{23} & d_{33} \end{vmatrix} + \begin{vmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{vmatrix} + \begin{vmatrix} d_{11} & d_{31} \\ d_{13} & d_{33} \end{vmatrix} \\ I_3 &= \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix} \end{aligned} \quad (2.14)$$

Notons que I_1 est la trace et I_3 est le déterminant de $\mathbf{D}$. Ces quantités invariantes permettent de définir des indices scalaires caractérisant le tenseur. La transformée inverse, d'un système de valeurs propres et vecteurs propres vers un tenseur, est donnée par la relation suivante :

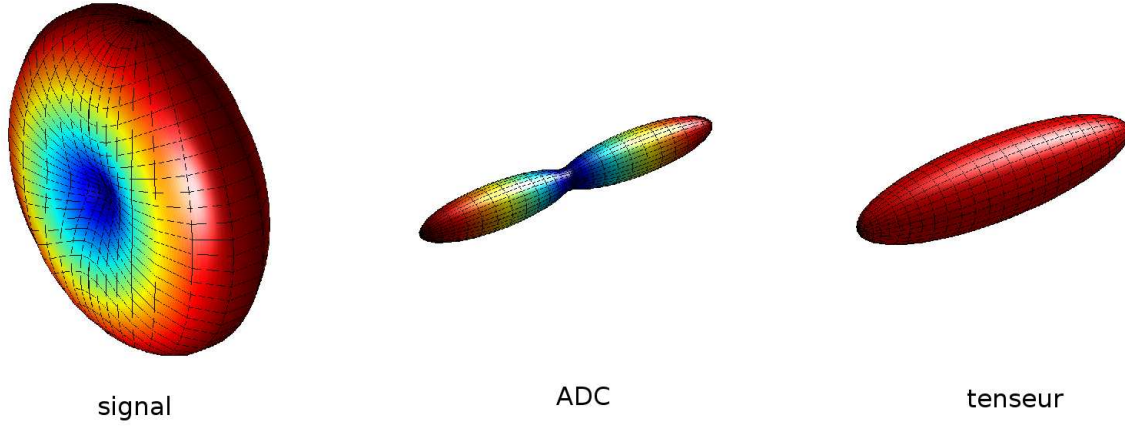


FIG. 2.1 – Différente représentation pour un tenseur ($\mathbf{D} = \{d_{11} = 2, d_{22} = 0.5, d_{33} = 0.5, d_{12} = 0, d_{13} = 0, d_{23} = 0\}$) : le signal de diffusion $S(\vec{g})$, l'ADC qui représente le profil de diffusion $d(\vec{g})$ et le tenseur qui correspond à la surface d'isoprobabilité du propagateur $\vec{g}^T \mathbf{D} \vec{g} = cste$.

$$\mathbf{D} = [\vec{e}_1, \vec{e}_2, \vec{e}_3] \text{diag}[\lambda_1, \lambda_2, \lambda_3] [\vec{e}_1, \vec{e}_2, \vec{e}_3]^{-1} = \mathbf{U} \mathbf{V} \mathbf{U}^{-1} \quad (2.15)$$

Puisque $\mathbf{U}$ est une matrice orthogonale ($\mathbf{D}$ symétrique), $\mathbf{U}^{-1} = \mathbf{U}^T$. Pour une matrice symétrique, l'équation précédente se simplifie :

$$\mathbf{D} = \sum_{i=1}^3 \lambda_i \vec{e}_i \vec{e}_i^T \quad (2.16)$$

Généralement, le tenseur de diffusion est représenté par un ellipsoïde. Toutefois, il est intéressant de noter que l'ellipsoïde ne correspond pas au profil de diffusion qui a en réalité la forme d'une cacahuète, mais à une surface d'isoprobabilité du propagateur. Le profil de diffusion $d(\vec{g})$ est tracé à partir de la relation (2.2), soit $d(\vec{g}) = \vec{g}^T \mathbf{D} \vec{g}$, pour tout $\vec{g}$ unitaire. La surface représentative de l'ellipsoïde est calculée à partir de la relation $\vec{g}^T \mathbf{D} \vec{g} = cste$. Ce n'est pas une fonction de $\vec{g}$, mais une fonction implicite pour laquelle on trace les valeurs de $\vec{g}$ vérifiant la relation ($\vec{g}$ n'est donc plus nécessairement unitaire). Finalement, le signal de diffusion peut être tracé à partir de la relation (2.1), soit $S(\vec{g}) = S_0 \exp(-b \vec{g}^T \mathbf{D} \vec{g})$, pour tout $\vec{g}$ unitaire. Cette différence de représentation est illustrée dans la figure 2.1.

Les indices scalaires caractérisant un tenseur d'ordre 2

Une fois les six paramètres du tenseur de diffusion calculés en chaque voxel (cf. partie 2.2), des méthodes de visualisation ont été développées pour apprécier la neuro-anatomie cérébrale. Pour visualiser toute l'information disponible, un ellipsoïde peut être placé en chaque voxel de l'image. Cette option n'est pas adaptée à la routine clinique puisque l'analyse d'une forme géométrique (même simple) en chaque voxel d'une image 3-D est compliquée (cf. figure 2.2).

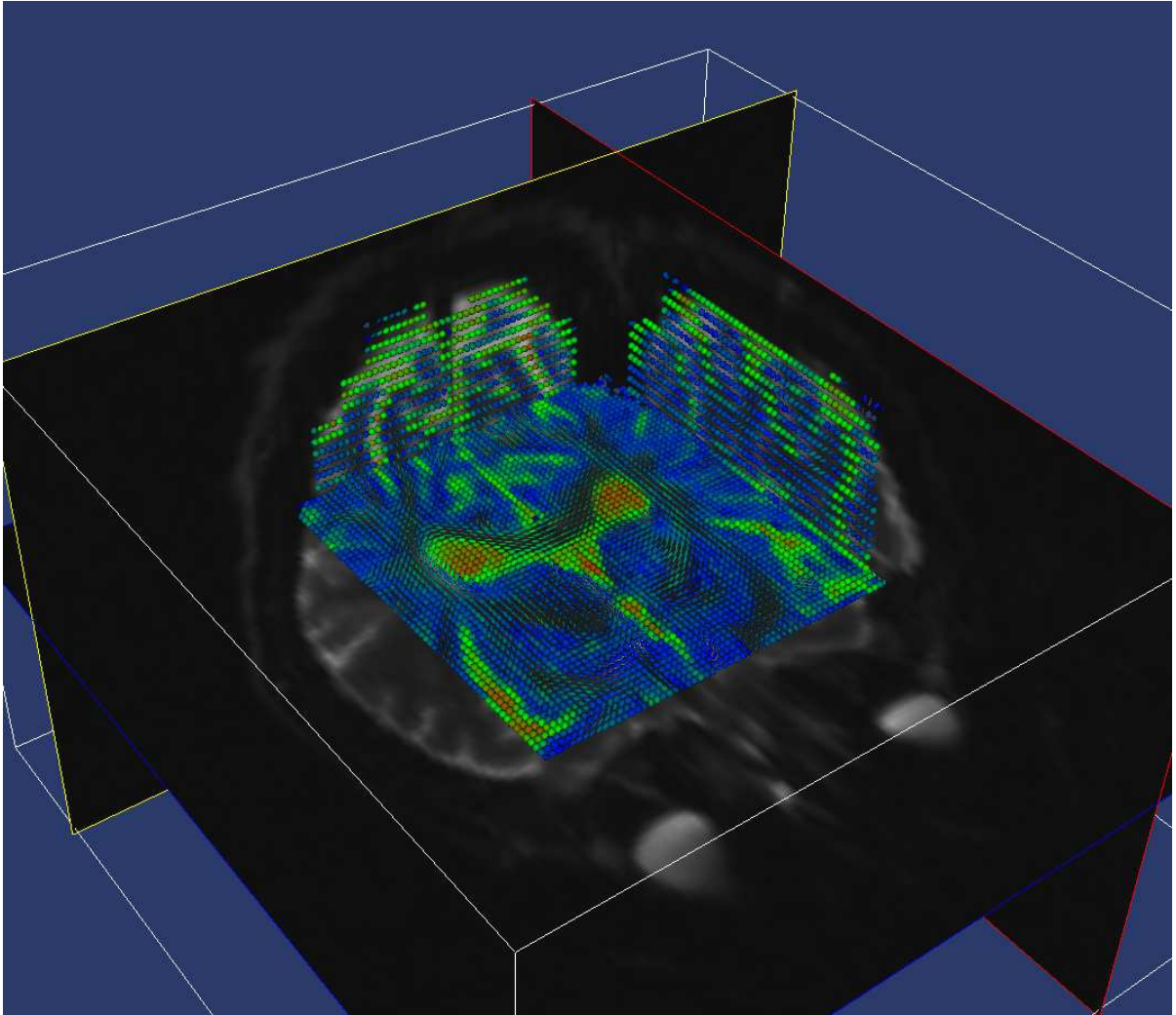


FIG. 2.2 – Visualisation du champ de tenseurs issu d'une modélisation à l'ordre 2.

De ce fait, des mesures scalaires ont été imaginées pour réduire les six paramètres du tenseur en chaque voxel. Les deux indices les plus utilisés sont la carte des coefficients de diffusion apparente (ou *Apparent Diffusion Coefficients* (ADC)) et la carte de fraction d'anisotropie. La carte d'ADC moyenne (indice noté également MD pour *Mean Diffusivity*) est obtenue en considérant le tiers de la trace de chaque tenseur de l'image, $I_1/3$ (cf. figure 2.3). Cet indice est invariant par rotation du tenseur et permet de caractériser la diffusivité du milieu : plus la diffusion est contrainte et plus l'indice est faible (par exemple dans la substance blanche, cf. figure 2.3), et inversement, moins la diffusion est contrainte et plus l'indice est grand (par exemple dans le liquide céphalo-rachidien au niveau des ventricules, cf. figure 2.3). L'anisotropie peut être caractérisée de plusieurs manières. Il est intéressant d'accéder à ce paramètre puisqu'il est directement relié au degré d'organisation des tissus. La façon la plus simple consiste à considérer le rapport entre la valeur propre maximale et minimale (λ_1/λ_3). Plus l'ellipse est allongée et plus cette quantité augmente. Malheureusement, cette mesure est sensible au bruit de mesure et ne tient surtout pas compte de la seconde valeur

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

propre. Une autre façon de caractériser l'anisotropie utilise la différence des trois valeurs propres. Cet indice vaut zéro pour une sphère ($\lambda_1 = \lambda_2 = \lambda_3$) et augmente lorsque l'ellipse s'éloigne de la sphère. Après normalisation entre 0 et 1 on obtient la Fraction d'Anisotropie (FA) [5, 7] (cf. figure 2.3) :

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \langle \lambda \rangle)^2 + (\lambda_2 - \langle \lambda \rangle)^2 + (\lambda_3 - \langle \lambda \rangle)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (2.17)$$

où $\langle \lambda \rangle$ représente la moyenne des valeurs propres. Il existe de nombreuses autres manières de représenter l'anisotropie des tenseurs de diffusion. A titre d'exemple, Westin *et coll.* ont proposé de classifier les tenseurs de diffusion en trois sous catégories [41] :

- cas linéaire ($\lambda_1 \gg \lambda_2 \approx \lambda_3$) : la diffusion se fait principalement dans la direction du vecteur propre associé à la plus grande valeur propre.

$$\mathbf{D} \approx \lambda_1 \mathbf{D}_1 = \lambda_1 \vec{e}_1 \vec{e}_1^T \quad (2.18)$$

- cas planaire ($\lambda_1 \approx \lambda_2 \gg \lambda_3$) : la diffusion se fait dans le plan défini par les deux vecteurs propres correspondant aux deux plus grandes valeurs propres.

$$\mathbf{D} \approx \lambda_1 \mathbf{D}_p = \lambda_1 (\vec{e}_1 \vec{e}_1^T + \vec{e}_2 \vec{e}_2^T) \quad (2.19)$$

- cas sphérique ($\lambda_1 \approx \lambda_2 \approx \lambda_3$) : la diffusion est isotrope dans chaque direction de l'espace.

$$\mathbf{D} \approx \lambda_1 \mathbf{D}_s = \lambda_1 (\vec{e}_1 \vec{e}_1^T + \vec{e}_2 \vec{e}_2^T + \vec{e}_3 \vec{e}_3^T) \quad (2.20)$$

Un tenseur de diffusion $\mathbf{D}$ peut être représenté par combinaison de ces trois cas :

$$\begin{aligned} \mathbf{D} &= \lambda_1 \vec{e}_1 \vec{e}_1^T + \lambda_2 \vec{e}_2 \vec{e}_2^T + \lambda_3 \vec{e}_3 \vec{e}_3^T \\ &= (\lambda_1 - \lambda_2) \vec{e}_1 \vec{e}_1^T + (\lambda_2 - \lambda_3) (\vec{e}_1 \vec{e}_1^T + \vec{e}_2 \vec{e}_2^T) + \lambda_3 (\vec{e}_1 \vec{e}_1^T + \vec{e}_2 \vec{e}_2^T + \vec{e}_3 \vec{e}_3^T) \\ &= (\lambda_1 - \lambda_2) \mathbf{D}_1 + (\lambda_2 - \lambda_3) \mathbf{D}_p + \lambda_3 \mathbf{D}_s \end{aligned} \quad (2.21)$$

où $(\lambda_1 - \lambda_2)$, $(\lambda_2 - \lambda_3)$ et λ_3 sont les coordonnées de $\mathbf{D}$ dans la base de tenseurs $[\mathbf{D}_1, \mathbf{D}_p, \mathbf{D}_s]$. Cette relation entre les valeurs propres du tenseur de diffusion permet de classifier le tenseur de diffusion par rapport à sa forme géométrique. En utilisant cette nouvelle décomposition, il est possible de quantifier la similarité aux cas linéaire, planaire et sphérique. Les coefficients obtenus sont normalisés de la manière suivante :

$$\begin{aligned} c_l &= \frac{\lambda_1 - \lambda_2}{\lambda_1} \\ c_p &= \frac{\lambda_2 - \lambda_3}{\lambda_1} \\ c_s &= \frac{\lambda_3}{\lambda_1} \end{aligned} \quad (2.22)$$

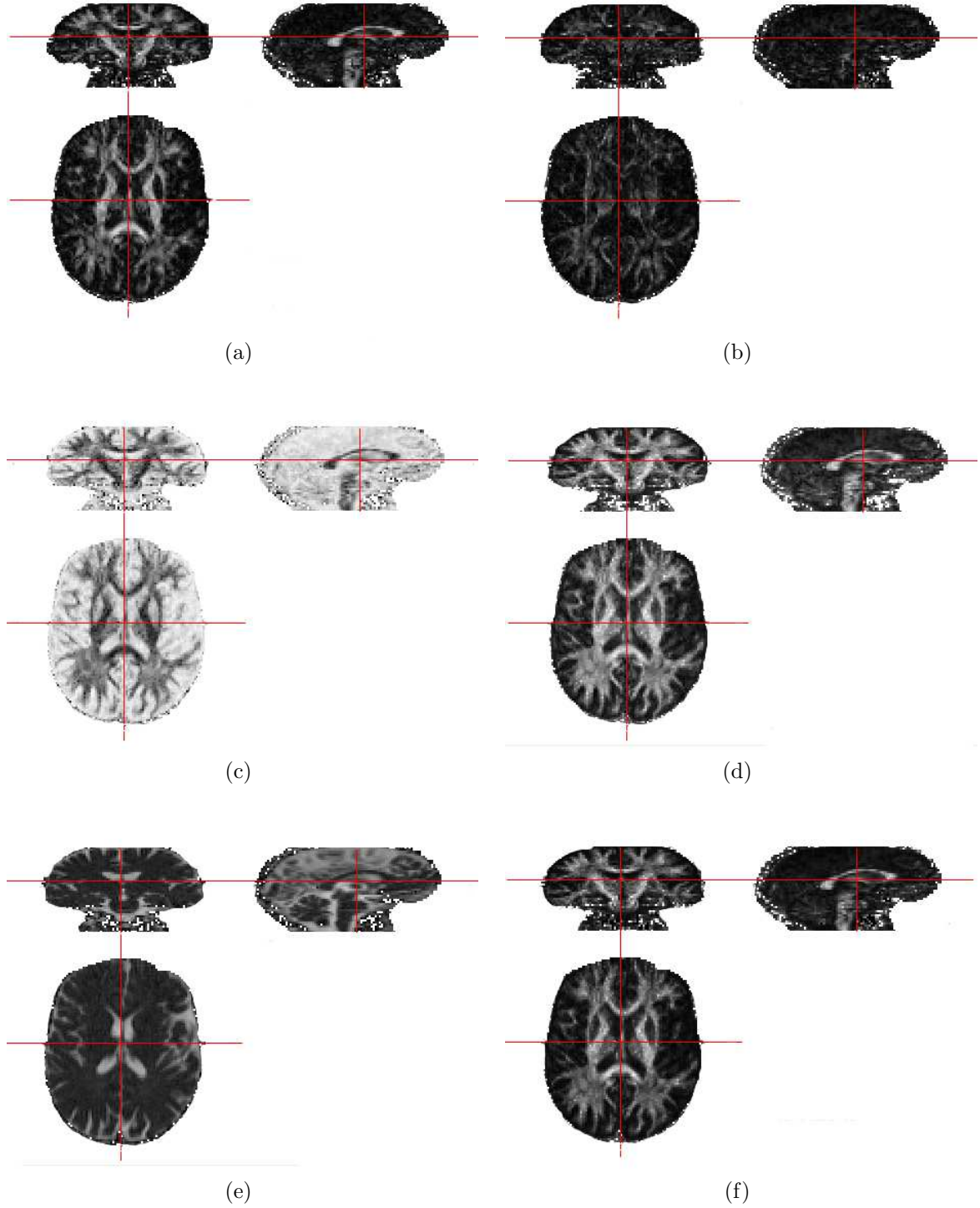


FIG. 2.3 – Modèle d'ordre 2 : visualisation triplanaire des indices (a) c_l , (b) c_p , (c) c_s , (d) c_a , (e) MD et (e) FA calculés pour un même sujet.

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

On a donc $c_l + c_p + c_s = 1$. Ces trois coefficients apportent une information complémentaire en caractérisant la forme des tenseurs. Finalement, l'anisotropie peut aussi être mesurée comme une différence avec le cas sphérique (cf. figure 2.3) :

$$c_a = c_l + c_p = 1 - c_s = 1 - \frac{\lambda_3}{\lambda_1} \quad (2.23)$$

2.1.3 Cas particulier : le tenseur d'ordre 4

Le tenseur d'ordre 2 ne permet pas de modéliser des profils de diffusion hétérogènes avec plusieurs orientations privilégiées. Or, ce type de profil est fréquemment observé au niveau des croisements de faisceaux de fibres. Les tenseurs d'ordre supérieur (en particulier ici d'ordre 4) permettent de s'affranchir de cette limitation. Lorsque $l = 4$, le tenseur de diffusion est représenté par une structure $3 \times 3 \times 3 \times 3$ **D**. Elle contient 15 éléments distincts ($d_{1111}, d_{2222}, d_{3333}, d_{1112}, d_{1113}, d_{2221}, d_{2223}, d_{3331}, d_{3332}, d_{1122}, d_{1133}, d_{2233}, d_{1123}, d_{2213}, d_{3312}$) de multiplicité $\mu = (1, 1, 1, 4, 4, 4, 4, 4, 4, 6, 6, 6, 12, 12, 12)$. Malheureusement, l'algèbre des tenseurs d'ordre 4 est peu développée dans la littérature et quasi inexistante dans les bibliothèques numériques actuelles. Ainsi, l'extension de notions classiques utilisées pour les tenseurs d'ordre 2 telles que la défini-positivité, l'inversion ou la diagonalisation demande réflexion. De plus, il est difficile de se représenter des objets de dimension supérieure à 3. Pour ces raisons, il est usuel d'introduire une matrice **M** de taille 6×6 équivalente au tenseur d'ordre 4 [36, 24] :

$$\mathbf{M} = \begin{pmatrix} d_{1111} & d_{1122} & d_{1133} & \sqrt{2}d_{1112} & \sqrt{2}d_{1113} & \sqrt{2}d_{1123} \\ * & d_{2222} & d_{2233} & \sqrt{2}d_{2221} & \sqrt{2}d_{2213} & \sqrt{2}d_{2223} \\ * & * & d_{3333} & \sqrt{2}d_{3312} & \sqrt{2}d_{3331} & \sqrt{2}d_{3332} \\ * & * & * & 2d_{1122} & 2d_{1123} & 2d_{2213} \\ * & * & * & * & 2d_{1133} & 2d_{3312} \\ * & * & * & * & * & 2d_{2233} \end{pmatrix} \quad (2.24)$$

où les $*$ représentent les éléments symétriques. Notons qu'il y a 21 termes au lieu des 15 éléments distincts du tenseur d'ordre 4. Ceci vient du fait que certains éléments de symétrie n'apparaissent pas (par exemple d_{1122} et d_{2211}).

L'extraction des maxima d'un tenseur d'ordre 4

D'après le théorème spectral, il est possible de calculer les 6 valeurs propres de la matrice **M**. Celles-ci n'ont aucune interprétation physique (notamment en terme de maximum du profil de diffusion $d(\vec{g})$). Elles sont donc difficilement utilisables, par exemple, pour mesurer l'anisotropie du tenseur. Nous présentons dans cette partie trois méthodes différentes d'extraction des maxima pour un tenseur d'ordre 4, cette opération étant nécessaire pour déterminer les directions privilégiées de diffusion correspondant à l'orientations locales des faisceaux de fibres :

- Résolution analytique : Dans [29, 31, 30], Qi *et coll.* définissent des D-(valeurs propres) correspondant aux points critiques (maxima, minima et points selles) de la fonction

de diffusion. Ces points critiques vérifient les équations suivantes pour certains $\lambda \in \mathbb{R}$ [31] :

$$\begin{cases} \mathbf{D}\vec{g}^3 = \lambda\vec{g} \\ \|\vec{g}\| = 1 \end{cases} \quad (2.25)$$

où :

$$(\mathbf{D}\vec{g}^3)_i = \sum_{j,k,l=1}^3 d_{ijkl} g_j g_k g_l \quad (2.26)$$

Les paramètres $(\lambda, \vec{g})$ vérifiant (2.25) sont respectivement une D-(valeur propre) et un D-(vecteur propre). Ces paramètres vérifient également la relation suivante :

$$\lambda = \mathbf{D}\vec{g}^4 \quad (2.27)$$

Toutes les solutions $(\lambda, \vec{g})$ vérifient le système d'équations à quatre inconnues défini par l'équation (2.25). Ce système peut être réécrit comme un problème à deux inconnues [31]. En pratique, les solutions ont trois formes caractéristiques, $\vec{g} = \pm[1, 0, 0]$, $\vec{g} = \pm[t, 1, 0]/\|[t, 1, 0]\|$ et $\vec{g} = \pm[u, v, 1]/\|[u, v, 1]\|$, avec t solution d'une équation de degré quatre et (u, v) solution d'un système de deux équations.

- Résolution différentielle : On peut également essayer de trouver directement les points critiques de $d(\vec{g})$, $\vec{g} = [g_1, g_2, g_3]$ sur la sphère [14]. Le problème d'optimisation sous contraintes (*i.e.*, $\|\vec{g}\| = 1$) peut être transformé en un problème d'optimisation sans contrainte dans un espace de dimension supérieure en utilisant les multiplicateurs de Lagrange. Soit $F(\vec{g}, \lambda) = d(\vec{g}) - \lambda(\|\vec{g}\|^2 - 1)$. Les points critiques $\vec{g}^*$ et les multiplicateurs associés λ^* doivent annuler le gradient de $F(\vec{g}^*, \lambda^*)$:

$$\frac{\partial F(\vec{g}^*, \lambda^*)}{\partial g_1} = \frac{\partial F(\vec{g}^*, \lambda^*)}{\partial g_2} = \frac{\partial F(\vec{g}^*, \lambda^*)}{\partial g_3} = \frac{\partial F(\vec{g}^*, \lambda^*)}{\partial \lambda} = 0 \quad (2.28)$$

- Recherche exhaustive : On peut également faire une recherche exhaustive des maxima de la fonction $d(\vec{g})$ sur la sphère (notée S^3). Dès lors, $d(\vec{g})$ admet un maximum local au point courant $\vec{p} \in S^3$ si il existe un voisinage V tel que pour tout $\vec{x}$ de V , on a : $d(\vec{x}) \leq d(\vec{p})$.

Alors qu'avec la recherche exhaustive, les maxima d'un tenseur d'ordre 4 sont directement extraits, les deux autres techniques donnent les points critiques à partir desquels il faut extraire les maxima. Pour ce faire, un critère sur le hessien encadré est considéré [16]. En effet, l'analyse du signe du hessien ne suffit pas pour prendre une décision dans le cas d'une optimisation contrainte. Le hessien est encadré par les dérivées premières de la contrainte :

$$\bar{\mathbf{H}}_{m+n}[F(\vec{g}, \lambda)] = \begin{pmatrix} \mathbf{0}_{m \times m} & \mathbf{B}_{m \times n} \\ \mathbf{B}_{m \times n}^T & \mathbf{H}[F(\vec{g}, \lambda)]_{n \times n} \end{pmatrix} \quad (2.29)$$

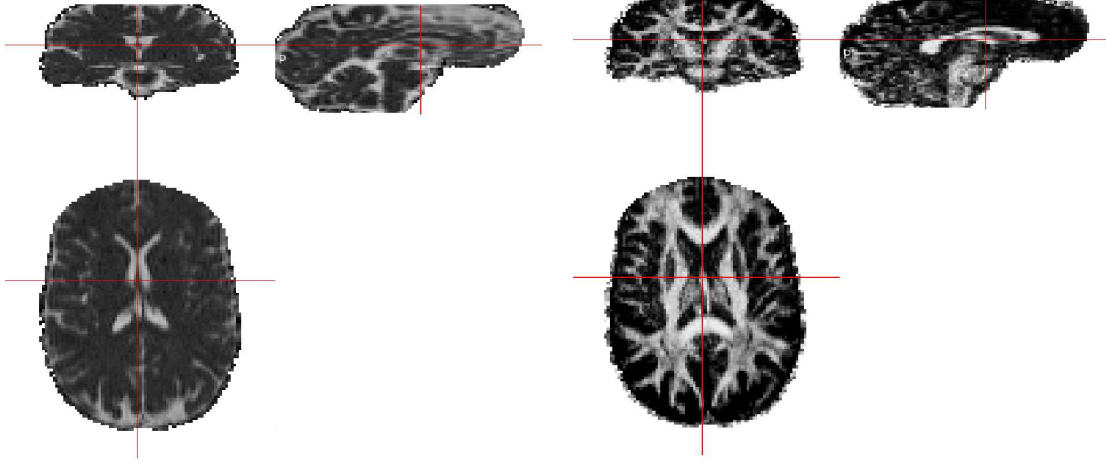


FIG. 2.4 – Modèle d'ordre 4 : visualisation tri-planaire d'une image de diffusion moyenne (MD, à gauche) et de la même image d'anisotropie généralisée (GA, à droite).

où $\mathbf{B}_{m \times n}$ est la matrice composée de la dérivée de $\|\vec{g}\|^2 - 1$ suivant les 3 variables de l'espace. Dans notre cas, $n = 3$ et le nombre de dimension du problème est $m = 1$ (le nombre de contraintes). On obtient finalement un critère sur le signe de $\bar{\mathbf{H}}$:

$$\begin{aligned} (-1)^m |\bar{\mathbf{H}}_m + r[F(\vec{g}, \lambda)]| &> 0 \quad \text{minimum strict} \\ (-1)^r |\bar{\mathbf{H}}_m + r[F(\vec{g}, \lambda)]| &> 0 \quad \text{maximum strict} \end{aligned} \quad (2.30)$$

$\forall r = m + 1, \dots, n$

où $|\cdot|$ représente le déterminant.

Les indices scalaires caractérisant un tenseur d'ordre 4

Comme pour les tenseurs d'ordre 2, des indices scalaires ont été proposés pour résumer l'information contenue dans les 15 paramètres indépendants du tenseurs d'ordre 4. Les deux indices importants sont la carte d'ADC moyenne (ou diffusion moyenne) et une carte mesurant l'anisotropie en chaque voxel de l'image. La carte d'ADC moyenne est obtenue en considérant l'expression générale suivante : $MD = \langle D \rangle = \frac{1}{4\pi} \int_{S_3} d(\vec{g}) d\vec{g}$ avec S_3 la sphère unité. La MD d'un tenseur d'ordre 4 est alors obtenue en considérant un cinquième de la trace de ce tenseur dans sa représentation 6×6 , $MD = \text{trace}(\mathbf{M})/5 = (d_{1111} + d_{2222} + d_{3333} + 2(d_{1122} + d_{1133} + d_{2233}))/5$ [27] (cf. figure 2.4). Comme expliqué précédemment, il n'est pas possible d'utiliser les valeurs propres de $\mathbf{M}$ pour définir une mesure d'anisotropie, ces dernières n'ayant pas d'interprétation physique au regard de l'organisation des micro-structures locales. Dans [27], l'anisotropie est exprimée comme la variance des valeurs de diffusion normalisées $D_N(\vec{g})$, où la normalisation se fait via l'expression :

$$D_N(\vec{g}) = \frac{d(\vec{g})}{\text{trg}(d(\vec{g}))} \quad (2.31)$$

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

où trg est la trace généralisée : $trg(d(\vec{g})) = \frac{3}{4\pi} \int_{S_3} d(\vec{g}) d\vec{g} = 3 < D > [27]$. On peut ensuite écrire la variance V du signal normalisé [27] :

$$\begin{aligned} V &= < D_N(\vec{g})^2 > - < D_N(\vec{g}) >^2 \\ &= \frac{1}{3} trg(D_N(\vec{g})^2) - \frac{1}{9} \\ &= \frac{1}{9} \left(\frac{dist(\mathbf{D}, \mathbf{0})}{< \mathbf{D} >^2} - 1 \right) \end{aligned} \quad (2.32)$$

où $< \mathbf{D} >$ représente la diffusion moyenne et avec [3] :

$$\begin{aligned} dist(\mathbf{D}_1, \mathbf{D}_2)^2 &= \frac{1}{4\pi} \int_{S_3} (d_1(\vec{g}) - d_2(\vec{g}))^2 d\vec{g} \\ &= d(\Delta_{400} + \Delta_{040} + \Delta_{004} + \Delta_{220} + \Delta_{022} + \Delta_{202})^2 \\ &\quad + d(\Delta_{211} + \Delta_{031} + \Delta_{013})^2 + d(\Delta_{121} + \Delta_{301} + \Delta_{103})^2 \\ &\quad + d(\Delta_{112} + \Delta_{310} + \Delta_{130})^2 + (b - d)(\Delta_{400} + \Delta_{040} + \Delta_{004})^2 \\ &\quad + (b - d)((\Delta_{310} + \Delta_{130})^2 + (\Delta_{301} + \Delta_{103})^2 + (\Delta_{031} + \Delta_{013})^2) \\ &\quad + (c - d)((\Delta_{400} + \Delta_{220})^2 + (\Delta_{400} + \Delta_{202})^2) \\ &\quad + (c - d)((\Delta_{040} + \Delta_{220})^2 + (\Delta_{400} + \Delta_{022})^2) \\ &\quad + (c - d)((\Delta_{004} + \Delta_{022})^2 + (\Delta_{004} + \Delta_{202})^2) \\ &\quad + (c - b)(\Delta_{310}^2 + \Delta_{301}^2 + \Delta_{130}^2 + \Delta_{031}^2 + \Delta_{103}^2 + \Delta_{013}^2) \\ &\quad + (a - b - 2(c - d))(\Delta_{400}^2 + \Delta_{040}^2 + \Delta_{004}^2) \\ &\quad + (b - d - 2(c - d))(\Delta_{220}^2 + \Delta_{022}^2 + \Delta_{202}^2) \end{aligned} \quad (2.33)$$

où $\mathbf{D}_1$ et $\mathbf{D}_2$ sont deux tenseurs d'ordre 4, $a = 1/9$, $b = 1/105$, $c = 1/63$, $d = 1/315$ et $\Delta_{i,j,k} = d_{2;1^i 2^j 3^k} - d_{1;1^i 2^j 3^k}$ avec i, j et k le nombre d'indice 1, 2 et 3 définissant la composante du tenseur. Il a été montré dans [27] que pour un tenseur d'ordre l , la borne supérieure de la variance V est donnée par :

$$sup(V) = \frac{l^2}{9(2l + 1)} \quad (2.34)$$

Ainsi, $\lim_{l \rightarrow +\infty} sup(V) = +\infty$. Il est alors d'usage d'appliquer une fonction monotone de $V : [0, +\infty[$ dans $[0, 1[$ [27]. L'Anisotropie Généralisée (GA) s'écrit alors (cf. figure 2.4) :

$$GA = 1 - \frac{1}{1 + (k_1 V)^{\epsilon(V)}} \quad (2.35)$$

avec :

$$\epsilon(V) = 1 + \frac{1}{1 + k_2 V} \quad (2.36)$$

où k_1 et k_2 sont deux coefficients d'échelle. Pour un tenseur d'ordre 4, les paramètres optimaux valent 250 et 5000 respectivement [27]. Cet indice peut également être calculé pour des

tenseurs d'ordre 2 ou plus généralement pour tout tenseur de diffusion, mais son obtention reste empirique puisqu'il faut fixer les valeurs de k_1 et k_2 .

2.2 L'estimation des paramètres du modèle en ITD

Nous venons d'introduire le modèle du tenseur de diffusion défini pour différents ordres l . Dans cette partie, nous cherchons à estimer les paramètres de ces modèles à partir des observations, ces paramètres modélisant soit le signal de diffusion, soit la diffusivité, soit le déplacement des molécules d'eau. La contrainte de positivité des tenseurs de diffusion sera également prise en compte lors de cette estimation. En effet, appliquer la fonction $\log$ à l'équation de Stejskal-Tanner (1.26) permet d'obtenir un système linéaire autorisant une détermination analytique des paramètres du tenseur de diffusion $\mathbf{D}$. Cette méthode donne une solution dans l'espace des matrices symétriques, rien n'assurant qu'elle soit définie positive. Par exemple, à l'ordre 2, le bruit ou les artefacts présents dans les images de diffusion peuvent entraîner l'estimation de tenseurs ayant des valeurs propres négatives ou nulles, ce qui n'est physiquement pas acceptable. Pour cette raison, des méthodes d'estimation plus avancées permettent de garantir la positivité du tenseur [4, 22, 13, 33, 20]. Dans [4] une paramétrisation élégante des tenseurs positifs est utilisée lors de l'estimation. Cette méthode sera présentée pour estimer des matrices symétriques, définies positives.

2.2.1 La méthode des moindres carrés : modélisation de la diffusivité

Pour simplifier les notations, $S_i(TE)$ (le signal de diffusion dans la direction $\vec{g}_i$) est noté S_i et $S(TE)_{g=0}$ (le signal en l'absence de gradient) est noté S_0 . En passant au logarithme, l'équation de Stejskal-Tanner s'écrit :

$$-\frac{1}{b} \ln \left(\frac{S_i}{S_0} \right) = \vec{g}^T D \vec{g} \quad (2.37)$$

Pour un tenseur d'ordre 2, l'équation (2.37) peut s'exprimer sous la forme d'un système d'équation linéaire en fonction des paramètres du tenseur de la forme suivante :

$$\vec{y} = \begin{pmatrix} -\frac{1}{b} \ln \left(\frac{S_1}{S_0} \right) \\ \vdots \\ -\frac{1}{b} \ln \left(\frac{S_i}{S_0} \right) \\ \vdots \\ -\frac{1}{b} \ln \left(\frac{S_N}{S_0} \right) \end{pmatrix} = \begin{pmatrix} g_{1;1}^2 & g_{1;2}^2 & g_{1;3}^2 & 2g_{1;1}g_{1;2} & 2g_{1;1}g_{1;3} & 2g_{1;2}g_{1;3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{i;1}^2 & g_{i;2}^2 & g_{i;3}^2 & 2g_{i;1}g_{i;2} & 2g_{i;1}g_{i;3} & 2g_{i;2}g_{i;3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{N;1}^2 & g_{N;2}^2 & g_{N;3}^2 & 2g_{N;1}g_{N;2} & 2g_{N;1}g_{N;3} & 2g_{N;2}g_{N;3} \end{pmatrix} \begin{pmatrix} d_{11} \\ d_{22} \\ d_{33} \\ d_{12} \\ d_{13} \\ d_{23} \end{pmatrix} \quad (2.38)$$

où N représente le nombre de mesures et $g_{i;j}$ la j ème composante de la i ème direction de gradient $\vec{g}_i$. De la même manière, pour un tenseur d'ordre 4, le système linéaire à résoudre prend la forme (en considérant le logarithme de l'équation (2.1)) :

$$\vec{y} = \begin{pmatrix} g_{1;1}^4 & g_{1;2}^4 & g_{1;3}^4 & 4g_{1;1}^3g_{1;2} & \dots & 12g_{1;2}^2g_{1;1}g_{1;3} & 12g_{1;3}^2g_{1;1}g_{1;2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{i;1}^4 & g_{i;2}^4 & g_{i;3}^4 & 4g_{i;1}^3g_{i;2} & \dots & 12g_{i;2}^2g_{i;1}g_{i;3} & 12g_{i;3}^2g_{i;1}g_{i;2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{N;1}^4 & g_{N;2}^4 & g_{N;3}^4 & 4g_{N;1}^3g_{N;2} & \dots & 12g_{N;2}^2g_{N;1}g_{N;3} & 12g_{N;3}^2g_{N;1}g_{N;2} \end{pmatrix} \begin{pmatrix} d_{1111} \\ d_{2222} \\ d_{3333} \\ d_{1112} \\ d_{1113} \\ d_{2221} \\ d_{2223} \\ d_{3331} \\ d_{3332} \\ d_{1122} \\ d_{1133} \\ d_{2233} \\ d_{1123} \\ d_{2213} \\ d_{3312} \end{pmatrix} \quad (2.39)$$

Finalement, l'écriture du système linéaire à résoudre peut se simplifier en $\vec{y} = \mathbf{A}\vec{x}$. Pour l'estimation des 6 paramètres du tenseur d'ordre 2, il faudra que $N \geq 6$, et de même, pour l'estimation des 15 paramètres du tenseur d'ordre 4, il faudra que $N \geq 15$. On résout ce système au sens des moindres carrés : on cherche $\vec{x}$ (les éléments indépendants du tenseur $\mathbf{D}$) minimisant l'erreur commise pendant l'estimation $\|\vec{r}\|^2 = \|\vec{y} - \mathbf{A}\vec{x}\|^2$, où chaque ligne de $\mathbf{A}$ contient les N_l monômes (2.6) du tenseur de diffusion. La solution s'écrit [18] :

$$\vec{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \vec{y} \quad (2.40)$$

Cette estimée minimise l'erreur quadratique entre la mesure et la prédiction. Il a été montré que la solution aux moindres carrés est le meilleur estimateur au sens du maximum de vraisemblance pour un bruit additif gaussien sur les observations $\vec{y}$ (dans notre cas, un bruit gaussien sur les ADC). Ce modèle de bruit ne correspond pas au bruit d'acquisition qui est généralement supposé Ricien sur le signal. De plus, aucune précaution n'est prise pour obtenir un tenseur défini positif. Pour dépasser ces limitations, il faut résoudre des systèmes plus complexes, bien souvent non-linéaires.

2.2.2 La méthode des quartiques : modélisation de la diffusivité

Lors de l'estimation des tenseurs de diffusion, il est souhaitable d'imposer une contrainte de positivité sur les tenseurs. La paramétrisation des tenseurs d'ordre pairs définis positifs présentée dans [4] permet de prendre en compte cette contrainte de positivité.

Paramétrisation des tenseurs de diffusion

Un tenseur de diffusion est représenté par un profil de diffusion pouvant être vu comme un polynôme positif sur la sphère unité (2.2). Tout polynôme positif peut s'écrire comme une somme de polynômes d'ordres inférieurs élevés au carré [8, 32]. Ainsi, en imagerie de

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

diffusion, tout polynôme homogène positif d'ordre l pair de 3 variables peut s'écrire comme la somme au carré de polynômes homogènes $p(g_1, g_2, g_3; \vec{c})$ d'ordre $l/2$, où $\vec{c}$ est un vecteur contenant les coefficients du polynôme :

$$d(\vec{g}, \vec{c}_1, \dots, \vec{c}_M) = \sum_{i=1}^M p(g_1, g_2, g_3; \vec{c}_i)^2 \quad (2.41)$$

Le nombre de termes dans la sommation a comme borne supérieure le nombre de monômes uniques d'un polynôme homogène d'ordre $l/2$ à 3 variables, M :

$$M \leq \frac{(2 + l/2)!}{2(l/2)!}$$

Par exemple, un tenseur d'ordre 2 peut s'écrire comme la somme de $M \leq 3$ polynômes d'ordre 1 au carré. De même, un tenseur d'ordre 4 pourra s'écrire comme une somme de $M \leq 6$ polynômes d'ordre 2 au carré. Ceci est en accord avec le théorème d'Hilbert qui énonce que $M = 3$ polynômes sont suffisants dans l'équation (2.41) pour un polynôme d'ordre 4 [15]. La paramétrisation dans l'équation (2.41) permet de générer tout l'espace des polynômes homogènes positifs avec pour seule contrainte que l'ordre du polynôme soit pair. Les inconnues sont les coefficients $\vec{c}_i$ des polynômes $p(\cdot)$ dont le nombre est borné par :

$$\frac{1}{8} \left(\frac{(2 + l/2)!}{(l/2)!} \right)^2 + \frac{(2 + l/2)!}{4(l/2)!},$$

soient 6 et 21 coefficients pour $l = 2$ et $l = 4$.

Formulation du problème d'optimisation

L'estimation d'un tenseur de diffusion défini positif se fait en minimisant la fonction d'énergie suivante par rapport aux coefficients polynômiaux [4] :

$$E(\vec{c}_1, \dots, \vec{c}_M) = \sum_{i=1}^N (S_i/S_0 - \exp(-bd(\vec{g}_i, \vec{c}_1, \dots, \vec{c}_M)))^2 \quad (2.42)$$

où $d(\vec{g}, \vec{c}_1, \dots, \vec{c}_M)$ est donné par (2.41). Le problème d'optimisation ainsi posé est non contraint, mais également non linéaire en fonction des paramètres $(\vec{c}_1, \dots, \vec{c}_M)$. Il est possible de se ramener à un problème linéaire avec une contrainte de positivité sur les paramètres à estimer en réécrivant $d(\vec{g}, \vec{c}_1, \dots, \vec{c}_M)$ comme une convolution sphérique [4] :

$$d(\vec{g}, \lambda(\cdot)) = \int_{S_{(M-1)}} \lambda(\vec{c}) p(g_1, g_2, g_3; \vec{c})^2 d\vec{c} \quad (2.43)$$

où $\lambda(\vec{c})$ est une fonction réelle non-négative. Ainsi, le profil de diffusion $d(\vec{g})$ n'est plus paramétré par l'ensemble des coefficients polynomiaux $(\vec{c}_1, \dots, \vec{c}_M)$, mais par la fonction $\lambda(\vec{c})$. Dès lors, en injectant (2.43) dans (2.42), le problème d'optimisation revient à déterminer la fonction $\lambda(\vec{c})$ qui minimise E . Pour permettre une résolution numérique de ce problème,

on approche l'intégrale (2.43) sur la sphère en considérant un jeu de M' vecteurs unitaires $(\vec{c}_1, \dots, \vec{c}_{M'})$ [4] :

$$d(\vec{g}, \lambda_1, \dots, \lambda_{M'}) = \sum_{i=1}^{M'} \lambda_i p(g_1, g_2, g_3; \vec{c}_i)^2 \quad (2.44)$$

où les λ_i sont des nombres réels positifs. L'équation (2.44) approxime l'espace des tenseurs d'ordre l positifs avec une précision contrôlée par M' .

Estimation des paramètres

Les tenseurs de diffusion définis positifs peuvent être calculés en se basant sur la relation (2.37). Cela revient à résoudre un système linéaire contraint de telle sorte que les estimées ne contiennent aucun élément négatif. Le système est de la forme $\vec{y} = \mathbf{P}\vec{x}$:

$$\begin{pmatrix} -\frac{1}{b} \ln \left(\frac{S_1}{S_0} \right) \\ \vdots \\ -\frac{1}{b} \ln \left(\frac{S_i}{S_0} \right) \\ \vdots \\ -\frac{1}{b} \ln \left(\frac{S_N}{S_0} \right) \end{pmatrix} = \begin{pmatrix} p(\vec{g}_1; \vec{c}_1)^2 & \dots & p(\vec{g}_1; \vec{c}_{M'})^2 \\ \vdots & \vdots & \vdots \\ \dots & p(\vec{g}_i; \vec{c}_j)^2 & \dots \\ \vdots & \vdots & \vdots \\ p(\vec{g}_N; \vec{c}_1)^2 & \dots & p(\vec{g}_N; \vec{c}_{M'})^2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{M'} \end{pmatrix} \quad (2.45)$$

$\mathbf{P}$ peut s'écrire comme le produit d'une matrice $\mathbf{G}$ de taille $N \times N_l$ contenant les monômes construits à partir des directions d'acquisition $\vec{g}_i$:

$$\mathbf{G} = \begin{pmatrix} (g_{1;1})^{n_1^l(1)} (g_{1;2})^{n_2^l(1)} (g_{1;3})^{n_3^l(1)} & \dots & (g_{1;1})^{n_1^l(N_l)} (g_{1;2})^{n_2^l(N_l)} (g_{1;3})^{n_3^l(N_l)} \\ \vdots & & \vdots \\ (g_{N;1})^{n_1^l(1)} (g_{N;2})^{n_2^l(1)} (g_{N;3})^{n_3^l(1)} & \dots & (g_{N;1})^{n_1^l(N_l)} (g_{N;2})^{n_2^l(N_l)} (g_{N;3})^{n_3^l(N_l)} \end{pmatrix} \quad (2.46)$$

et d'une matrice $\mathbf{C}$ de taille $N_l \times M'$ contenant seulement les monômes construits avec les éléments des vecteurs $\vec{c}_j$:

$$\mathbf{C} = \begin{pmatrix} \mu_1^l(c_{1;1})^{n_1^l(1)} (c_{1;2})^{n_2^l(1)} (c_{1;3})^{n_3^l(1)} & \dots & \mu_{N_l}^l(c_{1;1})^{n_1^l(N_l)} (c_{1;2})^{n_2^l(N_l)} (c_{1;3})^{n_3^l(N_l)} \\ \vdots & & \vdots \\ \mu_1^l(c_{M';1})^{n_1^l(1)} (c_{M';2})^{n_2^l(1)} (c_{M';3})^{n_3^l(1)} & \dots & \mu_{N_l}^l(c_{M';1})^{n_1^l(N_l)} (c_{M';2})^{n_2^l(N_l)} (c_{M';3})^{n_3^l(N_l)} \end{pmatrix}^T \quad (2.47)$$

Pour résoudre ce système en imposant que tous les λ_j soient positifs, un algorithme NNLS (*Non-Negative Least Square*) est mis en œuvre [21]. La solution obtenue contient au plus $\text{rang}(\mathbf{P})$ éléments non nuls. Il est intéressant de remarquer qu'un tenseur positif peut être estimé à partir des équations (2.38) et (2.39) en définissant un problème d'optimisation sous contraintes. Dans ce cas, la contrainte n'est pas triviale (il faut estimer des valeurs et les contraindre pour obtenir un tenseur symétrique défini positif). Ici, après reformulation du problème d'optimisation, il suffit d'imposer une simple contrainte de positivité directement

sur les paramètres estimés. Les N_l coefficients du tenseur d'ordre l d_k ($k \in [1, \dots, N_l]$) sont donnés par la relation $\mathbf{C}\vec{x}$.

2.2.3 Modélisation de la probabilité de déplacement : la fonction de distribution des orientations

Introduction

A l'intérieur d'un voxel, les fibres peuvent avoir des orientations hétérogènes. Dans un premier temps, nous montrons que la modélisation de la diffusivité n'est pas le reflet direct de l'organisation des structures sous-jacentes (*i.e.*, la direction de diffusivité maximale ne correspond pas nécessairement avec l'orientation des faisceaux de fibres). Pour ce faire, considérons deux faisceaux orthogonaux a et b qui se croisent, chacun d'entre eux occupant une fraction f et $(1 - f)$ du volume du voxel. Soient $\mathbf{D}_a$ et $\mathbf{D}_b$ les tenseurs de diffusion correspondant à chacun des faisceaux et $P_a(\vec{R}, \tau)$ et $P_b(\vec{R}, \tau)$ les propagateurs associés. La probabilité, d'une molécule d'eau choisie aléatoirement de parcourir une distance $\vec{R}$ pendant un temps τ s'écrit [23] :

$$P(\vec{R}, \tau) = fP_a(\vec{R}, \tau) + (1 - f)P_b(\vec{R}, \tau) \quad (2.48)$$

Si on fait l'hypothèse que les signaux s'ajoutent indépendamment, on obtient [23] :

$$S(\vec{g}, b) = f \exp(-b\vec{g}^T \mathbf{D}_a \vec{g}) + (1 - f) \exp(-b\vec{g}^T \mathbf{D}_b \vec{g}) = \exp(-bd(\vec{g})) \quad (2.49)$$

Et donc :

$$d(\vec{g}) = -\frac{1}{b} \ln [f \exp(-b\vec{g}^T \mathbf{D}_a \vec{g}) + (1 - f) \exp(-b\vec{g}^T \mathbf{D}_b \vec{g})] \quad (2.50)$$

La figure 2.5-A présente la probabilité de déplacement $P(\vec{R}, \tau)$ pour $\|\vec{R}\| = \text{constant}$. La figure 2.5-B donne la mesure de diffusivité $d(\vec{g})$ en fonction de la direction. On remarque que les maxima de la probabilité de déplacement sont alignés avec les directions des faisceaux de fibres passant dans le voxel alors que la diffusivité ne l'est pas. En effet, les diffusivités des deux compartiments sont combinées en appliquant une fonction non linéaire (2.50), alors que les probabilités de déplacement sont combinées en appliquant une fonction linéaire (2.48).

Jusqu'à présent, les méthodes présentées estimaient les paramètres d'un modèle de la diffusivité. Pour les raisons évoquées précédemment, il est essentiel d'introduire une modélisation du propagateur dont les maxima sont alignés avec les directions des faisceaux de fibres de la substance blanche : la fonction de distribution des orientations (ou, en anglais, *Orientation Distribution Function* : ODF).

La Fonction de Distribution des Orientations (ODF)

Nous venons de voir que les maxima d'un tenseur d'ordre 4, modélisant la diffusivité, ne sont pas alignés avec les directions de diffusion des molécules d'eau dans le milieu, d'où la nécessité d'introduire une nouvelle modélisation. Il est possible de modéliser le signal contenu

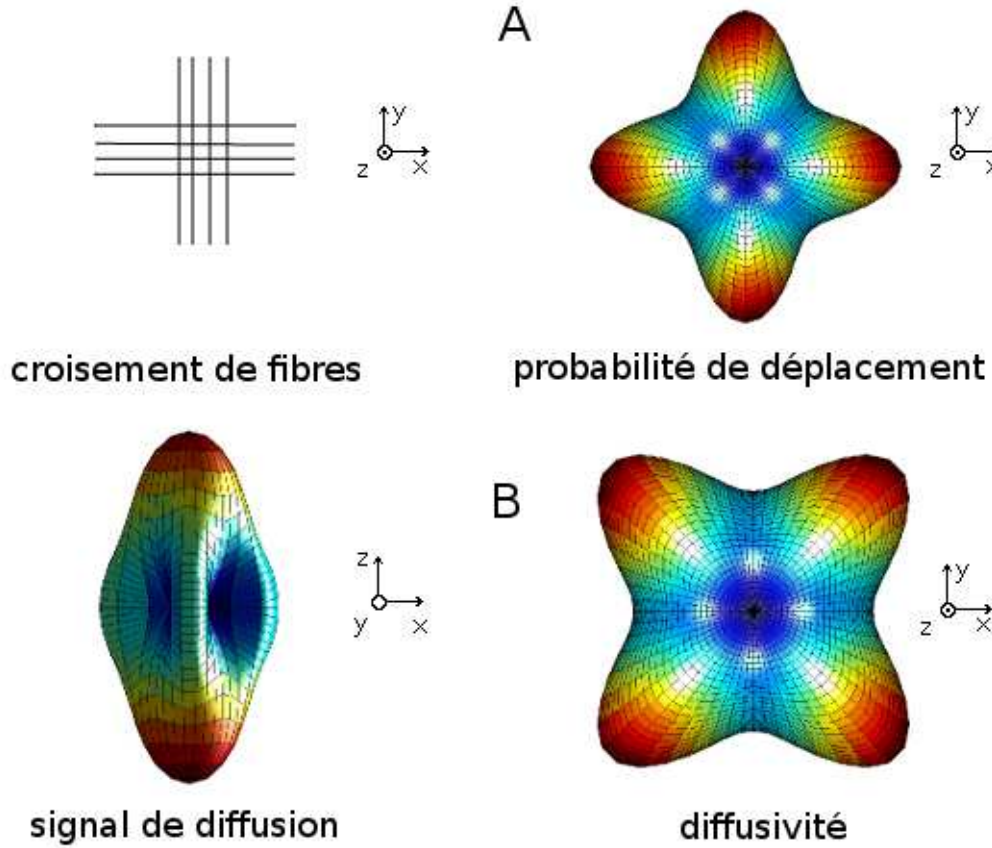


FIG. 2.5 – Visualisation de la probabilité de déplacement, du signal de diffusion (cette image a été tournée de 90° par rapport à la configuration du croisement de fibres) et de la diffusivité pour un croisement de fibre à 90 degrés.

dans chaque voxel de différentes façons. Le propagateur peut être calculé en remarquant que le signal mesuré est relié au propagateur par une transformée de Fourier notée $\mathcal{F}$ [39] :

$$\frac{S_i}{S_0} = \mathcal{F}[P(r\vec{g}_i, \tau)] \quad (2.51)$$

Intuitivement, on remarque qu'il faut échantillonner S le long d'un grand nombre de directions $\vec{g}_i$ pour différentes valeurs de b . Cette technique requiert donc un temps d'acquisition très long. Une alternative consiste à calculer la Fonction de Distribution des Orientations (ODF). Deux méthodes permettent de retrouver cette information. La première considère la projection radiale du propagateur P [38] :

$$\psi(\vec{u}) = \int_0^\infty P(r\vec{u}, \tau) dr \quad (2.52)$$

où $\vec{u}$ est un vecteur unité. On parle alors d'ODF de diffusion. La deuxième méthode considère le signal observé comme une convolution sphérique de l'ODF ψ avec la réponse impulsionnelle d'une seule fibre R [37] :

$$S(\theta, \phi) = \psi(\theta, \phi) \otimes R(\theta) \quad (2.53)$$

où θ est l'élévation et ϕ l'azimuthe en coordonnées sphériques. On parle alors d'ODF de fibres.

Extension de la méthode des quartiques

La méthode des quartiques permet d'estimer des tenseurs définis positifs modélisant la diffusivité. L'extension de cette méthode permet d'estimer une ODF de fibres également définie positive. Une première approche estime l'ODF de fibre à partir du signal de diffusion. Ce dernier est alors modélisé en utilisant une reconstruction multi-fibres [40] :

$$\frac{S(\vec{g}, b)}{S_0} = \int_{S_3} \sum_{j=1}^{M'} \underbrace{\lambda_j p(v_1, v_2, v_3; \vec{c}_j)^2}_{w(\vec{v})} B(\vec{v}, \vec{g}, b) d\vec{v} \quad (2.54)$$

où S_3 représente la sphère unité, v_1, v_2 et v_3 les trois composantes du vecteur unitaire $\vec{v}$, $B(\vec{v}, \vec{g}, b)$ une fonction de base, et $w(\vec{v})$ est une fonction sphérique non-négative assimilable à une fonction de pondération du mélange. On estime donc une ODF de fibre (2.53) avec $B(\cdot)$ représentant la réponse d'une fibre et $w(\vec{v})$ représentant $\psi(\theta, \phi)$ (on choisit donc de modéliser ψ avec des quartiques pour garantir la positivité). La fonction d'énergie à minimiser lors de l'estimation s'écrit alors :

$$E(\lambda_1, \dots, \lambda_{M'}) = \sum_{i=1}^N \left(S_i/S_0 - \sum_{j=1}^{M'} \int_{S_3} \lambda_j p(v_1, v_2, v_3; \vec{c}_j)^2 B(\vec{v}, \vec{g}_i, b) d\vec{v} \right)^2 \quad (2.55)$$

Dans [40], un noyau gaussien est proposé pour le choix de $B(\vec{v}, \vec{g}_i, b)$ afin de refléter l'atténuation du signal induit par la réponse d'une fibre unique orientée dans la direction $\vec{v}$ (ce noyau modélise donc le processus de diffusion) :

$$B(\vec{v}, \vec{g}, b) = \exp(-\delta(\vec{v}^T \vec{g})^2) \quad (2.56)$$

où δ est fixé à 200 et n'influe que très peu le résultat [40]. Cette modélisation donne un maximum (respectivement un minimum) du signal de diffusion pour un gradient dans une direction perpendiculaire (respectivement parallèle) à l'orientation des fibres sous-jacente $\vec{v}$, ce qui est en accord avec l'acquisition. Finalement, l'estimation se fait en résolvant le système de la forme $\vec{y} = \mathbf{B}\vec{x}$ en utilisant un algorithme NNLS pour garantir la positivité des paramètres estimés. La matrice $\mathbf{B}$ est de taille $N \times M'$ et de terme général $b_{i,j}(\vec{g}_i, \vec{c}_j) = \int_{S_2} p(v_1, v_2, v_3; c_j)^2 B(\vec{v}, \vec{g}_i, b) d\vec{v}$. Le système s'écrit :

$$\begin{pmatrix} \frac{S_1}{S_0} \\ \vdots \\ \frac{S_i}{S_0} \\ \vdots \\ \frac{S_N}{S_0} \end{pmatrix} = \begin{pmatrix} b_{1,1}(\vec{g}_1; \vec{c}_1) & \dots & b_{1,M'}(\vec{g}_1; \vec{c}_{M'}) \\ \vdots & \vdots & \vdots \\ \dots & b_{i,j}(\vec{g}_i; \vec{c}_j) & \dots \\ \vdots & \vdots & \vdots \\ b_{N,1}(\vec{g}_N; \vec{c}_1) & \dots & b_{N,M'}(\vec{g}_N; \vec{c}_{M'}) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{M'} \end{pmatrix} \quad (2.57)$$

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

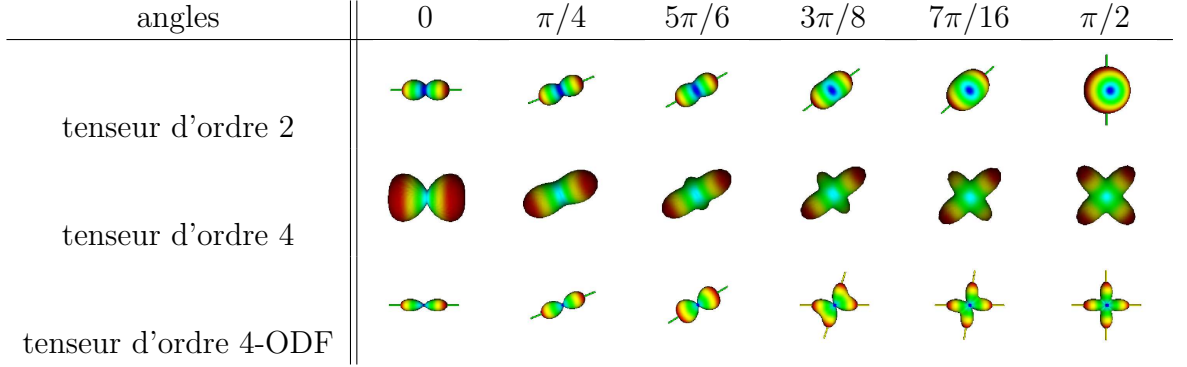


FIG. 2.6 – Estimation des modèles de diffusion et d'orientation de fibre associées pour un signal idéal (pas de bruit) d'un croisement de fibres avec un angle $\phi \in [0, \pi/4, 5\pi/6, 3\pi/8, 7\pi/16, \pi/2]$.

Comme avec la méthode des quartiques, les N_l coefficients de l'ODF de fibres d'ordre l définis positifs d_k ($k \in [1, \dots, N_l]$) sont donnés par la relation $\mathbf{C}\vec{x}$.

La deuxième façon de faire estime l'ODF de fibres à partir des N_l coefficients du tenseur d'ordre l modélisant la diffusivité. Ces coefficients sont placés dans un vecteur $\vec{t}$. Ce qui diffère avec la méthode précédente sont les observées y_i qui ne sont plus des observations brutes, mais qui sont calculées à partir du tenseur estimé. Avec les notations introduites précédemment, le système à résoudre s'écrit $\vec{y} = \mathbf{B}\vec{x}$ avec $\vec{y} = \exp(-b\mathbf{G}\vec{t})$:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} b_{1,1}(\vec{g}_1; \vec{c}_1) & \dots & b_{1,M'}(\vec{g}_1; \vec{c}_{M'}) \\ \vdots & \vdots & \vdots \\ \dots & b_{i,j}(\vec{g}_i; \vec{c}_j) & \dots \\ \vdots & \vdots & \vdots \\ b_{N,1}(\vec{g}_N; \vec{c}_1) & \dots & b_{N,M'}(\vec{g}_N; \vec{c}_{M'}) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{M'} \end{pmatrix} \quad (2.58)$$

Finalement, les N_l coefficients de l'ODF de fibres d'ordre l d_k ($k \in [1, \dots, N_l]$) sont donnés par la relation $\mathbf{C}\vec{x}$.

2.2.4 Comparaison des différentes méthodes d'estimation de tenseurs

Dans cette section, une comparaison des différentes méthodes d'estimation de tenseurs est proposée sur données simulées (cf. figure 2.6). Différentes configurations de croisement de fibres sont simulées pour montrer visuellement l'intérêt de considérer un modèle de tenseur d'ordre 4 plutôt qu'un modèle d'ordre 2 ainsi que l'intérêt de considérer une modélisation de l'ODF plutôt que de la diffusivité pour caractériser l'orientation des fibres.

Les données synthétiques sont simulées en utilisant un mélange de gaussiennes. Soit $\mathbf{D}_1$ un tenseur diagonal de termes diagonaux $1,7 \times 10^{-3}$, 2×10^{-4} et 2×10^{-4} . Le tenseur $\mathbf{D}_2(\phi)$ est

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle du tenseur de diffusion sur une base de polynômes homogènes

calculé en tournant $\mathbf{D}_1$ d'un angle ϕ autour du troisième vecteur propre. Soit $\vec{g}$ un vecteur de la sphère unité. Pour deux faisceaux de fibres, le signal simulé est donné par :

$$S(\vec{g}, b) = \sum_{i=1}^2 f_i \exp(-b \vec{g}^T \mathbf{D}_i \vec{g}) \quad (2.59)$$

où f_i et $\mathbf{D}_i$ sont, respectivement, la proportion et le tenseur de diffusion associés au faisceau i . Dans notre expérience, les f_i valent 1/2 et b vaut 3000.

A partir d'un ensemble de $N = 30$ signaux $S(\vec{g}_i, b)$ calculés pour différentes configurations de croisement de fibres $\phi \in [0, \pi/4, 5\pi/6, 3\pi/8, 7\pi/16, \pi/2]$, les tenseurs d'ordre 2, 4 et 4-ODF associés sont estimés. Quand ϕ varie de 0 à $\pi/2$, l'organisation des fibres passe d'un faisceau homogène organisé autour d'une direction privilégiée à un croisement orthogonal de deux faisceaux de fibres. On remarque immédiatement que seul le modèle d'ordre 4-ODF permet de restituer l'information d'inhomogénéité des structures au niveau voxel, dès lors que l'angle entre les faisceaux est plus grand que $5\pi/6$.

Chapitre 3

Le modèle de diffusion sur une base d'harmoniques sphériques

Sommaire

3.1	Décomposition sur une base d'harmoniques sphériques	50
3.1.1	Les Harmoniques Sphériques	50
3.1.2	Formulation générale du modèle	51
3.1.3	La fonction de distribution des orientations	52
3.2	Lien entre les bases de polynômes homogènes et des harmoniques sphériques	56

EN IRMd, il existe un état de l'art important sur les algorithmes de reconstruction qui ont pour objectif de synthétiser les informations de diffusion et d'orientation contenues dans les images pondérées en diffusion. Le résultat de tels algorithmes est une Fonction Symétrique sur la Sphère unité (FSS) en chaque voxel de l'image représentée par une forme géométrique. Parmi ces modèles, nous venons de présenter l'ITD généralisée (ITDG) (cf. section 2.1). Mais il est également possible d'utiliser l'imagerie Q-ball qui permet de modéliser soit l'ADC, soit la Fonction de Distribution des Orientations (ODF) [38, 11, 10, 1]. Jusqu'à présent, une FSS était représentée sur une base de Polynômes Homogènes (PH) contraints sur la sphère unité (ITDG). Toutefois, il est possible de modéliser ce signal en utilisant une base d'Harmoniques Sphériques (HS) (voir section 3.1). Les coordonnées sphériques $e = (r, \theta, \phi)^T$ permettent de décrire la diffusion de manière élégante sur la base d'HS. En effet, la décomposition du signal sur cette base est équivalente à une transformée de Fourier en coordonnées sphériques et introduit des propriétés utiles pour l'analyse du profil angulaire de diffusion (voir section 3.1). Lorsque le même ordre est considéré, une relation de passage existe entre ces deux bases (bijection) (voir section 3.2). Cette bijection est ensuite utilisée pour exprimer un tenseur issu de l'ITDG modélisant l'ADC dans la base d'HS afin de calculer facilement l'ODF associée. Finalement, une relation analytique entre ces deux bases est donnée.

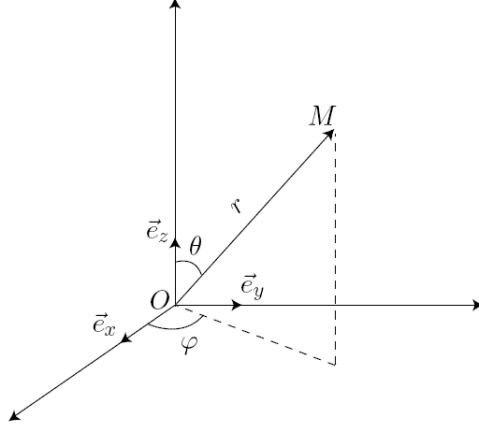


FIG. 3.1 – Repère sphérique.

3.1 Décomposition sur une base d'harmoniques sphériques

3.1.1 Les Harmoniques Sphériques

Le formalisme (physique) utilisé dans cet exposé est le suivant : θ représente les coordonnées polaires ($\theta \in [0, \pi]$) et ϕ représente les coordonnées azimutales ($\phi \in [0, 2\pi]$) (cf. figure 3.1). Les Harmoniques Sphériques (HS) $Y_l^m(\theta, \phi)$ forment une base de fonctions complexes continues bornées. Elles sont définies pour un ordre $l = 0, \dots, \infty$ et un degré $m = -l, \dots, l$ de la façon suivante (voir également la figure 3.2) :

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos(\theta)) \exp(im\phi) \quad (3.1)$$

où $P_l^m(x)$ sont les polynômes de Legendre associés calculés avec les équations suivantes :

$$\begin{aligned} P_l(x) &= \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \\ P_l^m(x) &= (-1)^m (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_l(x), \quad m \geq 0 \\ P_l^{-m}(x) &= (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \end{aligned} \quad (3.2)$$

Pour décomposer la diffusion qui prend ses valeurs dans R^{+*} , il faut définir une base modifiée d'harmoniques sphériques qui est réelle, symétrique et orthogonale. Cette base modifiée est donnée par :

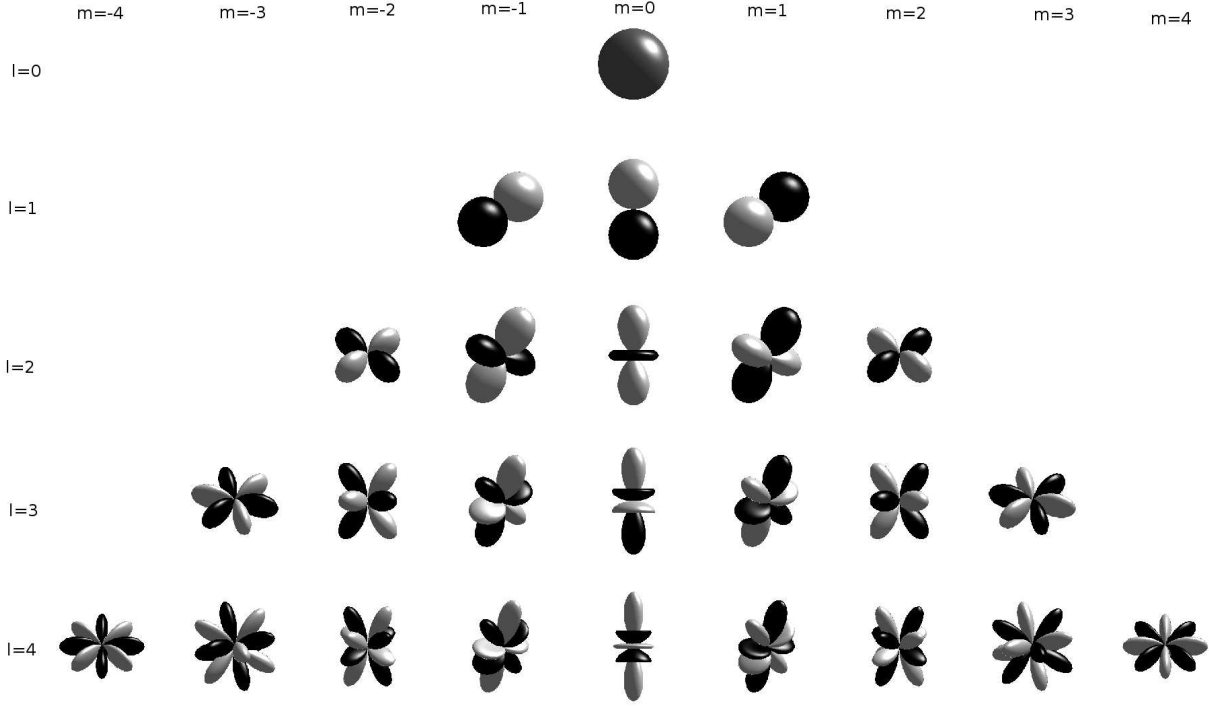


FIG. 3.2 – Base des harmoniques sphériques modifiée pour les ordres 0,1,2,3 et 4. Une couleur noire représente une valeur positive, alors qu'une couleur blanche représente une valeur négative.

$$Y_j(\theta, \phi) = \begin{cases} \sqrt{2} \operatorname{Re}(Y_l^m(\theta, \phi)), & m > 0 \\ Y_l^0(\theta, \phi), & m = 0 \\ \sqrt{2} \operatorname{Im}(Y_l^{|m|}(\theta, \phi)), & m < 0 \end{cases} \quad (3.3)$$

avec $j(l, m) = (l^2 + l + 2)/2 + m$. Cette base d'harmoniques sphériques modifiée a deux propriétés intéressantes. Premièrement, les harmoniques paires sont antipodalement symétriques, et deuxièmement, les harmoniques impaires sont antisymétriques (cf. figure 3.2). Dans la suite, nous nous intéresserons particulièrement à cette base jusqu'à l'ordre 4, impliquant ainsi 15 termes Y_j dont les expressions sont données en annexe-A.

3.1.2 Formulation générale du modèle

Le signal sur la base des harmoniques sphériques modifiée Y_j s'écrit comme une série de Laplace tronquée [11] :

$$d(\theta, \phi) = \sum_{j=1}^R c_j Y_j(\theta, \phi) \quad (3.4)$$

où R est le nombre de termes dans la base d'HS modifiées. Il est alors possible d'estimer les coefficients c_j en résolvant le système suivant, par exemple au sens des moindres carrés :

$$\begin{pmatrix} -\frac{1}{b} \ln \left(\frac{S_1}{S_0} \right) \\ \vdots \\ -\frac{1}{b} \ln \left(\frac{S_i}{S_0} \right) \\ \vdots \\ -\frac{1}{b} \ln \left(\frac{S_N}{S_0} \right) \end{pmatrix} = \begin{pmatrix} Y_1(\theta_1, \phi_1) & \dots & Y_R(\theta_1, \phi_1) \\ \vdots & \vdots & \vdots \\ Y_1(\theta_N, \phi_N) & \dots & Y_R(\theta_N, \phi_N) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_R \end{pmatrix} \quad (3.5)$$

Ces coefficients estimés modélisent la diffusivité. Pour les mêmes raisons que précédemment (*i.e.*, pour aligner le modèle sur les structures anatomiques sous-jacentes), il faut calculer l'ODF.

3.1.3 La fonction de distribution des orientations

Lorsque l'on décompose les Coefficients de Diffusion Apparents (ADC) sur la base d'harmoniques sphériques, on modélise la diffusivité. Les maxima de cette quantité ne correspondent pas à l'orientation des fibres sous-jacentes. Il est donc intéressant de définir une fonction dont les maxima correspondent à l'orientation des fibres : la Fonction de Distribution des Orientations (ODF). Dans la littérature, trois méthodes ont été proposées : l'ODF de diffusion par projection radiale, l'ODF de fibres par déconvolution sphérique et l'ODF de diffusion angle solide qui permet de pénaliser les petits déplacements des molécules d'eau et donc d'obtenir naturellement une forme piquée.

L'ODF par projection radiale

Tuch [38] a montré que l'ODF pouvait être estimée à partir des acquisitions du signal sur une seule sphère (liée à la b -valeur et appelée en anglais *shell*), sans reconstruire complètement la d.d.p. (densité de probabilité) de diffusion. On ne s'intéresse ici qu'à l'information angulaire. On peut reconstruire l'ODF ψ par projection radiale de la d.d.p. de diffusion P pour des valeurs de Δ suffisamment grandes :

$$\psi(\vec{u}) = k \int_0^\infty P(r\vec{u}, \Delta) f(r) dr \quad (3.6)$$

où $\vec{u}$ est un vecteur unité, k un coefficient de normalisation et f est une fonction qui permet d'associer un poids aux déplacements : ici $f(r) = 1$. L'ODF est donc une fonction sur la sphère unité qui décrit la probabilité angulaire de diffusion.

Tuch a montré qu'il était possible de calculer l'ODF directement à partir du signal issu d'acquisitions, sur une seule sphère, grâce à la transformée de Funk-Radon (FRT) [38]. Cette transformation est une généralisation de la transformée de Radon en coordonnées sphériques. Il faut donc lier la FRT avec la projection radiale du propagateur simplifiée ψ . La transformée de Funk-Radon est notée G . Elle fait correspondre à un point sur la sphère unité la direction d'intégration correspondante sur l'équateur. Soit f une fonction sphérique, on obtient :

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle de diffusion sur une base d'harmoniques sphériques

$$G[f(\vec{u})] = \int_S \delta(\vec{v}^T \vec{u}) f(\vec{v}) d\vec{v} = \int_{\vec{v} \perp \vec{u}} f(\vec{v}) d\vec{v} \quad (3.7)$$

où $\vec{u}$ et $\vec{v}$ sont des vecteurs unités. En remplaçant f dans (3.7) par l'expression du signal de diffusion, il est possible de relier ce signal au propagateur via la FRT [38]. Néanmoins, le point difficile de la FRT se situe au niveau de l'intégration. Il est possible d'obtenir une solution élégante en effectuant les calculs sur la base des harmoniques sphériques modifiée [10]. En remplaçant le signal par son extension en HS (3.4) on obtient :

$$\begin{aligned} G[S(\vec{u})/S_0] &= \frac{1}{S_0} \int_{\|\vec{v}\|=1} \delta(\vec{v}^T \vec{u}) S(\vec{v}) d\vec{v} \\ &= \frac{1}{S_0} \int_{\|\vec{v}\|=1} \delta(\vec{v}^T \vec{u}) \sum_{j=1}^R c_j Y_j(\vec{v}) d\vec{v} \\ &= \frac{1}{S_0} \sum_{j=1}^R c_j \overbrace{\int_{\|\vec{v}\|=1} \delta(\vec{v}^T \vec{u}) Y_j(\vec{v}) d\vec{v}}^{I_j} \end{aligned} \quad (3.8)$$

En utilisant le corollaire du théorème de Funk-Hecke [10], les I_j peuvent être calculés :

$$I_j = \int_{\|\vec{v}\|=1} \delta(\vec{v}^T \vec{u}) Y_j(\vec{v}) d\vec{v} = 2\pi P_{l(j)}(0) Y_j(\vec{u}) \quad (3.9)$$

où $P_{l(j)}(0)$ est le polynôme de Legendre associé à l'HS modifiée j de degré $l(j)$ évalué en 0 :

$$P_l(0) = \begin{cases} 0 & \text{si } l \text{ est impair} \\ (-1)^{l/2} \frac{1.3.5...(l-1)}{2.4.6...l} & \text{si } l \text{ est pair} \end{cases} \quad (3.10)$$

La FRT d'une fonction sphérique décomposée sur la base d'HS modifiée dans une direction unitaire $\vec{u}$ est finalement donnée par :

$$G[S(\vec{u})/S_0] = \sum_{j=1}^R \frac{2\pi P_{l(j)}(0)}{S_0} c_j Y_j(\vec{u}) \quad (3.11)$$

Ainsi, les coefficients de l'ODF sur cette base d'HS modifiée s'obtiennent par simple produit matriciel :

$$C' = \begin{pmatrix} \ddots & & & \\ & 2\pi \frac{(-1)^{l(j)/2}}{S_0} \frac{1.3.5...(l(j)-1)}{2.4.6...l(j)} & & \\ & & \ddots & \end{pmatrix} \begin{pmatrix} \vdots \\ c_j \\ \vdots \end{pmatrix} \quad (3.12)$$

L'ODF piquée

En pratique, l'échantillonnage des gradients est faible (~ 30 directions) et la b-valeur également (~ 1000 s/mm²). Pour ces raisons, le besoin d'un modèle plus piqué est apparu. Dans [10], une transformation linéaire des coefficients des HS décrivant l'ODF permet de transformer l'ODF lissée en une ODF piquée (cette approche s'inspire des méthodes de déconvolution sphérique). Connaissant la fonction de réponse R représentant l'ODF d'une fibre, $\psi_{\text{piquée}}$, l'ODF de fibres piquée, est reliée à l'ODF de diffusion ψ par la relation suivante :

$$\psi(\vec{u}) = \int_{\|\vec{v}\|=1} R(\vec{u} \cdot \vec{v}) \psi_{\text{piquée}}(\vec{v}) d\vec{v} \quad (3.13)$$

Remplaçons ψ et $\psi_{\text{piquée}}$ par leur approximation sur une base d'HS modifiée, $\psi(\vec{u}) = \sum_{j=1}^R c'_j Y_j(\vec{u})$ et $\psi_{\text{piquée}}(\vec{u}) = \sum_{j=1}^R f_j Y_j(\vec{u})$:

$$\sum_{j=1}^R c'_j Y_j(\vec{u}) = \sum_{j=1}^R f_j \int_{\|\vec{v}\|=1} R(\vec{u} \cdot \vec{v}) Y_j(\vec{v}) d\vec{v} \quad (3.14)$$

En utilisant le théorème de Funk-Hecke pour résoudre l'intégration, on obtient [10] :

$$f_j = \frac{c'_j}{r'_j} \quad \text{avec} \quad r'_j = 2\pi \int_{-1}^1 P_{l(j)}(t) R(t) dt \quad (3.15)$$

Reste à choisir un modèle pour la réponse impulsionnelle R d'une fibre. Dans la littérature, un tenseur axialement symétrique est choisi (*i.e.*, un tenseur ayant comme valeurs propres $[\lambda_1, \lambda_2, \lambda_2]$ tel que $\lambda_1 \gg \lambda_2$). En exploitant le lien analytique entre tenseur et ODF [10], on obtient pour R :

$$R(t) = \frac{1}{Z'} \frac{1}{\sqrt{(\lambda_2/\lambda_1 - 1)t^2 + 1}} \quad (3.16)$$

où Z' est une constante de normalisation. Pour un tenseur présentant une symétrie axiale, $Z' = 8\pi b \sqrt{\lambda_2^2 \lambda_1}$ [10]. Les coefficients des HS de la nouvelle ODF piquée sont obtenus en remplaçant R dans (3.15) après intégration :

$$f_j = \frac{4b\sqrt{\lambda_1 \lambda_2}}{A'_{l(j)}(1 - \lambda_2/\lambda_1)} c'_j \quad (3.17)$$

où $A'_{l(j)}(\alpha) = \int_{-1}^1 (1 - \alpha t^2)^{-1/2} P_{l(j)}(t) dt$. Ces coefficients sont donnés jusqu'à l'ordre 4 en annexe-B.

L'ODF angle solide

Pour une sphère de rayon r , un élément de volume est donné par : $d\nu = r^2 dr d\Omega$ avec $d\Omega = \sin\theta d\theta d\phi$ l'angle solide. Soit $ODF(\vec{u})d\Omega$ la probabilité de diffusion dans la direction $\vec{u}$ à travers l'angle solide $d\Omega$. Cette quantité est calculée en intégrant le propagateur $P(\vec{r})d\nu = P(r\vec{u})r^2 dr d\Omega$ pour toutes les valeurs de r en gardant la direction $\vec{u}$ constante :

Mesure et modélisation de la diffusion des molécules d'eau

Le modèle de diffusion sur une base d'harmoniques sphériques

$$ODF(\vec{u})d\Omega = \int_0^\infty P(r\vec{u})r^2drd\Omega \quad (3.18)$$

Après simplification, on remarque que calculer une ODF angle solide revient à prendre $f(r) = r^2$ dans (3.6). f permet donc de donner plus d'importance aux grands déplacements. Dans le cas précédent ($f(r) = 1$), la projection radiale donne un poids équivalent aux points qui sont près ou éloignés de l'origine. On obtient comme expression de l'ODF avec $f(r) = r^2$ [1] :

$$\psi(u) = \frac{1}{4\pi} + \frac{1}{16\pi^2} G[\nabla_b^2 \ln(-\ln(S(u)))] \quad (3.19)$$

où ∇_b^2 est l'opérateur de Laplace-Beltrami défini par :

$$\nabla_b^{2*} = \frac{1}{\sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial *}{\partial \phi} \right) + \frac{1}{\sin^2(\theta)} \frac{\partial^2 *}{\partial \phi^2} \quad (3.20)$$

Implémentation

Tout d'abord, il faut calculer par une procédure des moindres carrés classique les coefficients c_j du signal doublement passé au logarithme :

$$\ln(-\ln(S(u)/S_0)) = \sum_{j=1}^R c_j Y_j(u) \quad (3.21)$$

Ensuite, il faut appliquer l'opérateur de Laplace-Beltrami. Pour ce faire, il suffit de multiplier chaque coefficient c_j par la valeur propre $-l(j)(l(j) + 1)$. Il reste à appliquer la transformée de Funk-Radon en utilisant l'équation (3.12). Finalement, les coefficients angle solide sur la base d'HS modifiée sont donnés par :

$$c'_j = \begin{cases} \frac{1}{2\sqrt{\pi}} & \text{si } j = 1 \\ -\frac{1}{8\pi} (-1)^{l(j)/2} \frac{1.3.5 \dots (l(j)+1)}{2.4.6 \dots (l(j)-2)} c_j & \text{si } j > 1 \end{cases} \quad (3.22)$$

Comparaison des modèles ODF sur une base d'harmoniques sphériques

Dans cette section, différentes configurations de croisement de fibres sont simulées pour montrer visuellement les limitations en termes de résolutions angulaires des modèles d'ODF par projection radiale, d'ODF piquée et d'ODF angle solide.

Les données synthétiques sont simulées en utilisant l'équation (2.59) (*i.e.*, un mélange de gaussiennes), avec $f_i = 1/2$ et $b = 3000$. A partir d'un ensemble de $N = 30$ signaux $S(\vec{g}_i, b)$ calculés pour différentes configurations de croisement de fibres $\phi \in [0, \pi/4, 5\pi/6, 3\pi/8, 7\pi/16, -\pi/2]$, les ODF obtenues par projection radiale (cf. figure 3.3-a), les ODF piquées correspondantes ($\lambda_1 = 1$ et $\lambda_2 = 0,1$) (cf. figure 3.3-b) et les ODF angle solide (cf. figure 3.3-c) sont calculées.

Quand ϕ varie de 0 à $\pi/2$, l'organisation des fibres passe d'un faisceau homogène organisé autour d'une direction privilégiée à un croisement orthogonal de deux faisceaux de fibres.

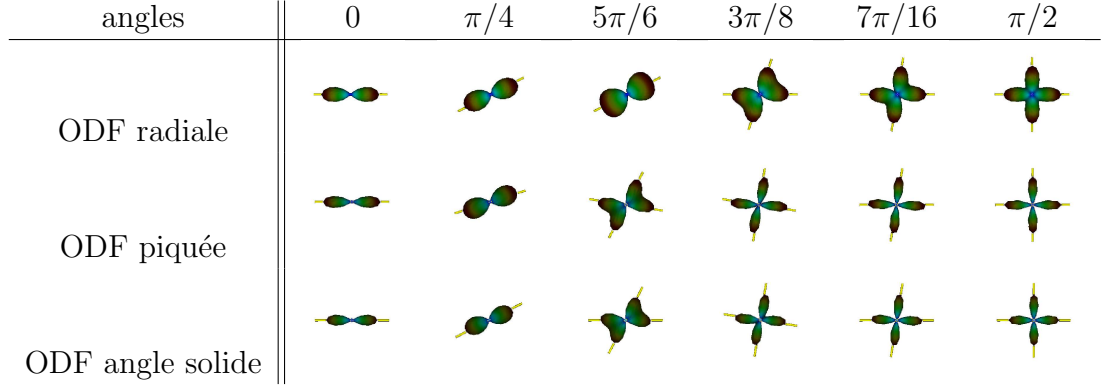


FIG. 3.3 – Estimation de l'ODF via la base des harmoniques sphériques modifiées.

On remarque que l'ODF angle solide et l'ODF piquée donnent les meilleurs résultats et permettent de dissocier des faisceaux se croisant avec un angle supérieur à $\pi/4$ (alors que la méthode par projection radiale ne dissocie les faisceaux qu'à partir d'un angle de $5\pi/6$). Les résultats obtenus avec l'ODF angle solide et l'ODF piquée sont similaires, néanmoins il est important de rappeler que seul l'ODF angle solide ne dépend d'aucun paramètre alors que l'ODF piquée est fonction de $(\lambda_1 \text{ et } \lambda_2)$.

3.2 Lien entre les bases de polynômes homogènes et des harmoniques sphériques

On a vu que le calcul de l'ODF sur la base d'HS modifiée s'effectue par simple produit matriciel (3.12), (3.22). Afin de simplifier les calculs, il peut être intéressant de naviguer entre le modèle exprimé sur une base de Polynômes Homogènes (PH) et le même modèle exprimé sur une base d'HS modifiée. Dans [10], il a été démontré que les HS paires d'ordre k et les PH d'ordre k sur la sphère sont des bases du même espace vectoriel. Il est ainsi possible de définir une transformation linéaire $\mathbf{M}$ entre ces deux bases qui permet d'exprimer les coefficients d'une série d'HS modifiée c_j en coefficients d'un polynôme homogène sur la sphère. Soit S une FSS décomposée sur la base d'HS modifiée :

$$c_j = \int_{\Omega} S(\theta, \phi) Y_j(\theta, \phi) d\Omega \quad (3.23)$$

où Y_j est le j ème terme de la base d'HS modifiée. On a vu précédemment qu'un polynôme homogène de degré l sur la sphère s'écrit (2.9) :

$$P(\vec{g}) = \sum_{k=1}^{N_l} \mu_k^l d_k(g_1)^{n_1^l(k)} (g_2)^{n_2^l(k)} (g_3)^{n_3^l(k)} \quad (3.24)$$

où N_l est le nombre d'éléments indépendants au degré l . Si P décrit la même FSS, alors en remplaçant S dans (3.23) par P , on obtient :

$$\begin{aligned} c_j &= \sum_{k=1}^{N_l} d_k \int_{\Omega} \mu_k^l(g_1)^{n_1^l(k)}(g_2)^{n_2^l(k)}(g_3)^{n_3^l(k)} Y_j(\theta, \phi) d\Omega \\ \Rightarrow \vec{c} &= \mathbf{M} \vec{a} \end{aligned} \quad (3.25)$$

Avec :

$$\begin{aligned} \mathbf{M} &= \begin{pmatrix} \mu_1^l \int_{\Omega} (g_1)^{n_1^l(k)}(g_2)^{n_2^l(k)}(g_3)^{n_3^l(k)} Y_1(\theta, \phi) d\Omega & \dots & \mu_{N_l}^l \int_{\Omega} (g_1)^{n_1^l(k)}(g_2)^{n_2^l(k)}(g_3)^{n_3^l(k)} Y_1(\theta, \phi) d\Omega \\ \vdots & & \vdots \\ \mu_1^l \int_{\Omega} (g_1)^{n_1^l(k)}(g_2)^{n_2^l(k)}(g_3)^{n_3^l(k)} Y_{N_l}(\theta, \phi) d\Omega & \dots & \mu_{N_l}^l \int_{\Omega} (g_1)^{n_1^l(k)}(g_2)^{n_2^l(k)}(g_3)^{n_3^l(k)} Y_{N_l}(\theta, \phi) d\Omega \end{pmatrix} \\ \vec{a} &= (d_1 \dots d_{N_l})^T \end{aligned} \quad (3.26)$$

Il a été montré dans [10] que la matrice $\mathbf{M}$ de taille $N_l \times N_l$ est une matrice de changement de base et est donc inversible. Ainsi, soit $\vec{c}$ un vecteur de N_l coefficients issus de la série de Laplace tronquée à l'ordre l . $\mathbf{M}^{-1}\vec{c}$ permet de calculer les coefficients associés au PH contraint sur la sphère. Il existe donc une bijection entre les deux bases. A l'ordre 4, cette matrice de passage est donnée en annexe-C.

Conclusion partielle

Nous venons de présenter une description de l'imagerie de diffusion qui a permis d'appréhender les problèmes inhérents à cette modalité. Après une description du processus physique en jeu lors de la formation des images de diffusion, le modèle du tenseur de diffusion, les contraintes, et les outils associés ont été présentés. Différentes représentations ont permis de modéliser soit le signal de diffusion, soit la diffusivité, soit la probabilité de déplacement des molécules d'eau. Les modèles du tenseur d'ordre 2 et 4 seront exploités respectivement en parties III et V. La partie suivante présente la problématique et les dernières modélisations statistiques pour la détection de changements en IRM de diffusion.

Bibliographie

- [1] I. AGANJ, C. LENGLET et G. SAPIRO : ODF reconstruction in Q-ball imaging with solid angle consideration. *In IEEE International Symposium on Biomedical Imaging, Boston, ISBI '09*, p. 1398 – 1401, 2009.
- [2] Y. ASSAF et P. BASSER : Composite hindered and restricted model of diffusion (charmed) MR imaging of the human brain. *NeuroImage*, 27(1):48 – 58, 2005.
- [3] A. BARMPOUTIS, M. HWANG, D. HOWLAND, J. FORDER et B. VEMURI : Regularized positive-definite fourth-order tensor field estimation from DW-MRI. *NeuroImage*, 45(1): 153 – 162, 2009.
- [4] A. BARMPOUTIS et B. VERMURI : A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints. *In IEEE International Symposium on Biomedical Imaging, Rotterdam, ISBI '10*, p. 1385 – 1388, 2010.
- [5] P. BASSER et D. JONES : Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR Biomed.*, 8(7-8):333 – 344, 1995.
- [6] P. BASSER et D. JONES : Diffusion-tensor MRI : theory, experimental design and data analysis - a technical review. *NMR Biomed.*, 17(7-8):456 – 467, 2002.
- [7] P. BASSER et C. PIERPAOLI : Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance*, 111(3):209 – 219, 1996.
- [8] C. BERG : The multidimensional moment problem and semigroups. *Moments in mathematics, American Mathematical Society*, 37:110 – 124, 1987.
- [9] E. CANALES-RODRIGUEZ, C. LIN, Y. ITURRIA-MEDINA, C. YEH, K. CHO et L. MELIE-GARCIA : Diffusion orientation transform revisited. *NeuroImage*, 49(2):1326 – 1339, 2010.
- [10] M. DESCOTEAUX : *High Angular Resolution Diffusion MRI : From Local Estimation to Segmentation and Tractography*. Thèse de doctorat, Nice - Sophia Antipolis, 2008.
- [11] M. DESCOTEAUX, E. ANGELINO, S. FITZGIBBONS et R. DERICHE : A fast and robust ODF estimation algorithm in Q-ball imaging. *In IEEE International Symposium on Biomedical Imaging, Arlington, ISBI '06*, p. 81 – 84, 2006.
- [12] M. DESCOTEAUX, R. DERICHE, D. BIHAN, J. MANGIN et C. POUPON : Diffusion propagator imaging : Using laplace's equation and multiple shell acquisitions to recons-

- truct the diffusion propagator. *In International Conference on Information Processing in Medical Imaging, Williamsburg, IPMI '09*, p. 1 – 13, 2009.
- [13] P. FILLARD, X. PENNEC, V. ARSIGNY et N. AYACHE : Clinical DT-MRI estimation, smoothing, and fiber tracking with log-euclidean metrics. *IEEE Transactions on Medical Imaging*, 26(11):1472 – 1482, 2007.
- [14] A. GHOSH, E. TSIGARIDAS, M. DESCOTEAUX, P. COMON, B. MOURRAIN et R. DERICHE : A polynomial based approach to extract the maxima of an antipodally symmetric spherical function and its application to extract fiber directions from the orientation distribution function in diffusion MRI. *In Medical Image Computing and Computer Assisted Intervention, Workshop on Computational Diffusion MRI, New York, MICCAI '08*, p. 237 – 248, 2008.
- [15] D. HILBERT : Ueber die Darstellung definiter Formen als Summe von Formenquadraten. *Mathematische Annalen*, 32:342–350, 1888.
- [16] E. IM : Hessian sufficiency for bordered hessian. *Research Letters in the Information and Mathematical Sciences*, 8:189 – 196, 2005.
- [17] B. JIAN, B. VEMURI, E. OZARSLAN, P. CARNEY et T. MARECI : A novel tensor distribution model for the diffusion-weighted MR signal. *NeuroImage*, 37(1):164 – 176, 2007.
- [18] D. JONES, M. HORSFIELD et A. SIMMONS : Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42(3):515 – 525, 1999.
- [19] J. KARGER et W. HEINK : The propagator representation of molecular transport in microporous crystallites. *Journal of Magnetic Resonance*, 51:1 – 7, 1983.
- [20] B. LANDMAN, P. BAZIN et J. PRINCE : Diffusion tensor estimation by maximizing rician likelihood. *In IEEE International Conference on Computer Vision, Rio, ICCV '07*, p. 1 – 8, 2007.
- [21] C. LAWSON et R. HANSON : *Solving Least Squares Problems*. Prentice-Hall, 1974.
- [22] J. MANGIN, C. POUPON, C. CLARK, D. LE BIHAN et I. BLOCH : Distortion correction and robust tensor estimation for MR diffusion imaging. *Medical Image Analysis*, 6(3):191 – 198, 2002.
- [23] L. MINATI et W. WEGLARZ : Physical foundations, models, and methods of diffusion magnetic resonance imaging of the brain : A review. *Concepts in Magnetic Resonance Part A*, 30A(5):278 – 307, 2007.
- [24] A. NEEMAN : *Visualization Techniques for Computational Mechanics*. Thèse de doctorat, Santa Cruz-California, 2009.
- [25] E. OZARSLAN et T. MARECI : Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50:955 – 965, 2003.
- [26] E. OZARSLAN, T. SHEPHERD, B. VEMURI, S. BLACKBAND et T. MARECI : Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, 31(3):1086 – 1103, 2006.

- [27] E. OZARSLAN, B. VERMURI et T. MARECI : Generalized scalar measures for diffusion MRI using trace, variance, and entropy. *Magnetic Resonance in Medicine*, 53:866 – 876, 2005.
- [28] W. PRICE : Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion, part 1 : basic theory. *Magnetic Resonance : an Educational Journal*, 9(5):299 – 336, 1997.
- [29] L. QI : Eigenvalues of a real supersymmetric tensor. *Journal of Symbolic Computation*, 40:1302 – 1324, 2005.
- [30] L. QI, F. WANG et Y. WANG : Z-eigenvalue methods for a global polynomial optimization problem. *Mathematical Programming*, 118:301 – 316, 2009.
- [31] L. QI, Y. WANG et E. WU : D-eigenvalues of diffusion kurtosis tensors. *Journal of Computational and Applied Mathematics*, 221:150 – 157, 2008.
- [32] W. RUDIN : Sums of squares of polynomials. *The American Mathematical Monthly*, 107:813 – 821, 2000.
- [33] R. SALVADOR, A. PENA, D. MENON, T. CARPENTER, J. PICKARD et E. BULLMORE : Formal characterization and extension of the linearized diffusion tensor model. *Human Brain Mapping*, 2005(2):144 – 155, 2005.
- [34] T. SCHULTZ et H. SEIDEL : Estimating crossing fibers : A tensor decomposition approach. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1635 – 1642, 2008.
- [35] E. STEJSKAL et J. TANNER : Spin diffusion measurements : Spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42:288 – 292, 1965.
- [36] A. TARANTOLA, B. COLL, K. MOSEGAARD et P. XU : The null information probability density for anisotropic tensors, 2000.
- [37] J. TOURNIER, F. CALAMANTE, D. GADIAN et A. CONNELLY : Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *Neuroimage*, 23(3):1176 – 1185, 2004.
- [38] D. TUCH : Q-Ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358 – 1372, 2004.
- [39] V. WEDEEN, P. HAGMANN, W. TSENG, T. REESE et R. WEISSKOFF : Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377 – 1386, 2005.
- [40] T. WELDESELAASSIE, A. BARMPOUTIS et M. ATKINS : Symmetric positive-definite cartesian tensor orientation distribution functions (CT-ODF). In *Medical Image Computing and Computer Assisted Intervention, Pékin*, num. 6361 de LNCS, p. 582 – 589, 2010.
- [41] C. WESTIN, S. MAIER, H. MAMATA, A. NABAVI, F. JOLESZ et R. KIKINIS : Processing and visualization of diffusion tensor MRI. *Medical Image Analysis*, 6(2):93 – 108, 2002.

PARTIE II

Détection de changements en IRM de diffusion

Introduction

LA deuxième partie traite de la thématique centrale de cette thèse : la détection de changements en imagerie de diffusion. La détection de changements correspond à la localisation de régions ayant évolué entre deux images. L'utilisation de ces approches ne se limite pas à l'imagerie médicale, mais couvre un spectre très large d'applications.

Tout d'abord, le premier chapitre introduit les concepts relatifs aux statistiques inférentielles. Des règles de décision issues de test d'hypothèses permettent de détecter des régions en évolution. Cette détection se fait à un niveau de probabilité d'erreur donné, appelée niveau de signification du test, à partir duquel il est possible d'établir une valeur de seuil de détection en considérant, si nécessaire, les problèmes liés aux comparaisons multiples. Ensuite, le deuxième chapitre présente un état de l'art général sur la détection de changements en imagerie médicale. Cet état de l'art permet d'explorer les différentes techniques de détection de changements développées pour l'imagerie médicale, puis s'intéresse plus particulièrement à la détection de changements en imagerie de diffusion. Finalement, le troisième chapitre est consacré à la description d'un cadre statistique général pour le suivi longitudinal en imagerie de diffusion. Ce cadre statistique peut être appliqué à différentes problématiques et sera mis en œuvre dans la suite de ce mémoire de thèse pour la détection automatique de changements dans des séquences temporelles d'images de diffusion.

Chapitre 4

Statistiques inférentielles

Sommaire

4.1	Les tests d'hypothèses	69
4.2	Les deux types d'erreurs possibles	70
4.3	Signification statistique : accepter ou rejeter l'hypothèse nulle	72
4.3.1	Tests d'hypothèses simples	72
4.3.2	Problème des comparaisons multiples	73

LA détection de changements consiste à définir une règle de décision entre deux hypothèses. Cette règle permet de répondre à la question suivante : quels sont les voxels de l'image où un changement significatif a eu lieu ? Ce chapitre commence par un rappel sur les tests d'hypothèses. Un test d'hypothèses est une procédure qui permet d'évaluer une hypothèse statistique à partir d'un jeu d'échantillons. Ensuite, le concept d'erreur, à la base des statistiques inférentielles, est présenté. Finalement, la prise de décision se fait à un niveau de probabilité d'erreur donné, appelée niveau de signification du test, à partir duquel il est possible d'établir une valeur seuil de détection en considérant, si nécessaire, les problèmes liés aux comparaisons multiples.

4.1 Les tests d'hypothèses

Pour introduire les tests d'hypothèses, nous allons étudier un exemple fictif. Soit une étude effectuée sur un intervalle de temps significatif qui a permis de montrer que les moyennes annuelles des étudiants de l'université suivent une loi normale $\mathcal{N}(12; 1,3)$. Maintenant, une université prétend avoir augmenté cette moyenne en utilisant une nouvelle méthode d'enseignement. Cette nouvelle méthode fut mise à l'essai entre 2001 et 2009. Les notes moyennes issues de cette étude sont présentées dans le tableau 4.1. Deux hypothèses s'affrontent : ou bien la nouvelle méthode est sans effet, ou bien, au contraire, elle permet réellement d'augmenter la moyenne des étudiants. L'énoncé du problème est le suivant : soit X la variable aléatoire représentant la note annuelle moyenne des étudiants universitaires et M son espérance mathématique, alors il faut tester les deux hypothèses suivantes :

Année	2001	2002	2003	2004	2005	2006	2007	2008	2009
Moyenne	12,8	13,9	14,2	13,1	12,9	13,2	12,1	12,3	14,5

TAB. 4.1 – Données de l'étude sur la nouvelle méthode d'enseignement.

Vérité Décision	H_0	H_1
	$1 - \alpha$ (VN) α (FP)	β (FN) $1 - \beta$ (VP)

 TAB. 4.2 – Erreurs de première (α) et deuxième (β) espèce.

$$\begin{aligned}
 H_0 : M &= 12 \\
 H_1 : M &> 12
 \end{aligned} \tag{4.1}$$

Ce test va permettre de répondre à la question suivante : faut-il opter pour la nouvelle méthode d'enseignement ? Ou encore, est-ce que les faits observés contredisent nettement la validité de l'hypothèse H_0 (ou hypothèse nulle) ? Notons que H_1 s'appelle l'hypothèse alternative. Ensuite, puisque le test d'hypothèses porte sur M , il est naturel de s'intéresser à la moyenne $\bar{X}$ des observations. $\bar{X}$ est aussi appelée la variable de décision. Si H_0 est vraie, la moyenne $\bar{X}$ suit une loi normale $\mathcal{N}(M = 12; \frac{1,3}{\sqrt{9}})$ (car l'expérience s'est déroulée sur 9 ans). La règle de décision est la suivante :

- si $\bar{X}$ est suffisamment grand (*i.e.*, supérieur à un seuil k , fonction d'une probabilité d'erreur α), on opte pour H_1 avec une probabilité α de se tromper (on choisit généralement $\alpha = 5\%$ ou 1%).
- si $\bar{X} < k$, on conserve H_0 .

Ce choix pour le seuil k (autrement dit d'une probabilité « de se tromper » α) va introduire deux types d'erreurs.

4.2 Les deux types d'erreurs possibles

D'après l'exemple de la partie précédente, un test permet d'accepter ou de rejeter l'hypothèse H_0 . Par ailleurs, soit H_0 , soit H_1 est vraie. La décision engendre quatre possibilités caractérisées par des probabilités liées aux deux types d'erreurs que l'on peut commettre (cf. tableau 4.2). Soient α et β les probabilités d'erreurs de première et deuxième espèce, respectivement (cf. figure 4.1) :

- α est la probabilité de choisir H_1 alors que H_0 est vraie, ou encore, la probabilité d'avoir un faux positif (FP).
- $1 - \alpha$ est la probabilité de choisir H_0 alors que H_0 est vraie, ou encore, la probabilité d'avoir un vrai négatif (VN).

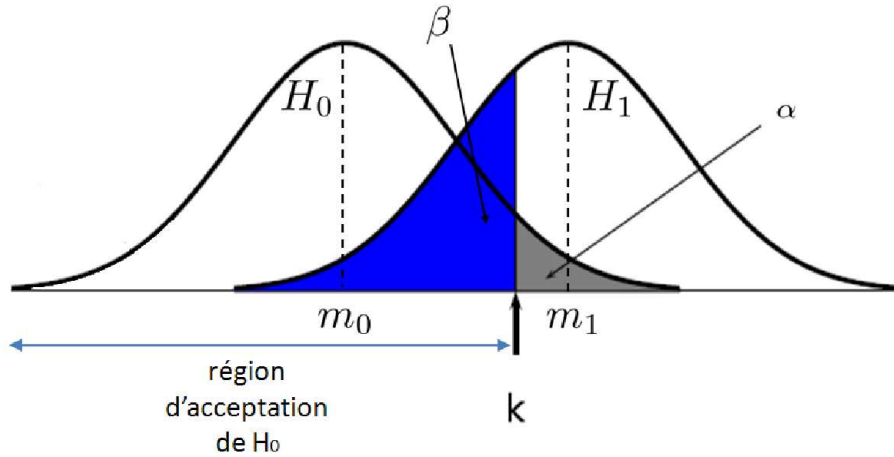


FIG. 4.1 – Erreurs de première (α) et de deuxième (β) espèce pour un test unilatéral de l'hypothèse $M = m_0$ au risque α .

- β est la probabilité de choisir H_0 alors que H_1 est vraie, ou encore, la probabilité d'avoir un faux négatif (FN).
- $1 - \beta$ est la probabilité de choisir H_1 alors que H_1 est vraie, ou encore, la probabilité d'avoir un vrai positif (VP). $1 - \beta$ est aussi la « puissance » du test, qui dépendra en particulier de la séparation des deux distributions.

En pratique, ces erreurs correspondent à des risques différents. Dans l'exemple précédent, le risque de première espèce consiste à accepter le nouvel enseignement alors qu'il est inefficace, alors que le risque de deuxième espèce consiste à laisser passer une occasion d'améliorer la moyenne d'un étudiant, et donc d'améliorer le système éducatif. En général, les lois de probabilités sous H_1 ne sont pas connues, et il est d'usage de fixer α en fonction du risque de première espèce désiré.

Pour mener à bien l'étude précédente sur les moyennes des étudiants, il faut choisir un niveau de signification α pour le test statistique. Un niveau α de 0,05 est choisi, c'est-à-dire qu'on accepte de se tromper dans 5 cas sur 100. Il est alors facile de calculer le seuil de détection k associé à α grâce aux tables statistiques de l'écart réduit pour la loi normale $\mathcal{N}(0; 1)$. On trouve :

$$k = 12 + \frac{1,3}{3} 1,64 = 12,71 \quad (4.2)$$

Finalement, la règle de décision s'écrit :

- si $\bar{X} > 12,71$, rejeter H_0 et accepter H_1 .
- si $\bar{X} < 12,71$, accepter H_0 .

Or les données relevées dans le tableau 4.1 indiquent que $\bar{x} = 13,03$. En conclusion, le test d'hypothèse nous informe qu'il faut rejeter H_0 , c'est-à-dire que la nouvelle méthode

d'enseignement est jugée efficace.

4.3 Signification statistique : accepter ou rejeter l'hypothèse nulle

4.3.1 Tests d'hypothèses simples

Pour un test d'hypothèses, la règle de décision peut se réécrire de la façon suivante :

$$\text{rejet de } H_0 \Leftrightarrow T \in W_\alpha \quad (4.3)$$

où α est fixé, T est une statistique de test et W_α est la région de rejet qui correspond à l'ensemble des valeurs qui conduisent à rejeter H_0 au profit de H_1 . Cette région dépend du choix du taux d'erreur de première espèce α . On détermine la région de rejet de la façon suivante : si H_0 est vraie, α correspond à la probabilité (sous l'hypothèse H_0) que la statistique T appartienne à la région de rejet W_α , et donc que H_0 soit rejetée à tort :

$$P(T \in W_\alpha | H_0) = \alpha \quad (4.4)$$

La loi de probabilité doit être entièrement connue sous H_0 (généralement, on ne sait pas modéliser H_1). Deux stratégies sont alors envisageables :

- soit on fixe le niveau de signification α , on en déduit le seuil de décision k et on réalise le test à partir de la statistique observée.
- soit on calcule, pour la statistique observée, la p-valeur. Rappelons que la p-valeur est la probabilité d'obtenir la même valeur (ou une valeur encore plus extrême) du test si l'hypothèse nulle est vraie. On décide alors de rejeter H_0 en comparant la p-valeur au niveau de signification α souhaité. Cette seconde stratégie est plus informative.

Prenons l'exemple d'une hypothèse H_0 sous laquelle la statistique de test suit une loi normale centrée réduite $\mathcal{N}(0; 1)$ et fixons l'erreur de première espèce, $\alpha = 0,05$. Notons que la distribution est symétrique et considérons le test bilatéral de seuil 0,05, soit une erreur à droite et à gauche de 0,025 (cf. figure 4.2). En utilisant les tables statistiques de l'écart réduit pour la loi normale $\mathcal{N}(0; 1)$, la règle de décision s'écrit maintenant :

$$\text{rejet de } H_0 \Leftrightarrow T \notin [-1,96; +1,96] \quad (4.5)$$

De la même manière, si on avait choisi $\alpha = 0,01$:

$$\text{rejet de } H_0 \Leftrightarrow T \notin [-2,56; +2,56] \quad (4.6)$$

Supposons maintenant que le test d'hypothèses ait donné une valeur de 3,1 pour T . L'hypothèse nulle H_0 est rejetée pour un seuil de 0,05 et un seuil de 0,01. L'intérêt de conserver la statistique de test T est de pouvoir calculer sa p-valeur, ici $p = 0,0019$.

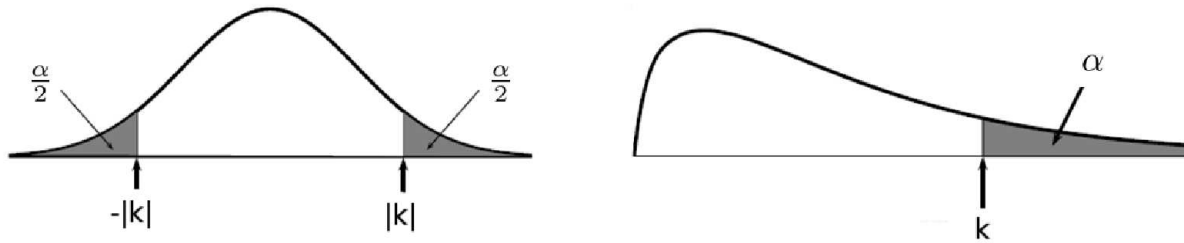


FIG. 4.2 – Zones de rejet pour un test bilatéral et unilatéral.

Décision \ Vérité	H_0	H_1	Total
	VN	FN	A
H_1	FP	VP	R
Total	m_0	m_1	m

 TAB. 4.3 – m tests d'hypothèses sont effectués simultanément. Le total des tests s'écrit : $m_0 + m_1 = A + R = m$. A et R représentent respectivement le nombre d'acceptations et de rejets.

4.3.2 Problème des comparaisons multiples

En détection de changements, lors de la comparaison de deux images, des milliers de tests d'hypothèses sont calculés simultanément (par exemple, si on effectue un test par voxel), ce qui pose le problème des comparaisons multiples. En ITD, dans le cas du suivi longitudinal, les images ont typiquement une taille d'environ $100 \times 100 \times 50$, et le cerveau occupe la moitié de l'image, soit 250 000 voxels. Il faudra donc faire 250 000 tests d'hypothèses. Soit m le nombre de tests d'hypothèses effectués simultanément. Dans le tableau 4.3, hormis m , A (le nombre de tests conduisant à accepter H_0) et R (le nombre de tests conduisant au rejet de H_0), toutes les autres quantités ne sont pas observables. Chaque test d'hypothèses a une erreur de première et de deuxième espèce. Soit $P(FP_i > 0)$ la probabilité du test i d'avoir au moins une fausse alarme. Il est toujours possible de contrôler l'erreur de première espèce de chaque test : $P(FP_i > 0) \leq \alpha$, $i = 1, \dots, m$, mais il est préférable de contrôler le nombre total de fausses alarmes (en anglais, *FamilyWise Error Rate* : FWER) ou le nombre de fausses alarmes parmi les détections (en anglais, *False Discovery Rate* : FDR). Le FDR représente la proportion attendue de rejets erronés parmi tous les rejets.

Certains auteurs ont donc tenté de contrôler le taux global d'erreur (FWER) [60, 1], défini dans le cadre d'un suivi longitudinal comme la probabilité de détecter à tort un ou plusieurs voxels sur l'ensemble des voxels analysés (*i.e.*, permet de contrôler l'erreur de première espèce sur l'ensemble des tests). La probabilité qu'au moins un test rejette H_0 alors que H_0 est vraie (*i.e.*, la probabilité d'avoir au moins une fausse alarme sur les m tests), s'écrit :

$$P(\text{au moins une fausse alarme sur les } m \text{ tests}) =$$

$$1 - P(\text{aucune fausse alarme sur les } m \text{ tests}) = 1 - (1 - \alpha)^m = \alpha_m \quad (4.7)$$

Si $m = 1$, on retrouve les résultats obtenus dans le cas d'un test d'hypothèses simple, à savoir $\alpha_m = \alpha$. Il ne faut donc pas chercher à tester les voxels séparément, mais plutôt à contrôler α_m avec une procédure tenant compte des différents tests d'hypothèses calculés. Si l'on teste chaque hypothèse nulle de manière indépendante, la procédure de Bonferroni, contrôlant le taux global d'erreur (FWER), est applicable [1]. Elle est basée sur l'inversion de l'équation (4.7) :

$$\alpha = 1 - (1 - \alpha_m)^{1/m} \quad (4.8)$$

En ne conservant que le premier terme de la série de Taylor, on obtient $\alpha = \frac{\alpha_m}{m}$. Donc, la correction de Bonferroni dit qu'il faut tester chaque hypothèse à un niveau de signification égal à $\frac{\alpha_m}{m}$. Cette procédure garantit que la probabilité d'avoir au moins une fausse alarme est inférieure à α_m : $P(FP > 0) \leq \alpha_m$. Cette méthode est basée sur une hypothèse d'indépendance entre voxel qui, en pratique, n'est pas vérifiée à cause de la corrélation spatiale des données. Pour prendre en compte cette corrélation, une alternative possible utilise la théorie des champs Gaussiens [60].

Les procédures contrôlant le FWER sont trop conservatrices (cf, figure 4.3), surtout quand le nombre de tests m devient grand [53]. Une alternative au FWER consiste à utiliser un critère permettant de contrôler le taux de faux positifs (FDR) [9, 10, 53]. La définition du taux de faux positifs (TFP) est assez intuitive. Soient FP le nombre total de faux positifs et R le nombre total de positifs (le nombre de tests rejetés). La TFP s'écrit alors de la manière suivante :

$$TFP = \begin{cases} \frac{FP}{R} & \text{si } R > 0 \\ 0 & \text{autrement} \end{cases} \quad (4.9)$$

Le FDR est alors défini comme l'espérance du taux de faux positifs (qui doit être inférieur à une valeur donnée α_m [9, 10, 53]) :

$$FDR = E[TFP] = E\left[\frac{FP}{R} | R > 0\right] P(R > 0) \leq \alpha_m \quad (4.10)$$

Avec le FDR, accepter 4 fausses détections sur 10 hypothèses nulles rejetées ($TFP = 2, 5$) est une erreur plus sérieuse que 20 fausses détections sur 100 hypothèses nulles rejetées ($TFP = 0, 2$). La procédure permet d'accepter plus d'erreurs quand beaucoup d'hypothèses nulles sont rejetées, mais moins quand peu d'hypothèses nulles sont rejetées. Dans la partie expérimentale, la procédure de Benjamini et Hochberg [9, 10] est utilisée pour contrôler le taux de faux positifs. D'autres procédures permettent d'estimer ce critère. Ces dernières sont présentées afin de retrouver la procédure de Benjamini et Hochberg. Cette relation repose sur le lien entre le FDR, et le positive FDR (pFDR) défini comme l'espérance du taux de faux positifs conditionnellement à l'existence d'au moins une détection : $pFDR = FDR/P(R > 0)$ [53]. Ces deux définitions sont étroitement liées, puisque, lorsque le nombre de tests m tend vers l'infini, $P(R > 0)$ tend vers 1. Si la loi de probabilité est connue sous H_0 , que les

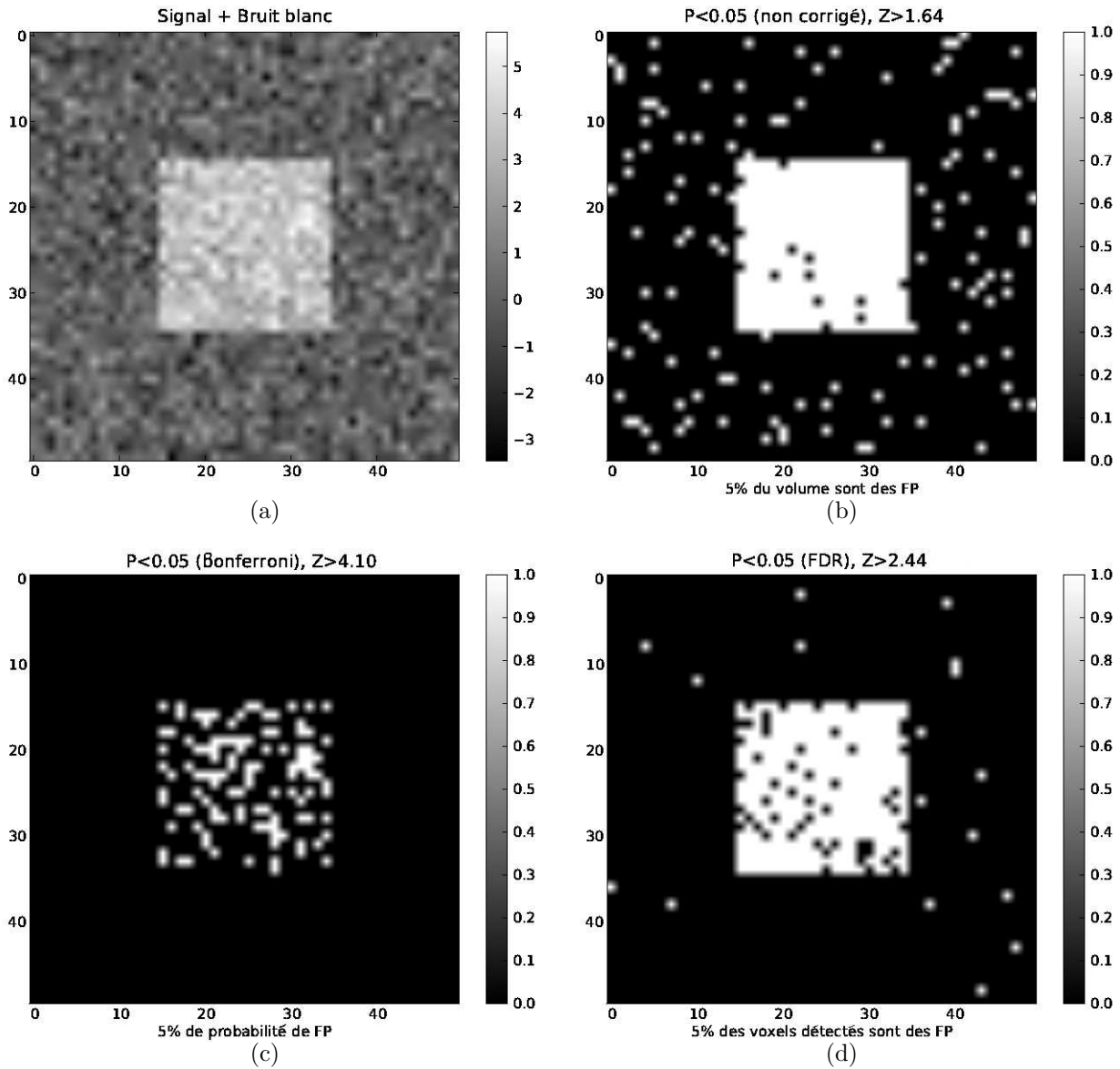


FIG. 4.3 – Comparaisons multiples : (a) carte des Z-scores d'un modèle signal constant égal à trois (zone de changements à détecter) + bruit blanc gaussien de moyenne nulle et de variance unité, (b) carte de détection après seuillage au niveau de signification $\alpha = 0,05$ sans prise en compte des corrections multiples, (c) carte de détection après seuillage avec la méthode de Bonferroni au niveau de signification $\alpha = 0,05$ et (d) carte de détection après seuillage avec la méthode FDR au niveau de signification $\alpha = 0,05$.

statistiques de test T_i sont indépendantes et que les p-valeurs p_i associées sont identiquement distribuées sous H_0 , alors il est possible de donner une formulation bayésienne au pFDR [53]. On définit pour cela des indicateurs pour chaque test d'hypothèse, $H = (H_1, H_2, \dots, H_m)$, où $H_i = 0$ quand la i ème hypothèse nulle est vraie et $H_i = 1$ quand la i ème hypothèse alternative est vraie. Pour une région de rejet W fixée (la même pour tous les tests), le pFDR s'écrit comme la probabilité conditionnelle que l'hypothèse nulle soit vraie sachant qu'on se trouve dans la zone de rejet :

$$pFDR = P(H = \mathbf{0}^m | T \in W^m) = \frac{P(H = \mathbf{0}^m)P(T \in W^m | H = \mathbf{0}^m)}{P(T \in W^m)} \quad (4.11)$$

A partir des statistiques de test $T = (T_1, T_2, \dots, T_m)$, il est possible de calculer les p-valeurs associées $p = (p_1, p_2, \dots, p_m)$. On fixe maintenant une région de rejet (cette fois définie sur les p-valeurs) $\Gamma = [0, \gamma]$ avec $\gamma \in [0, 1]$. Puisque sous H_0 les p-valeurs suivent une distribution uniforme sur l'intervalle $[0, 1]$, $P(p \in \Gamma^m | H = \mathbf{0}^m) = \gamma$ (cf. figure 4.4). Un estimateur de la probabilité $P(p \in \Gamma^m)$ est donné par la probabilité de comptage suivante (cf. figure 4.4) :

$$\hat{P}(p \in \Gamma^m) = \hat{P}(p \leq \gamma^m) = \frac{\#\{p_i \leq \gamma\}}{m} = \frac{R(\gamma)}{m} \quad (4.12)$$

où $\#\{A\}$ représente le cardinal de l'ensemble A et $R(\gamma)$ représente le nombre de rejets pour une zone de rejet fixée Γ . Finalement, les procédures permettant d'estimer le taux de faux positifs estiment $P(H = \mathbf{0}^m)$. Une stratégie d'estimation possible est présentée dans [53], et on obtient :

$$\widehat{FDR}(\gamma) = \frac{m\gamma P(H = \mathbf{0}^m)}{R(\gamma)} \quad (4.13)$$

La procédure de Benjamini et Hochberg pour des p-valeurs classées par ordre croissant $0 < p_{(1)} < p_{(2)} < \dots < p_{(m)}$, s'écrit [9, 10] :

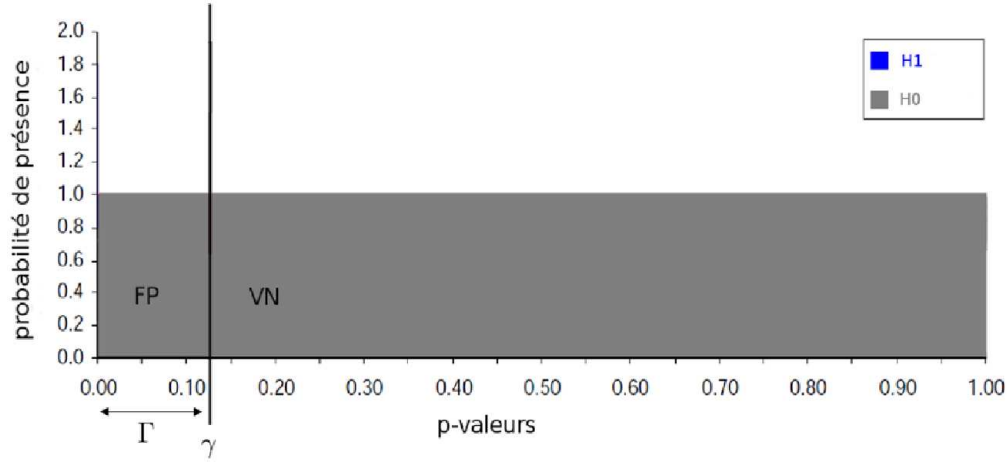
$$\hat{k} = \max \left\{ k : p_{(k)} \leq \frac{k}{m} \alpha_m, 0 \leq k \leq m \right\} \quad (4.14)$$

Cette procédure garantit un taux de faux positifs inférieur à α_m . Le seuil de décision FDR au niveau α_m est donné par : $T_{FDR} = T_{(\hat{k})}$. Toutes les hypothèses nulles $H_{(i)}, i = 1, \dots, \hat{k}$ sont alors rejetées. La relation entre cette procédure et la méthode précédente se fait en choisissant l'estimée la plus stricte pour $\hat{P}(H = \mathbf{0}^m)$, à savoir $\hat{P}(H = \mathbf{0}^m) = 1$. L'estimation du rang $\hat{l}$ entraînant le rejet des hypothèses nulles $H_{(i)}, i = 1, \dots, \hat{l}$ se fait de la manière suivante :

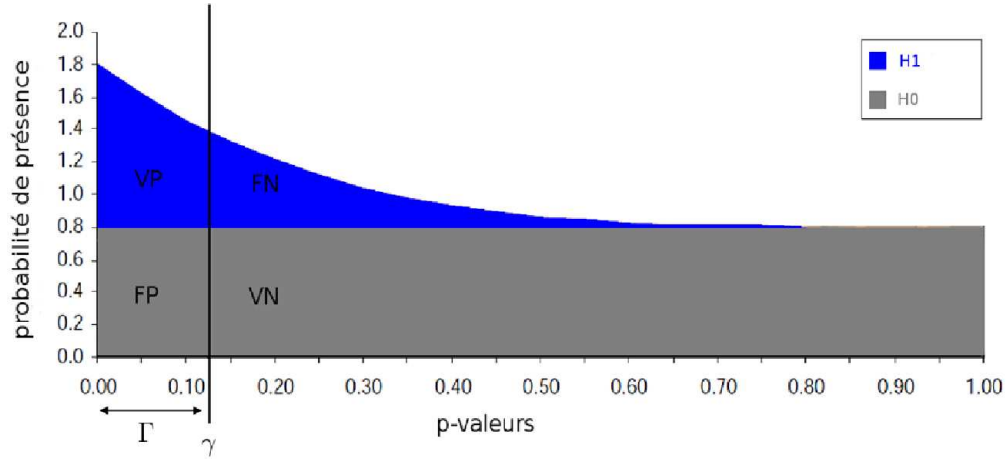
$$\hat{l} = \max \left\{ l : \widehat{FDR}(p_{(l)}) \leq \alpha_m, 0 \leq l \leq m \right\} \quad (4.15)$$

En remplaçant $\widehat{FDR}(p_{(l)}) = \frac{mp_{(l)}}{l}$, on obtient l'équivalence avec la procédure de Benjamini et Hochberg :

$$\hat{l} = \max \left\{ l : p_{(l)} \leq \frac{l}{m} \alpha_m, 0 \leq l \leq m \right\} \quad (4.16)$$



(a)



(b)

FIG. 4.4 – Deux configurations pour la distribution cumulée des p-valeurs sous H_0 et H_1 : (a) si l'hypothèse H_0 est vraie, les p-valeurs sont distribués uniformément dans l'intervalle $[0, 1]$ et (b) si l'hypothèse H_1 est vraie, les p-valeurs suivent une distribution inconnue, les valeurs proches de zéro étant les plus probables. En pratique, on observe le mélange de ces deux distributions. γ est la frontière de la zone de rejet sur les p-valeurs $\Gamma = [0, \gamma]$.

Notons qu'une estimation plus fine de $\hat{P}(H = \mathbf{0}^m)$ (*i.e.*, $\hat{P}(H = \mathbf{0}^m) < 1$) aurait donné un rang $\hat{l} \geq \hat{k}$. Un résultat de la procédure de Benjamini et Hochberg sur un cas synthétique est présenté en figure 4.3. Ce résultat est comparé avec la correction de Bonferroni et la prise de décision directe sans prise en compte des corrections multiples. Dans la suite de ce mémoire de thèse, nous utiliserons la procédure de Benjamini et Hochberg pour tenir compte des comparaisons multiples lors du calcul des cartes de détection.

Chapitre 5

La détection de changements en imagerie médicale

Sommaire

5.1	Etat de l'art sur la détection de changements en imagerie médicale	80
5.1.1	Les problématiques de recalage dans la détection de changements .	80
5.1.2	La détection de changements basée sur un recalage des images . .	82
5.2	Etat de l'art sur la détection de changements en imagerie de diffusion	83
5.2.1	Détection de changements sur des indices scalaires	84
5.2.2	Détection de changements multivariée sur des vecteurs de paramètres	87
5.2.3	Détection de changements multivariée sur la structure matricielle des tenseurs de diffusion	88

AVANT d'aborder la présentation des stratégies de détection de changements utilisées pendant cette thèse, une description des différentes approches de la littérature est proposée. Le spectre des applications relatives à la détection de changements étant large, nous nous intéressons dans ce chapitre uniquement au domaine de l'imagerie médicale. L'étude bibliographique commence par un bref descriptif général des modélisations statistiques élaborées pour la détection de changements en imagerie médicale. Cette description s'articule autour de la présentation des problématiques de recalage inhérentes à ces approches. Puis, l'étude bibliographique s'intéresse plus particulièrement aux techniques existantes en imagerie de diffusion. La nature multivariée des données, ainsi que la géométrie de l'espace défini par les tenseurs de diffusion rendent l'élaboration de méthodes de détection de changements délicate.

5.1 Etat de l'art sur la détection de changements en imagerie médicale

Nous proposons un aperçu des travaux réalisés sur ce sujet dans le domaine de l'imagerie médicale, d'autres applications étant présentées dans [42]. Les modèles mathématiques pour la détection de changements ont connu ces dernières années un développement considérable qui a donné lieu à de nombreux travaux en relation avec le suivi de pathologies diverses (maladies neurodégénératives : Alzheimer, sclérose en plaques ; tumeurs cérébrales). Ces modèles interviennent également dans l'étude du développement et du vieillissement du cerveau [43]. Comprendre les changements du cerveau et décrypter son fonctionnement sont les clés pour une détection précoce de diverses pathologies. Le traitement de l'image apporte ici une aide précieuse dans l'analyse (automatique et reproductible) tant qualitative que quantitative, des évolutions observées. La détection de changements met en jeu différentes techniques de traitement d'images : la segmentation, le recalage, la détection, la quantification et la modélisation des changements. La comparaison de zones préalablement segmentées est une démarche envisageable pour le suivi dans le temps de certaines pathologies (comme la sclérose en plaques). Les méthodes de segmentation visent à isoler la zone d'intérêt (lésions, tumeurs) des autres structures anatomiques. Cette segmentation peut être semi-manuelle ou automatique et a donné lieu à de nombreux travaux reposant sur les méthodes classiques de segmentation (déterministes, statistiques, orientées contours, croissance de régions, modèles déformables, ...). Une revue de ces différentes approches et des problèmes de validation associés peut être trouvée dans [5].

Notre bibliographie s'intéresse aux modélisations statistiques pour la détection de changements. Toutes ces modélisations supposent que les images à comparer sont dans un référentiel commun. Le recalage, utilisé pour mettre les images à comparer dans ce référentiel commun, est la clé de voute de ces stratégies. Cette section commence par un descriptif détaillé des problématiques de recalage. Ensuite, différentes approches possibles pour la détection de changements en imagerie médicale sont proposées. Elles sont basées sur un recalage préalable des images à comparer.

5.1.1 Les problématiques de recalage dans la détection de changements

La détection de changements à partir de modélisations statistiques dépend fortement de l'étape de recalage. Le recalage permet de mettre les images dans un même référentiel afin de réaliser une détection fine (sans segmentation explicite) des différences inter-images, liées au processus d'évolution de la zone étudiée [5, 13]. Le recalage est alors une étape importante qui est confrontée à de nombreuses difficultés liées à l'application.

Le recalage a pour objectif la mise en correspondance (dans un même espace 3D) des différentes informations anatomiques et fonctionnelles utilisées dans suivi de structures anatomiques normales ou pathologiques. Les différentes méthodes de recalage diffèrent par le

niveau de représentation de l'image utilisé (intensité, surfaces, lignes de crête, graphes, ...), les transformations géométriques considérées (rigide, affine, déformable) ainsi que par les critères de similarité guidant la mise en correspondance [32]. Des méthodes de recalage rigide, affine ou déformable utilisant l'information portée par les niveaux de gris et s'appuyant sur des fonctions de similarité quadratique, des mesures de corrélation ou l'information mutuelle sont ainsi classiquement utilisées [32]. Lors de la mise en correspondance de deux images, une fonction de coût (la fonction de similarité) est optimisée. Pour des images scalaires, à chaque itération de l'optimisation, une interpolation des données permet de rééchantillonner l'image déformée afin de mettre à jour la fonction de coût.

Comme cela a été illustré dans [13], le recalage rigide, voire affine, est souvent insuffisant, même lors du suivi dans le temps d'un même patient. Des déformations résiduelles de faibles amplitudes (locales ou globales) peuvent être observées, nécessitant le recours à des techniques de recalage permettant davantage de degrés de liberté. Un recalage précis est particulièrement important lorsque des différences subtiles doivent être détectées. Par ailleurs, le recalage des différentes modalités (IRM, ITD, imagerie spectroscopique de RMN) susceptibles d'intervenir dans la détection de changements est rendu délicat par la présence de zones en évolution. Des différences inter-images importantes (en intensité, texture, morphologie et topologie) peuvent ainsi être observées dans certaines zones (apparition de nécroses, d'œdèmes, ...). De plus, l'évolution d'une pathologie peut par ailleurs, par diffusion et croissance de la zone infectée, modifier et comprimer les tissus sains environnants. Ces différents effets ne sont pas pris en compte dans les fonctions de similarité classiques. Le recalage d'images prenant en compte des zones en évolution a fait l'objet d'un nombre limité de travaux. L'approche la plus simple consiste à écarter ces zones du domaine de définition du critère de similarité [51]. Cette approche suppose une segmentation préalable de la zone en évolution et ne prend pas en compte les compressions des tissus sains environnants. Une autre approche efficace, envisageable pour les transformations géométriques à faible nombre de degrés de liberté, consiste à utiliser des fonctions de similarité robustes [31]. La robustesse de ces critères aux données manquantes ou aberrantes permet de s'affranchir d'une segmentation préalable de la zone en évolution. Le problème devient notablement plus difficile si un recalage déformable est nécessaire (que ce soit dans le suivi d'un même patient ou dans le recalage inter-patients ou d'un patient avec un atlas). Il est en particulier important dans un recalage à grand nombre de degrés de liberté de distinguer les modifications induites par l'évolution de la pathologie des autres changements inter-images. Les approches récentes dans ce domaine (principalement utilisées pour le recalage de l'image d'un patient présentant une tumeur avec un atlas représentatif d'une population saine) s'appuient sur des modèles biomécaniques ou physio-pathologiques de l'évolution des tumeurs (prolifération-diffusion des cellules cancéreuses) [62, 63]. Ces modèles (qui nécessitent une segmentation préalable de la tumeur) peuvent être exploités pour représenter la contribution de l'évolution de la tumeur dans le champ de déformation global.

En ITD, l'application d'une transformation géométrique sur une image de tenseurs nécessite une étape supplémentaire de réorientation qui permet de conserver à la fois l'orientation et la forme des structures anatomiques [3] (les différentes stratégies d'interpolation et de

réorientation des tenseurs d'ordre 2 et d'ordre 4 seront respectivement présentées dans les chapitres 7 et 13). A chaque étape de l'optimisation, la réorientation des tenseurs complique la dérivation du gradient de la fonction de coût [14, 61]. Ainsi, beaucoup de recalage en ITD se font à partir d'une représentation scalaire des tenseurs ne nécessitant aucune réorientation (images d'indice invariant par rotation : FA, MD, ...) [12]. Une alternative possible consiste à réorienter les tenseurs après chaque étape de l'optimisation [2]. La transformation n'est alors pas prise en compte dans la définition de la fonction de coût.

5.1.2 La détection de changements basée sur un recalage des images

Les méthodes de détection de changements basées sur un recalage préalable des images reposent sur deux stratégies différentes : les méthodes qui utilisent le recalage pour gommer toutes les différences inter-images et qui effectuent la détection de changements par analyse du champ de déformation, et les méthodes d'analyse statistique des changements inter-images qui utilisent le recalage pour compenser tous les changements de « non intérêt » tout en conservant les changements d'intérêt comme des tumeurs ou des lésions en évolution.

L'analyse des champs de déformation

Pour cette catégorie d'approche, l'objectif du recalage est de compenser toutes les différences inter-images. A partir du recalage déformable de deux images d'un même patient (à une résolution suffisamment fine), l'analyse du champ de déformation estimé permet d'effectuer une détection de changements [55, 41, 23]. Les évolutions apparaissent alors comme des zones de déformation spécifique qui peuvent être caractérisées par leur propriétés différentielles (cartes de Jacobiens). Par exemple, dans [55, 41], une telle approche est utilisée pour détecter des lésions en évolution. Ces lésions en évolution apparaissent comme des contractions ou des dilatations dans le champ des déformations. Dans [23], une stratégie similaire permet de quantifier des variations locales de volume.

L'analyse des changements inter-images

Pour cette catégorie d'approches, l'objectif du recalage est de compenser tous les changements de « non intérêt ». L'analyse des changements inter-images consiste alors en une détection des zones d'évolution, éventuellement suivie d'une quantification des modifications [35, 27, 37, 13, 5, 50]. La détection peut être menée à partir de la différence inter-images seuillée [35] ou par des tests d'hypothèses prenant en compte une modélisation du bruit et du signal [27, 13, 4, 50]. Des éléments de contexte spatial peuvent également être intégrés dans la décision par filtrage des images initiales [12], par utilisation de voisinages étendus [12, 13] ou par utilisation d'approches *a contrario* [44]. Certaines approches peuvent intégrer des informations provenant de modalités IRM différentes [13, 4]. Dans ce contexte, l'imagerie du tenseur de diffusion fait actuellement l'objet de recherches actives en tant que modalité complémentaire. Par exemple, certaines études ont révélé l'infiltration de la substance blanche apparemment normale par du tissu tumoral [39, 56, 52]. Les modalités d'imagerie

anatomique classiques (T1, T2, FLAIR) ne permettent en effet pas d’observer les changements dans les régions de tissu uniforme. L’ITD fournit des informations nouvelles sur les micro-structures locales (fibres), en particulier dans la substance blanche [11, 54]. La prochaine section évalue l’apport de cette modalité pour la détection de changements.

5.2 Etat de l’art sur la détection de changements en imagerie de diffusion

Un nombre très limité de travaux méthodologiques ont pour l’heure évalué l’apport de l’ITD pour le suivi longitudinal. Dans [45], une méthode de détection et de quantification de l’infiltration de la substance blanche par des lésions tumorales de type gliomes est développée. La détection est réalisée par des techniques spécifiques de visualisation des données ITD. Une quantification de l’infiltration basée sur la fraction d’anisotropie et un modèle de mélange probabiliste est décrite dans [52]. Des modifications de la fraction d’anisotropie et de la diffusion moyenne des tissus, au sein de la tumeur et en périphérie de celle-ci, ont également récemment été mises en évidence sur un modèle animal (le rat) [28].

Rappelons que les données issues d’IRM de diffusion sont intrinsèquement multivariées puisque plusieurs mesures sont effectuées en chaque voxel de l’image. Ces mesures quantifient la diffusion des molécules d’eau dans des directions de gradient spécifiées par le protocole d’acquisition et uniformément distribuées sur la sphère unité. Une analyse statistique menée directement sur ces mesures de diffusion est envisageable. Toutefois, ce volume de données (parfois des centaines de directions de gradient différentes sont acquises) est fréquemment synthétisé en un nombre restreint de paramètres. En supposant le processus de diffusion gaussien [40], la diffusion (qui peut être anisotrope) est alors caractérisée par une matrice de covariance du déplacement des particules (appelée également tenseur de diffusion). Cette matrice de covariance doit être symétrique et définie positive. Des tests statistiques peuvent alors être élaborés soit sur la structure matricielle du tenseur de diffusion, soit sur les éléments indépendants du tenseur de diffusion vectorisés. De plus, l’espace des matrices symétriques, définies positives n’est pas un espace vectoriel. Les tests statistiques peuvent alors intégrer cette contrainte. Finalement, une réduction de dimension supplémentaire permet de calculer des paramètres scalaires invariants par rotation (en particulier la FA et la MD). La mesure de ces paramètres est reproductible lors d’examens successifs, dans les mêmes conditions d’acquisition [21]. Des tests statistiques classiques (par exemple un T-test) peuvent alors être mis en œuvre sur ces quantités scalaires.

Nous proposons dans cette section un état de l’art sur les techniques existantes pour la détection de changements en imagerie de diffusion. Un résumé synthétique de ces approches est présenté dans le tableau 5.1. Deux grandes familles de tests statistiques permettent d’effectuer une détection de changements : d’un côté, les tests paramétriques où une hypothèse est formulée sur la distribution des données, et de l’autre, les tests non-paramétriques où aucun *a priori* n’est requis sur la distribution des données. Ces techniques sont présentées dans la suite selon la classification suivante : 1- les méthodes de détection de changements

calculées sur des indices scalaires, 2- les méthodes de détections de changements multivariées sur des vecteurs de paramètres et 3- les méthodes de détections de changements multivariées sur la structure matricielle des tenseurs de diffusion.

5.2.1 Détection de changements sur des indices scalaires

Tests paramétriques

Des indices scalaires tels que la FA ou la MD permettent de caractériser le modèle de diffusion. La MD caractérise la diffusivité du milieu, alors que la FA est directement reliée au degré d'organisation des tissus. Un test du rapport de vraisemblance généralisé (ou, en anglais, *Generalized Likelihood Ratio Test*, GLRT, initialement proposé dans [13]) [12], un T-test [59], et un F-test [34] permettent la quantification des changements au niveau de ces images d'indice (FA et MD). L'utilisation d'un test d'hypothèses multivarié (le test de Hotelling T^2) rend possible la comparaison simultanée de plusieurs indices [34]. Le vecteur d'intérêt est de dimension deux et contient la FA et la MD. Ces approches paramétriques ont une limite commune : elles n'exploitent qu'une partie de l'information contenue dans les tenseurs de diffusion et ne permettent ainsi de détecter que des changements de FA, de MD ou une combinaison de FA et de MD. Il est également intéressant de remarquer que la relation non triviale entre le signal de diffusion (dont le bruit est supposé Ricien) et la FA ou la MD fait qu'il est difficile de connaître la nature du bruit sur ces deux indices. Si l'on fait l'hypothèse que les intensités du signal DWI suivent une distribution normale multivariée, il a été montré que la FA suit asymptotiquement une distribution normale quand le nombre de directions de gradient tend vers l'infini [15].

Test non-paramétrique

Un test non-paramétrique de permutation permet de s'affranchir de l'hypothèse sur la distribution des données. La méthode proposée dans [17] fait référence dans la littérature et est présentée ici de manière détaillée. Les échantillons contiennent les coefficients de diffusion apparente. En chaque voxel de deux images recalées d'un même patient, les deux échantillons à comparer sont appariés par direction de gradient (on fait l'hypothèse que le même protocole est utilisé pour l'acquisition des deux images). On veut tester si ces deux échantillons sont identiques en utilisant comme grandeur caractéristique de l'échantillon la FA (obtenue pour chaque échantillon après estimation du tenseur de diffusion), c'est-à-dire : $H_0 : FA_1 = FA_2$ vs. $H_1 : FA_1 \neq FA_2$. Pour ce faire, la différence des valeurs de FA des échantillons à comparer est calculée : $\hat{\theta} = FA_1 - FA_2$. Il s'agit donc de tester si $\hat{\theta}$ est statistiquement différent de zéro. Une procédure de permutation est utilisée pour obtenir une estimation de la distribution des données, et, finalement, décider si l'hypothèse nulle doit être acceptée ou de rejetée. En chaque voxel de l'image, les labels (*i.e.*, échantillons d'appartenance) des coefficients de diffusion apparente sont permutés aléatoirement. La distribution sous H_0 est construite à partir des différences de FA obtenues pour l'ensemble des permutations : $\hat{\theta}_i^* = FA_{1;i}^* - FA_{2;i}^*$ pour la i ème permutation. Bien souvent, l'ensemble des permutations possibles est bien trop grand pour être entièrement exploré. Un échantillon aléatoire de cet ensemble est alors

Détection de changements en IRM de diffusion

La détection de changements en imagerie médicale

Test	Niveau de représentation de la diffusion	Type de test	Niveau de représentation au niveau de l'image
Histogramme [19]	- indice scalaire	/	fibre
T-test T-test apparié [12]	- coefficients de diffusion apparente	paramétrique	voxel
T-test F-test [59, 34]	- indice scalaire	paramétrique	voxel-groupe
GLRT [12]	- indice scalaire - inadéquation des données par rapport au modèle - tenseur vectorisé	paramétrique	région
Hotelling T^2 [34, 48, 59]	- vecteur d'indices scalaire - tenseur vectorisé	paramétrique	voxel-groupe
LRT Log-LRT [48, 49, 46]	- tenseur - vecteurs propres - valeurs propres	paramétrique	voxel-groupe
Différence [17]	- indice scalaire	non-paramétrique	voxel
Hotelling T^2 [59]	- tenseur vectorisé	non-paramétrique	/
Test des signes [12]	- coefficients de diffusion apparente	non-paramétrique	voxel
Test des rangs signés de Wilcoxon [12]	- coefficients de diffusion apparente	non-paramétrique	voxel
Watson bipolaire [47]	- directions principales	paramétrique	voxel-groupe
Cramér Log-Cramér [59]	- tenseur vectorisé	non-paramétrique	voxel-groupe

TAB. 5.1 – Récapitulatif des différentes méthodes de détection de changements en imagerie de diffusion classées par type de test.

considéré. On réalise ainsi N permutations à partir desquelles une p-valeur est estimée :

$$p = \frac{\#\{|\hat{\theta}_i^*| \geq |\hat{\theta}|\}}{N}, \quad i \in [1, N],$$

où $\#\{A\}$ représente le cardinal de l'ensemble A . On obtient une carte de p-valeurs sans permutation, ainsi que N cartes de p-valeurs associées à chaque permutation. Ces $N + 1$ cartes sont seuillées en utilisant un seuil arbitraire afin d'obtenir des clusters (*i.e.*, groupes de voxel connexes). Ensuite, une p-valeur corrigée pour les comparaisons multiples est calculée pour chaque cluster en utilisant une analyse « supra threshold cluster » [30]. Un histogramme est calculé à partir du plus grand (en taille) cluster de chacune des N cartes de permutation seuillées. Une p-valeur corrigée est alors estimée pour chaque cluster de la carte sans permutation :

$$p = \frac{\#\{\hat{\eta}_i^* \geq \hat{\eta}\}}{N}, \quad i \in [1, N],$$

où $\hat{\eta}$ représente la taille du cluster dans la carte sans permutation et les $\hat{\eta}_i^*$ représentent les tailles des N plus grands clusters utilisées pour construire l'histogramme. Lors du calcul des p-valeurs, la combinaison de ces deux étapes permet notamment de s'affranchir des changements pouvant être introduits par des artefacts d'acquisition (bruit, erreur de recalage, ...) et de toute hypothèse sur les distributions sous-jacentes.

Visualisation basée sur un *a priori* anatomique

Les tests statistiques peuvent également intégrer un *a priori* sur les fibres de la substance blanche. Il s'agit toujours d'approches scalaires avec une définition différente de la population comparer. Les fibres sont calculées en utilisant un algorithme de tractographie permettant de suivre la direction principale de diffusion en chaque voxel de l'image. La notion de « fibres » est un peu abusive puisqu'il s'agit en fait d'un modèle. L'ensemble des « fibres » obtenues par tractographie peuvent être regroupées en faisceaux de fibres. Une fois cette opération terminée, une analyse quantitative sur différents indices (FA, MD) calculés le long des faisceaux de fibres de la substance blanche peut être menée. Cette approche a montré son intérêt médical [18]. De nombreux travaux dans ce domaine s'appuient sur l'utilisation d'un modèle (*i.e.*, un atlas) des faisceaux de la substance blanche [25, 33, 29]. On fait ainsi l'hypothèse que toutes les images de diffusion peuvent être transformées vers un espace commun. Néanmoins, cette étape peut être mise en défaut lorsque de fortes distorsions géométriques sont présentes dans les images (par exemple lorsque des tumeurs sont imagées). De plus, ces approches ne sont pas appropriées aux extrémités des faisceaux où la variabilité est importante (par exemple quand les fibres sont en éventail). Pour éviter l'utilisation d'un modèle des faisceaux de fibres, les fibres d'un faisceau peuvent être alignées pour construire un modèle de fibre moyenne [19]. En chaque échantillon de ce modèle, les tenseurs dans une section de faisceau sont réorientés en appliquant la transformation estimée lors de l'alignement des fibres et un tenseur moyen est calculé en utilisant une métrique riemannienne. Cette méthode respecte les propriétés des tenseurs de diffusion. A partir des tenseurs de diffusion moyens, des indices scalaires sont calculés et tracés en fonction de leur position spatiale le long du faisceau. Une telle approche semi-automatique (interprétation visuelle des

résultats) permet d'analyser un nombre limité de faisceaux dans un temps raisonnable. Pour analyser automatiquement l'ensemble du cerveau, des tests statistiques doivent être mis en place. De plus, cette méthode se base sur des indices scalaires issus des tenseurs de diffusion, l'information utilisée ne reflète alors qu'une partie de l'information contenue dans le tenseur de diffusion, ce qui peut éventuellement empêcher la détection de certaines modifications. Idéalement, l'étude statistique portera directement sur l'ensemble du tenseur de diffusion, soit représentée sous une forme vectorisée, soit dans sa représentation naturelle sous forme matricielle.

5.2.2 Détection de changements multivariée sur des vecteurs de paramètres

Tests paramétriques

Une première approche consiste à comparer directement les images de coefficients de diffusion apparente. Cette comparaison a été menée à l'aide d'un T-test ou d'un T-test apparié par direction de gradients (pour obtenir le même jeu de directions de gradient dans les deux examens à comparer après recalage, une interpolation sphérique des coefficients de diffusion apparente est utilisée) [12]. Un test du rapport de vraisemblance généralisé, ou en anglais, *Generalized Likelihood Ratio Test* (GLRT) a également été proposé pour mesurer l'inadéquation des données par rapport au modèle de diffusion considéré [12].

L'organisation de la substance blanche peut également être caractérisée par le modèle de diffusion utilisé. Un test paramétrique basé sur une distribution de Watson bipolaire opérant sur les directions principales de diffusion (vecteurs propres principaux des tenseurs de diffusion) permet de quantifier les changements d'organisation de la substance blanche [47]. Si l'on souhaite comparer les six éléments indépendants du tenseur de diffusion, il est possible de les regrouper dans un vecteur de paramètres à six dimensions et d'effectuer un test de Hotelling T^2 [48]. Le test de Hotelling T^2 fait l'hypothèse que le vecteur de paramètres en entrée suit une loi normale multivariée et que les deux échantillons à comparer ont la même matrice de covariance. Ces hypothèses semblent en accord avec les résultats issus de simulations de Monte-Carlo [36] et les résultats théoriques [8]. De la même manière, un autre test paramétrique permet de comparer deux échantillons de tenseurs vectorisés : le GLRT [12].

Tests non-paramétriques

Afin d'éviter les hypothèses inhérentes aux tests paramétriques précédents, différentes solutions ont été envisagées. Les tests des signes et des rangs signés de Wilcoxon permettent de comparer de manière non-paramétrique deux échantillons de coefficients de diffusion apparente appariés par direction de gradient [12]. Le test des signes prend en compte la différence de signe entre les éléments appariés en testant si les différences positives et négatives sont équiprobables. Dans le test des rangs signés de Wilcoxon, les différences de signe sont triées par rapport à la valeur absolue des différences entre éléments appariés. Cette fois, le test vérifie si la somme des rangs affectés d'un signe positif est égale à la somme des rangs affectés d'un signe négatif. Ces méthodes n'exploitent pas complètement l'information contenue dans

les amplitudes des différences observées.

Lorsque l'échantillon contient directement des tenseurs de diffusion sous leurs formes vectorisées, une permutation multivariée consiste à effectuer des sous-tests sur chaque élément du vecteur observé. Les p-valeurs obtenues pour chaque sous-test sont alors combinées à l'aide d'une fonction d'inclusion [59]. Une autre solution non-paramétrique a été proposée dans [59] : les test de Cramér. Ce dernier test, moins connu, vérifie si les deux échantillons sont égaux en distributions. Notons $d_{i,j}$ la moyenne de toutes les distances euclidiennes entre les tenseurs vectorisés des échantillons i et j . Le test de Cramér s'écrit alors : $T_{\text{Cramér}} = d_{1,2} - 1/2(d_{1,1} + d_{2,2})$. Tous ces tests multivariés non-paramétriques considèrent les tenseurs de diffusion comme des observations alors que les réelles observées sont les valeurs du signal pondéré en diffusion. De plus, permuter directement les éléments du tenseur un à un a beaucoup moins de sens que de permuter le signal observé dans une direction de gradient donnée. Finalement, toutes ces méthodes mettent en œuvre un test multivarié non-paramétrique sur les éléments du tenseur vectorisés (idem pour les tests paramétriques), sans contrainte sur ces éléments. Ces tests ne respectent donc pas nécessairement la géométrie de l'espace des tenseurs de diffusion.

Le cadre statistique de détection de changements ainsi défini omet la propriété fondamentale de positivité des tenseurs de diffusion. Récemment, des métriques non-euclidiennes ont été proposées pour respecter la géométrie de l'espace des tenseurs d'ordre 2 symétriques et définis positifs [6, 38]. Afin de respecter cette géométrie, le test de Cramér a été adapté à un cadre non-euclidien spécifique défini sur l'espace log-euclidien [59] : le test de Log-Cramér. Avec une métrique log-euclidienne [6], des calculs euclidiens sont menés dans le domaine des logarithmes des matrices, ce qui est suffisant pour garantir au résultat d'être un tenseur de diffusion. Par conséquent, la prise en compte de la métrique log-euclidienne dans le test de Log-Cramér est immédiate. Le test ainsi défini respecte la géométrie de l'espace des tenseurs de diffusion sans réel surcoût de calcul (il suffit de calculer le logarithme des tenseurs de diffusion).

5.2.3 Détection de changements multivariée sur la structure matricielle des tenseurs de diffusion

Tests paramétriques

Lors d'une détection de changements, considérer les tenseurs de diffusion vectorisés est une limitation importante. En effet, les différences mesurées peuvent alors provenir d'un changement des propriétés de diffusion (variation des valeurs propres du tenseur de diffusion), d'un changement d'organisation des structures sous-jacentes (variation de l'orientation du tenseur de diffusion) ou une combinaison de ces deux phénomènes. Ces tests ne permettent pas de dissocier l'origine des modifications observées. Dans un cadre applicatif donné, il peut être intéressant de différencier l'origine des changements. Par exemple, en sclérose en plaques, seules des modifications de la diffusion sont attendues. Un test permettant de quantifier de tels changements, en considérant les autres causes d'évolution comme non significatives, est

un atout certain pour l'étude de la pathologie. A cette fin, dans [48, 49, 46], un test du rapport de vraisemblance, ou en anglais, *Likelihood Ratio Test* (LRT) est directement construit sur la matrice symétrique représentant le tenseur de diffusion. Il est alors aisé de découpler les informations de diffusion et d'orientation en utilisant une décomposition spectrale. La construction de statistiques sur les matrices de valeurs propres ou de vecteurs propres assure alors une meilleure description de la statistique. Les tests d'hypothèses ainsi développés permettent, par exemple, d'étudier avec une plus grande sensibilité des pathologies où la diffusion est altérée et l'orientation des fibres conservée [48, 49, 46]. La positivité des tenseurs est garantie en utilisant la métrique log-euclidienne (Log-LRT).

Cette approche a été proposée dans le cadre d'études de groupes. Soient deux groupes contenant des images de diffusion à comparer. La première étape consiste à ramener toutes ces images vers un espace commun. Ainsi, en chaque voxel de l'image, les deux échantillons à comparer sont directement construits en considérant les tenseurs à cet emplacement dans chacun des groupes. L'étape de construction des échantillons à comparer est moins évidente lors du suivi dans le temps d'un même patient. Dans la littérature, la méthode la plus utilisée considère un voisinage spatial de taille et de forme fixées autour du voxel d'étude [13, 12]. Peu de méthodes ont été proposées pour comparer des échantillons de tenseurs de diffusion à comparer dans la cadre du suivi longitudinal d'un patient. Cette problématique est abordée au travers des différentes contributions de cette thèse. Le prochain chapitre propose ainsi une description structurée des tests statistiques multivariés utilisés dans la suite pour le suivi longitudinal en IRMd. Les tests du rapport de vraisemblance (Log-LRT) seront utilisés pour comparer des tenseurs de diffusion [48, 49, 46].

Solutions alternatives

Signalons également que dans la littérature, d'autres méthodes ont été imaginées pour aborder le problème de détection de changements en imagerie de diffusion. L'objectif de ces méthodes est d'éviter l'utilisation de métriques complexes respectant la géométrie de l'espace des tenseur de diffusion. Le principe de ces méthodes est donc de se ramener sur un espace où une distance euclidienne est adaptée. Pour ce faire, ces méthodes réduisent la dimension du problème initial en passant d'un espace à six dimensions, à un espace de dimension plus faible qui permet tout de même de capturer la variabilité dans les données. Deux approches ont été proposées, dont voici un bref descriptif :

- **Apprentissage de la variété** où réside les tenseurs [57]. En utilisant l'Isomap, la structure de la variété sous-jacente aux données d'ITD est déterminée. Ensuite, une analyse statistique est effectuée sur les variétés estimées après aplanissement.
- **Apprentissage par noyaux** de la distribution des tenseurs grâce à des méthodes à noyaux [26]. Les tenseurs sont représentés dans un espace de plus grande dimension, l'espace de Hilbert à noyau reproduisant ou, en anglais, reproducing kernel Hilbert space (RKHS). Pour apprendre la distribution des tenseurs, deux analyses en composantes principales (ACP) modifiées ont été proposées. Rappelons tout d'abord que

l'ACP est une technique d'analyse statistique linéaire fréquemment utilisée en reconnaissance de formes. L'objectif de l'ACP est de déterminer une base orthogonale ajustant au mieux les données selon un critère de variance. Pour ce faire, la méthode détermine les directions successives de variance maximale qui correspondent également à l'optimum d'un critère géométrique. En imagerie du tenseur de diffusion, une analyse statistique non-linéaire est souhaitable. L'ACP modifiée se fait alors, soit dans l'espace des tenseurs avec une métrique riemannienne (l'analyse en géodésiques principales [22]), soit à partir de l'apprentissage par noyaux dans un nouvel espace de caractéristiques (l'ACP à noyaux [26]). L'ACP à noyaux est une forme d'analyse ACP non-linéaire. L'idée est de transformer les données initiales vers un espace de caractéristiques dans lequel la distribution des données est de nouveau considérée comme linéaire et où une ACP classique peut être appliquée. Dans [58], une analyse morphométrique basée sur une ACP à noyaux est conduite sur des patients atteints de schizophrénie (méthode MBM : *Manifold Based Morphometry*).

Tests statistiques adaptés aux tenseurs de diffusion

Sommaire

6.1	Loi multinormale pour les matrices symétriques définies positives	92
6.1.1	Distribution multinormale vectorisée	92
6.1.2	Distribution multinormale matricielle pour les matrices symétriques	93
6.1.3	Distribution multinormale matricielle pour les matrices symétriques définies positives	94
6.2	Tests statistiques multivariés sur des populations de tenseurs .	96
6.2.1	Le test du rapport de vraisemblance à un échantillon (RV1)	96
6.2.2	Le test du rapport de vraisemblance à deux échantillons (RV2) . .	101
6.2.3	La distribution empirique sous l'hypothèse nulle	108

AFIN de comparer des images d'IRMd intrinsèquement multivariées (au moins six, voire des centaines de directions de mesures sont disponibles en chaque voxel de l'image), des indices scalaires caractérisant le phénomène physique de diffusion ont été utilisés en premier lieu. Toutefois, ces indices offrent la possibilité de détecter uniquement des changements spécifiques. Par exemple, la FA ne permet pas de quantifier un changement d'orientation des micro-structures sous-jacentes. Le tenseur de diffusion contient une information plus riche qui permet à elle seule de caractériser complètement la diffusion en trois dimensions. Les approches proposées dans ce chapitre ont pour objectif de comparer directement les tenseurs entre eux. De telles techniques nécessitent de prendre en compte la nature même des tenseurs de diffusion qui sont des matrices symétriques définies positives. Tout d'abord, une loi multinormale pour les tenseurs de diffusion est présentée. Un cadre statistique général de détection de changements est ensuite défini [46]. Des tests paramétriques multivariés sont présentés [48, 49, 46]. Ils autorisent la détection de changements de diffusion, de changements d'orientation, ou une combinaison de ces deux phénomènes.

6.1 Loi multinormale pour les matrices symétriques définies positives

Parmi tous les modèles, nous focalisons notre étude sur les lois permettant de modéliser les données de diffusion. Dans la littérature, deux méthodes ont principalement été utilisées pour construire une distribution sur des matrices symétriques définies positives. Dans [24], les éléments de la matrice sont réarrangés dans un vecteur. Ce vecteur est ensuite modélisé par une loi multinormale. Dans [16], aucune opération de vectorisation n'est requise et la distribution multinormale est construite directement pour chaque élément de la matrice. Les distributions multinormales pour des matrices symétriques sous leurs formes vectorisées puis matricielles sont présentées dans cette section. En imagerie de diffusion, lors de suivis longitudinaux, les tenseurs sont des matrices définies positives. Une distribution multinormale tenant compte de cette propriété est également présentée. Les variables aléatoires sont indicées avec un « $\sim$ ». Par exemple, $\tilde{X}$ est une variable aléatoire de réalisation observée X . Dans la suite, nous confondrons, pour la simplicité des notations, les variables aléatoires avec leurs réalisations (lorsqu'il n'y a pas d'ambiguïté).

6.1.1 Distribution multinormale vectorisée

Soit $\mathbf{D}$ une matrice de taille $p \times p$. On note $diag(\mathbf{D})$ le vecteur colonne formé avec les $p \times 1$ éléments situés sur la diagonale de $\mathbf{D}$ et $antidiag(\mathbf{D})$ le vecteur colonne formé avec les $p(p-1)/2 \times 1$ éléments situés dans la matrice triangulaire supérieure de $\mathbf{D}$ (sans les termes diagonaux) pris colonne par colonne. On obtient le vecteur $vec(\mathbf{D})$ de taille $p(p+1)/2 \times 1$ en concaténant les éléments diagonaux et anti-diagonaux de $\mathbf{D}$. Par exemple, lorsque $p = 3$, on obtient :

$$vec(\mathbf{D}) = \begin{pmatrix} diag(\mathbf{D}) \\ \sqrt{2} antidiag(\mathbf{D}) \end{pmatrix} = \begin{pmatrix} d_{11} \\ d_{22} \\ d_{33} \\ \sqrt{2}d_{12} \\ \sqrt{2}d_{13} \\ \sqrt{2}d_{23} \end{pmatrix} \quad (6.1)$$

L'opérateur $vec(\cdot)$ permet de passer de la norme de Frobenius sur les matrices symétriques à une mesure euclidienne sur des vecteurs :

$$\begin{aligned} \|\mathbf{D}\|^2 &= tr(\mathbf{D}^2) \\ \|vec(\mathbf{D})\|^2 &= vec(\mathbf{D})^T vec(\mathbf{D}) \end{aligned} \quad (6.2)$$

Pour une matrice $\mathbf{D}$ symétrique, ces deux métriques donnent le même résultat. Notons $Sym(p)$ l'espace des matrices symétriques de dimension p et $Sym^+(p)$ l'espace des matrices symétriques définies positives de dimension p . $\mathbf{D} \in Sym(p)$ suit une loi multinormale définie pour les matrices symétriques vectorisées de moyenne $\mathbf{M} \in Sym(p)$ et de variance $\Sigma_{q \times q} \in Sym^+(q)$ (avec $q = p(p+1)/2$), si :

$$vec(\mathbf{D}) \sim \mathcal{N}_p(vec(\mathbf{M}), \Sigma_{q \times q}) \quad (6.3)$$

6.1.2 Distribution multinormale matricielle pour les matrices symétriques

Une autre manière de modéliser les tenseurs de diffusion consiste à les considérer directement comme des matrices aléatoires. Soit $\mathbf{D}$ une matrice aléatoire, dont les éléments d_{ij} sont p^2 variables aléatoires. Cette représentation permet d'étudier directement les p^2 éléments d_{ij} , mais également les valeurs propres et les vecteurs propres de ces matrices. Il faut déterminer la loi de probabilité conjointe de chacune de ces variables aléatoires. Il existe plusieurs ensembles classiques de matrices aléatoires qui ont été largement étudiés. Comme dans [48, 46], nous utilisons l'Ensemble Gaussien Orthogonal (GOE) pour représenter $\mathbf{D}$. Cet ensemble est défini pour des matrices symétriques réelles $p \times p$. Or par construction, les tenseurs de diffusion sont des matrices symétriques. Comme le nom GOE le suggère, la distribution est invariante par transformations orthogonales. $\mathbf{D} \in Sym(p)$ suit une loi multinormale centrée réduite issue de l'ensemble GOE $\mathcal{N}_{pp}(\mathbf{0}, \mathbf{I}_p)$, si sa distribution est de la forme :

$$f_{\mathbf{D}}(\mathbf{D}) = \frac{1}{(2\pi)^{q/2}} \exp\left(-\frac{1}{2}tr(\mathbf{D}^2)\right) \quad (6.4)$$

Soient d_{ij} les éléments de $\mathbf{D}$ ($d_{ij} = d_{ji}$). Le terme dans l'exponentielle de (6.4) est égal à la somme des carrés suivante :

$$tr(\mathbf{D}^2) = \sum_{i,j=1}^p d_{ij}^2 = \sum_{i=1}^p d_{ii}^2 + 2 \sum_{i<j}^p d_{ij}^2 = \sum_{i=1}^p d_{ii}^2 + \sum_{i<j}^p (\sqrt{2}d_{ij})^2 \quad (6.5)$$

La densité de probabilité ainsi définie est le produit de $p(p+1)/2$ lois normales univariées. Les p termes diagonaux de $\mathbf{D}$ suivent une loi normale centrée réduite $\mathcal{N}(0, 1)$ et les $q - p = p(p-1)/2$ termes anti-diagonaux suivent une loi normale centrée de variance $1/2$ $\mathcal{N}(0, 1/2)$. Ainsi, lorsque $p = 3$, $\mathbf{D}$ suit une loi de la forme :

$$\mathbf{D} \sim \begin{pmatrix} \mathcal{N}(0, 1) & \mathcal{N}(0, 1/2) & \mathcal{N}(0, 1/2) \\ * & \mathcal{N}(0, 1) & \mathcal{N}(0, 1/2) \\ * & * & \mathcal{N}(0, 1) \end{pmatrix} \quad (6.6)$$

où chaque élément de $\mathbf{D}$ est indépendant et l'étoile signifie copie par symétrie. La somme des carrés dans (6.5) correspond donc à la somme des carrés de q variables normales centrées réduites indépendantes. On en déduit ainsi que la somme des carrés suit une loi du khi-deux à q degrés de liberté : $tr(\mathbf{D}^2) \sim \chi^2(q)$.

Grâce au changement de variable $\mathbf{D} \rightarrow \mathbf{D}' : \mathbf{D}' = \phi(\mathbf{D}) = \Sigma^{1/2}\mathbf{D}\Sigma^{1/2} + \mathbf{M}$, on calcule l'expression générale de la densité multinormale pour des matrices symétriques pour $\mathbf{M} \in Sym(p)$ et $\Sigma \in Sym^+(p)$. Le passage de $\mathbf{D}$ à $\mathbf{D}'$ ainsi défini est biunivoque et est donné par :

$$f_{\mathbf{D}} = \frac{\phi^{-1}(f_{\mathbf{D}'})}{|det(J)|} \quad (6.7)$$

où $\det(J)$ est le déterminant du Jacobien issu de la transformation $\phi : J(\mathbf{D} \rightarrow \mathbf{D}') = |\Sigma|^{-(p+1)/2}$. En appliquant le changement de variable dans (6.5), on obtient [46] :

$$\text{tr}(\mathbf{D}^2) = \text{tr} [(\Sigma^{-1/2}(\mathbf{D}' - \mathbf{M})\Sigma^{-1/2})^2] = \text{tr} [((\mathbf{D}' - \mathbf{M})\Sigma^{-1})^2] \quad (6.8)$$

On arrive à la définition générale suivante : $\mathbf{D}' \in \text{Sym}(p)$ suit une distribution multinormale $\mathcal{N}_{pp}(\mathbf{M}, \Sigma)$, avec $\mathbf{M} \in \text{Sym}(p)$ et $\Sigma \in \text{Sym}^+(p)$, si $\mathbf{D}'$ a pour densité [46] :

$$f_{\mathbf{D}'}(\mathbf{D}'; \mathbf{M}, \Sigma) = \frac{1}{(2\pi)^{q/2} |\Sigma|^{(p+1)/2}} \exp \left(-\frac{1}{2} \text{tr} [((\mathbf{D}' - \mathbf{M})\Sigma^{-1})^2] \right) \quad (6.9)$$

Considérons maintenant un cas particulier de l'équation (6.9). On définit le changement de variables $\mathbf{D}'' = \sigma^2 \mathbf{D} + \mathbf{M}$, où $\sigma > 0$ représente un facteur d'échelle global et $\mathbf{D} \sim \mathcal{N}_{pp}(\mathbf{0}, \mathbf{I}_p)$. Dans ce cas, les composantes de la matrice aléatoire $\mathbf{D}''$ sont indépendantes (*i.e.*, non corrélées) car $\Sigma = \sigma^2 \mathbf{I}_p$ est diagonale. On obtient la définition suivante : soit $\mathbf{D}'' \in \text{Sym}(p)$ suit une distribution multinormale $\mathcal{N}_{pp}(\mathbf{M}, \sigma^2 \mathbf{I}_p)$, avec $\mathbf{M} \in \text{Sym}(p)$ et $\sigma^2 > 0$, si $\mathbf{D}''$ a pour densité :

$$f_{\mathbf{D}''}(\mathbf{D}''; \mathbf{M}, \sigma^2 \mathbf{I}_p) = \frac{1}{(2\pi)^{q/2} \sigma^q} \exp \left(-\frac{1}{2\sigma^2} \text{tr} [(\mathbf{D}'' - \mathbf{M})^2] \right) \quad (6.10)$$

Puisque la somme des carrés dans (6.5) suit une loi du khi-deux à q degrés de liberté, alors :

$$1/\sigma^2 \text{tr} [(\mathbf{D}'' - \mathbf{M})^2] \sim \chi^2(q) \quad (6.11)$$

Finalement, $\mathbf{D}''$ suit une loi de la forme :

$$\mathbf{D}'' \sim \begin{pmatrix} \mathcal{N}(\mu_{11}, \sigma^2) & \mathcal{N}(\mu_{12}, \sigma^2/2) & \mathcal{N}(\mu_{13}, \sigma^2/2) \\ * & \mathcal{N}(\mu_{22}, \sigma^2) & \mathcal{N}(\mu_{23}, \sigma^2/2) \\ * & * & \mathcal{N}(\mu_{33}, \sigma^2) \end{pmatrix} \quad (6.12)$$

où les termes de $\mathbf{M}$ sont notés μ_{ij} . Signalons une propriété intéressante : soit $\mathbf{D} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ la décomposition spectrale de $\mathbf{D}$ avec $\mathbf{U} \in O(p)$ (l'espace des matrices orthogonales de dimension p) et $\mathbf{\Lambda} \in \text{diag}(p)$ (l'espace des matrices diagonale de dimension p), alors $\text{tr}(\mathbf{D}^2) = \text{tr}(\mathbf{\Lambda}^2)$ (ce qui montre l'invariance de la densité de probabilité (6.4) par des transformations orthogonales).

6.1.3 Distribution multinormale matricielle pour les matrices symétriques définies positives

L'espace des tenseurs de diffusion est contraint. En effet, les tenseurs sont des matrices Symétriques Définies Positives (SDP). La contrainte de positivité conditionne les valeurs propres du tenseur qui doivent être toutes strictement positives. Le déterminant d'un tenseur est donc strictement positif :

$$\mathbf{D} \in \text{Sym}^+(p) \Rightarrow \det(\mathbf{D}) > 0 \quad (6.13)$$

Ainsi, l'espace des matrices SDP est inclus dans l'espace vectoriel des matrices symétriques ($\text{Sym}^+(p) \subset \text{Sym}(p)$). Lorsque $p = 2$, l'espace des tenseurs symétriques définis positifs peut

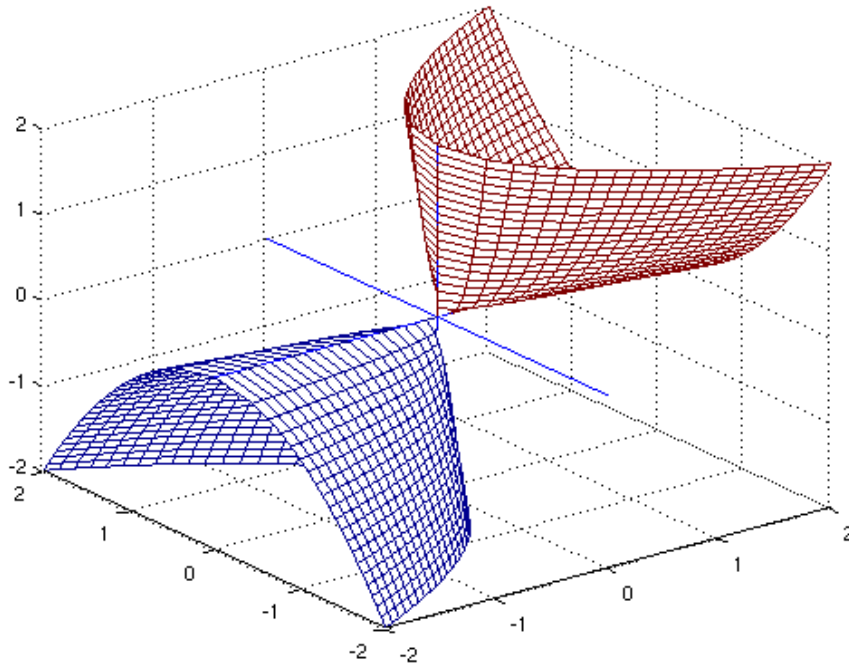


FIG. 6.1 – Frontière des tenseurs 2×2 symétriques définis positifs.

être représenté dans un espace à trois dimensions (correspondant aux trois composantes indépendantes du tenseur) par l'équation : $d_{11}d_{22} - 2d_{12}^2 > 0$. La frontière de cet espace (*i.e.*, $\det(\mathbf{D}) = 0$) est un cône à deux nappes (cf. figure 6.1). Ce cône délimite l'espace des matrices symétriques positives $Sym^+(2)$. À l'intérieur du cône, les matrices ont un déterminant positif : le cône rouge contient les matrices ayant des valeurs propres toutes positives, alors que le cône bleu contient les matrices ayant des valeurs propres toutes négatives. L'espace des tenseurs est donc un espace convexe. Ainsi, toute opération convexe de tenseurs produit un tenseur dans l'espace des matrices définies positives. Parmi ces opérations se trouvent la moyenne pondérée et l'interpolation linéaire. Néanmoins, toute opération non-convexe de tenseurs n'est pas assurée de produire un tenseur défini positif. Par exemple, la soustraction de deux tenseurs SDP ne conduit pas forcément à un tenseur SDP. De même, la multiplication par un scalaire négatif conduit en dehors de l'espace des tenseurs : si $\mathbf{D} \in Sym^+(p)$, $\forall \alpha \leq 0, \alpha \mathbf{D} \notin Sym^+(p)$. Ce manque de structure vectorielle est problématique, notamment lors de la comparaison de deux tenseurs où il faudra effectuer des soustractions de tenseurs (opération non convexe). Il existe plusieurs cadres mathématiques qui autorisent d'effectuer des opérations sur les tenseurs SDP tout en assurant le résultat d'être un tenseur SDP. Pennec [38] propose un cadre riemannien alors que Arsigny et Schwartzman [7, 48, 46] utilisent un cadre log-euclidien pour lever les limitations du calcul euclidien sur les matrices symétriques définies positives.

Dans la suite, l'espace log-euclidien est utilisé pour effectuer des calculs sur les tenseurs. Pour assurer la validité du modèle gaussien aux frontières de l'espace des matrices SDP, l'hypothèse de positivité sur les tenseurs de diffusion est relaxée en travaillant dans le domaine

log-euclidien, c'est-à-dire en travaillant sur le logarithme de chaque tenseur. La distance log-euclidienne entre deux tenseurs $\mathbf{D}_1$ et $\mathbf{D}_2$ s'écrit comme la distance euclidienne entre leurs logarithmes : $d(\mathbf{D}_1, \mathbf{D}_2) = \|\log(\mathbf{D}_1) - \log(\mathbf{D}_2)\|^2$. Le logarithme $\mathbf{L}$ d'un tenseur $\mathbf{D} \in \text{Sym}^+(p)$ s'obtient en appliquant la relation suivante : $\mathbf{L} = \log(\mathbf{D}) = \mathbf{U} \log(\mathbf{\Lambda}) \mathbf{U}^T$, avec $\mathbf{\Lambda}$ la matrice des valeurs propres et $\mathbf{U}$ la matrice des vecteurs propres de $\mathbf{D}$. Finalement, on fait l'hypothèse que $\mathbf{L} \in \text{Sym}(p)$ suit une distribution normale pour les matrices symétriques : $\mathbf{L} \sim \mathcal{N}_{pp}(\mathbf{M}, \Sigma)$. Cette hypothèse doit bien sûr être validée pour les données à modéliser (nous verrons par la suite en section 11.1.1 que cette distribution rend bien compte des données de diffusion). Elle implique que $\mathbf{D}$ suit une loi log-normale définie par :

$$f_{\mathbf{D}}(\mathbf{D}; \mathbf{M}, \sigma^2 \mathbf{I}_p) = \frac{J(\mathbf{D})}{(2\pi)^{q/2} \sigma^q} \exp \left(-\frac{1}{2\sigma^2} \text{tr} [(\log(\mathbf{D}) - \mathbf{M})^2] \right) \quad (6.14)$$

où $J(\mathbf{X}) = J(\mathbf{Y} \rightarrow \mathbf{X}) = |\partial \mathbf{Y} / \partial \mathbf{X}|$ est le jacobien de la transformée log suivante : $\mathbf{Y} = \log \mathbf{X}$.

6.2 Tests statistiques multivariés sur des populations de tenseurs

Une description structurée des tests statistiques multivariés, utilisés dans la suite pour le suivi longitudinal, est présentée dans cette section. Ces tests ont été développés récemment par des statisticiens et appliqués à la comparaison de groupes [48, 49, 46]. Ils offrent la possibilité d'étudier une population (test indicé avec un 1) ou de comparer deux populations (test indicé avec un 2). Ces tests quantifient des changements du tenseur de diffusion (test indicé avec TD), des valeurs propres (test indicé avec VP) ou des vecteurs propres (test indicé avec VEC). Les différents tests considérés dans la suite sont synthétisés dans le tableau 6.1.

6.2.1 Le test du rapport de vraisemblance à un échantillon (RV1)

Le rapport de vraisemblance (RV) teste deux hypothèses : H_0 (les log tenseurs ont une moyenne connue *a priori*) et H_1 (la moyenne des log tenseurs est différente de cet *a priori*). Considérons un échantillon $\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n \in \text{Sym}(p)$ indépendant et identiquement distribué (i.i.d.) suivant la loi de probabilité $\mathcal{N}_{pp}(\mathbf{M}, \sigma^2 \mathbf{I}_p)$. L'idée consiste à comparer les probabilités de l'observation sous chacune des deux hypothèses. Pour nos applications, on veut tester si $\mathbf{M}$ a pour matrice de valeurs propres $\mathbf{\Lambda}_0$, en supposant la matrice des vecteurs propres $\mathbf{U}$ connue :

$$H_0 : \mathbf{M} = \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^T \text{ vs. } H_1 : \mathbf{M} \neq \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^T \quad (6.15)$$

pour une matrice $\mathbf{\Lambda}_0 \in \text{diag}(p)$ et $\mathbf{U} \in O(p)$.

Le rapport de vraisemblance d'un échantillon

Soit $f_{\mathbf{L}_i}$ la loi de probabilité multinormale matricielle pour les matrices symétriques (6.10) et $\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n$ un échantillon observé issu de cette loi. On appelle vraisemblance (notée V) la statistique :

	T_{1VP}	T_{2TD}	T_{2VP}	T_{2VEC}
observations	$\mathcal{N}_{pp}(\mathbf{M}, \sigma^2 \mathbf{I})$	$\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I})$ $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I})$	$\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I})$ $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I})$	$\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I})$ $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I})$
hypotheses	$H_0 : (\mathbf{M}) \in \mathcal{M}_{1, \Lambda_0}$ $H_1 : (\mathbf{M}) \notin \mathcal{M}_{1, \Lambda_0}$	$H_0 : \mathbf{M}_1 = \mathbf{M}_2$ $H_1 : \mathbf{M}_1 \neq \mathbf{M}_2$	$H_0 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2, \Lambda}$ $H_1 : (\mathbf{M}_1, \mathbf{M}_2) \notin \mathcal{M}_{2, \Lambda}$	$H_0 : \mathbf{M}_1 = \mathbf{M}_2$ $H_1 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2, \Lambda}$
estimées	$\hat{\mathbf{M}}_0 = \mathbf{U} \hat{\Lambda}_0 \mathbf{U}^T$ $\hat{\mathbf{M}}_1 = \mathbf{U} \hat{\Lambda} \mathbf{U}^T$ $\hat{\sigma}_1^2 = s^2$	$\hat{\mathbf{M}}_{1,0} = \bar{\mathbf{L}}$ $\hat{\mathbf{M}}_{2,0} = \bar{\mathbf{L}}$ $\hat{\mathbf{M}}_{1,1} = \bar{\mathbf{L}}_1$ $\hat{\mathbf{M}}_{2,1} = \bar{\mathbf{L}}_2$ $\hat{\sigma}_1^2 = s_{12}^2$	$\hat{\mathbf{M}}_{1,0} = \mathbf{V}_1 \hat{\Lambda} \mathbf{V}_1^T$ $\hat{\mathbf{M}}_{2,0} = \mathbf{V}_2 \hat{\Lambda} \mathbf{V}_2^T$ $\hat{\mathbf{M}}_{1,1} = \bar{\mathbf{L}}_1$ $\hat{\mathbf{M}}_{2,1} = \bar{\mathbf{L}}_2$ $\hat{\sigma}_1^2 = s_{12}^2$	$\hat{\mathbf{M}}_{1,0} = \bar{\mathbf{L}}$ $\hat{\mathbf{M}}_{2,0} = \bar{\mathbf{L}}$ $\hat{\mathbf{M}}_{1,1} = \mathbf{V}_1 \hat{\mathbf{D}} \mathbf{V}_1^T$ $\hat{\mathbf{M}}_{2,1} = \mathbf{V}_2 \hat{\mathbf{D}} \mathbf{V}_2^T$ $\hat{\sigma}_1^2 = s_{12}^2 + \frac{n_1 n_2}{qn} \text{tr} [(\Lambda_1 - \Lambda_2)^2]$
test	$\frac{n-1}{ps^2} \text{tr} [(\hat{\Lambda} - \Lambda_0)^2]$	$\frac{(n-2)n_1 n_2}{qn^2 s_{12}^2} \text{tr} [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2]$	$\frac{(n-2)n_1 n_2}{pn^2 s_{12}^2} \text{tr} [(\Lambda_1 - \Lambda_2)^2]$	$\frac{2n_1 n_2 (q(n-2)+2p)}{q(q-p)n\hat{\sigma}_1^2} (\text{tr}(\Lambda_1 \Lambda_2) - \text{tr}(\bar{\mathbf{L}}_1 \bar{\mathbf{L}}_2))$
distribution	$F(p, q(n-1))$	$F(q, q(n-2))$	$F(p, q(n-2))$	$F(q-p, q(n-2)+p)$

TAB. 6.1 – Les tests statistiques multivariés sur des populations de tenseurs.

$$V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}, \sigma^2 \mathbf{I}_p) = \prod_{i=1}^n f_{\mathbf{L}_i}(\mathbf{L}_i; \mathbf{M}, \sigma^2 \mathbf{I}_p) \quad (6.16)$$

L'interprétation est la suivante. Par hypothèse, les variables aléatoires $\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n$ sont indépendantes. Donc la densité de probabilité que l'échantillon théorique $(\tilde{\mathbf{L}}_1, \tilde{\mathbf{L}}_2, \dots, \tilde{\mathbf{L}}_n)$ ait pour réalisation l'échantillon observé $(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n)$ est égal au produit des densité de probabilités que $\mathbf{L}_i$ prenne la valeur $\mathbf{L}_i$, soit :

$$P(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n) = V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}, \sigma^2 \mathbf{I}_p) \quad (6.17)$$

Il est possible de calculer de manière analytique la log-vraisemblance. Pour ce faire, on rappelle l'expression de la distribution multinormale $\mathcal{N}_{pp}(\mathbf{M}, \sigma^2 \mathbf{I}_p)$ pour une matrice $\mathbf{L}_i \in \text{Sym}(p)$ (6.10) :

$$f_{\mathbf{L}_i}(\mathbf{L}_i; \mathbf{M}, \sigma^2 \mathbf{I}_p) = \frac{1}{(2\pi)^{q/2} \sigma^q} \exp \left(-\frac{1}{2\sigma^2} \text{tr} [(\mathbf{L}_i - \mathbf{M})^2] \right) \quad (6.18)$$

La log-vraisemblance (notée l) peut alors s'écrire :

$$\begin{aligned} l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}, \sigma^2 \mathbf{I}_p) &= \log(V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}, \sigma^2 \mathbf{I}_p)) \\ &= \log\left(\prod_{i=1}^n f_{\mathbf{L}_i}(\mathbf{L}_i; \mathbf{M}, \sigma^2 \mathbf{I}_p)\right) = \sum_{i=1}^n \log(f_{\mathbf{L}_i}(\mathbf{L}_i; \mathbf{M}, \sigma^2 \mathbf{I}_p)) \\ &= \sum_{i=1}^n \left(\log \left(\frac{1}{(2\pi)^{q/2} \sigma^q} \right) \right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (\text{tr} [(\mathbf{L}_i - \mathbf{M})^2]) \\ &= -\frac{nq}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (\text{tr} [(\mathbf{L}_i - \mathbf{M})^2]) \end{aligned} \quad (6.19)$$

La décomposition en somme quadratique est ensuite utilisée pour réécrire cette expression [48, 46] :

$$\sum_{i=1}^n \text{tr} [(\mathbf{L}_i - \mathbf{M})^2] = \sum_{i=1}^n \text{tr} [(\mathbf{L}_i - \bar{\mathbf{L}})^2] + n \text{tr} [(\bar{\mathbf{L}} - \mathbf{M})^2] = nqs^2 + n \text{tr} [(\bar{\mathbf{L}} - \mathbf{M})^2] \quad (6.20)$$

où $\bar{\mathbf{L}} = \frac{1}{n} \sum_{i=1}^n \mathbf{L}_i$ représente la moyenne de l'échantillon et s^2 est obtenu par simple identification :

$$s^2 = \frac{1}{qn} \sum_{i=1}^n \text{tr} [(\mathbf{L}_i - \bar{\mathbf{L}})^2] \quad (6.21)$$

En injectant (6.20) dans (6.19), on obtient :

$$l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}, \sigma^2 \mathbf{I}_p) = -\frac{nq}{2} \log(2\pi\sigma^2) - \frac{nq}{2\sigma^2} \left(s^2 + \frac{1}{q} \text{tr} [(\bar{\mathbf{L}} - \mathbf{M})^2] \right) \quad (6.22)$$

L'estimateur de la moyenne $\mathbf{M}$ de l'échantillon au sens du maximum de vraisemblance, pour toute valeur $\sigma > 0$ fixée, est donné par :

$$\hat{\mathbf{M}} = \underset{\mathbf{M}}{\operatorname{argmax}} \left(l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}, \sigma^2 \mathbf{I}_p) \right) = \underset{\mathbf{M}}{\operatorname{argmin}} \left(\operatorname{tr} [(\bar{\mathbf{L}} - \mathbf{M})^2] \right) \quad (6.23)$$

Si aucune contrainte n'est imposée à $\mathbf{M}$, alors $\hat{\mathbf{M}} = \bar{\mathbf{L}}$. Dans le cas contraire (par exemple en fixant les vecteurs propres ou les valeurs propres de $\mathbf{M}$), l'estimée $\hat{\mathbf{M}} \neq \bar{\mathbf{L}}$. En dérivant la log-vraisemblance par rapport à σ^2 et en remplaçant $\mathbf{M}$ par son estimée $\hat{\mathbf{M}}$, on obtient :

$$\frac{\partial l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \hat{\mathbf{M}}, \sigma^2 \mathbf{I}_p)}{\partial \sigma^2} = -\frac{nq}{2\sigma^2} + \frac{nq}{2\sigma^4} \left(s^2 + \frac{1}{q} \operatorname{tr} [(\bar{\mathbf{L}} - \hat{\mathbf{M}})^2] \right) \quad (6.24)$$

L'estimation au sens du maximum de vraisemblance de σ^2 se calcule en annulant cette dérivée :

$$\hat{\sigma}^2 = s^2 + \frac{1}{q} \operatorname{tr} [(\bar{\mathbf{L}} - \hat{\mathbf{M}})^2] \quad (6.25)$$

Finalement, la vraisemblance de l'échantillon sous H_0 est donnée par : $V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}_0, \sigma_0^2 \mathbf{I})$ et la vraisemblance de l'échantillon sous H_1 par : $V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}_1, \sigma_1^2 \mathbf{I})$. L'expression du RV s'écrit alors :

$$RV = \frac{V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}_1, \sigma_1^2 \mathbf{I}_p)}{V(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \mathbf{M}_0, \sigma_0^2 \mathbf{I}_p)} \quad (6.26)$$

Si $RV > 1$ alors H_1 est plus vraisemblable que H_0 . En général, les moyennes $\mathbf{M}_0$ et $\mathbf{M}_1$ et les variances σ_0^2 et σ_1^2 ne sont pas connues et il faut alors considérer leurs estimées (par exemple au sens du maximum de vraisemblance (6.23) et (6.25)). On obtient ainsi les estimées de la variance sous H_0 et H_1 au sens du maximum de vraisemblance :

$$\begin{aligned} \hat{\sigma}_0^2 &= s^2 + \frac{1}{q} \operatorname{tr} [(\bar{\mathbf{L}} - \hat{\mathbf{M}}_0)^2] \\ \hat{\sigma}_1^2 &= s^2 + \frac{1}{q} \operatorname{tr} [(\bar{\mathbf{L}} - \hat{\mathbf{M}}_1)^2] \end{aligned} \quad (6.27)$$

La variance est différente sous H_0 et H_1 car les contraintes sous ces hypothèses ne sont pas nécessairement les mêmes. Les estimées $\hat{\mathbf{M}}_0$ et $\hat{\mathbf{M}}_1$ peuvent donc aussi être différentes. En remplaçant ces estimées dans (6.22), les log-vraisemblances maximisées associées ont pour expression :

$$\begin{aligned} l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \hat{\mathbf{M}}_0, \hat{\sigma}_0^2 \mathbf{I}_p) &= -\frac{nq}{2} - \frac{nq}{2} \log(2\pi \hat{\sigma}_0^2) \\ l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \hat{\mathbf{M}}_1, \hat{\sigma}_1^2 \mathbf{I}_p) &= -\frac{nq}{2} - \frac{nq}{2} \log(2\pi \hat{\sigma}_1^2) \end{aligned} \quad (6.28)$$

Finalement, en considérant le logarithme (L) du rapport de vraisemblance (RV) dans (6.26), on obtient :

$$\begin{aligned} LRV &= l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \hat{\mathbf{M}}_1, \hat{\sigma}_1^2 \mathbf{I}_p) - l(\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n; \hat{\mathbf{M}}_0, \hat{\sigma}_0^2 \mathbf{I}_p) \\ &= \frac{nq}{2} \log \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} \right) = \frac{nq}{2} \log \left(1 + \frac{\hat{\sigma}_0^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2} \right) \end{aligned} \quad (6.29)$$

En remarquant que cette expression a le même sens de variation que $nq \frac{\hat{\sigma}_0^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2}$, on obtient le test suivant :

$$T_{LRV} = nq \frac{\hat{\sigma}_0^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2} = \frac{n}{\hat{\sigma}_1^2} \left(\text{tr} \left[(\bar{\mathbf{L}} - \hat{\mathbf{M}}_0)^2 \right] - \text{tr} \left[(\bar{\mathbf{L}} - \hat{\mathbf{M}}_1)^2 \right] \right) \quad (6.30)$$

On remarque que l'expression du test ne dépend plus de σ_0^2 , et nécessite d'estimer uniquement $\mathbf{M}_0$, $\mathbf{M}_1$ et σ_1^2 .

Le test sur les valeurs propres

Soit un échantillon i.i.d. $\mathbf{L}_1, \mathbf{L}_2, \dots, \mathbf{L}_n$ suivant une loi $\mathcal{N}_{pp}(\mathbf{M}, \sigma^2 \mathbf{I}_p)$, avec $\mathbf{M}$ supposée avoir une matrice de vecteur propre $\mathbf{U}$ connue. On veut tester si $\mathbf{M}$ a pour matrice des valeurs propres $\mathbf{\Lambda}_0$ (hypothèse H_0), ou bien si elle a pour matrice des valeurs propres $\mathbf{\Lambda} \neq \mathbf{\Lambda}_0$ (hypothèse H_1). Tout d'abord, il faut déterminer les estimées au sens du maximum de vraisemblance de $\mathbf{M}$ sous les hypothèses H_0 et H_1 et de σ sous H_1 . Par hypothèse, $\mathbf{M}$ est supposée connue sous H_0 et est de la forme $\hat{\mathbf{M}}_0 = \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^T$. Sous H_1 , on cherche à estimer $\mathbf{M}$ connaissant sa matrice des vecteurs propres $\mathbf{U}$ mais pas sa matrice des valeurs propres $\mathbf{\Lambda}$. D'après (6.23), l'estimation de $\mathbf{\Lambda}$ sous H_1 est donnée par l'équation suivante :

$$\hat{\mathbf{\Lambda}} = \underset{\mathbf{\Lambda}}{\text{argmin}} \left(\text{tr} \left[(\bar{\mathbf{L}} - \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T)^2 \right] \right) = \underset{\mathbf{\Lambda}}{\text{argmin}} \left(\text{tr} \left[(\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U} - \mathbf{\Lambda})^2 \right] \right) \quad (6.31)$$

Le minimum est atteint lorsque la matrice diagonale $\mathbf{\Lambda}$ est la plus proche de $\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U}$ en considérant la norme de Frobenius. On a donc $\hat{\mathbf{\Lambda}} = \text{diag}(\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U})$, la projection orthogonale de $\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U}$ sur l'espace des matrices diagonales. Ensuite, en remplaçant $\hat{\mathbf{M}}_1 = \mathbf{U} \hat{\mathbf{\Lambda}} \mathbf{U}^T$ dans (6.25), on obtient l'estimation de la variance sous H_1 :

$$\hat{\sigma}_1^2 = s^2 + \frac{1}{q} \text{tr} \left[(\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U} - \hat{\mathbf{\Lambda}})^2 \right] \quad (6.32)$$

De manière synthétique, on obtient :

$$\begin{aligned} H_0 &: \hat{\mathbf{M}}_0 = \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^T \\ H_1 &: \hat{\mathbf{M}}_1 = \mathbf{U} \hat{\mathbf{\Lambda}} \mathbf{U}^T, \text{ avec } \hat{\mathbf{\Lambda}} = \text{diag}(\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U}) \\ &\hat{\sigma}_1^2 = s^2 + \frac{1}{q} \text{tr} \left[(\mathbf{U}^T \bar{\mathbf{L}} \mathbf{U} - \hat{\mathbf{\Lambda}})^2 \right] \end{aligned} \quad (6.33)$$

En remplaçant les estimées de $\mathbf{M}_0$ et $\mathbf{M}_1$ (6.33) dans (6.30), on arrive au test suivant :

$$T_{LRV} = \frac{n}{\hat{\sigma}_1^2} \left(\text{tr} \left[(\bar{\mathbf{L}} - \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^T)^2 \right] - \text{tr} \left[(\bar{\mathbf{L}} - \mathbf{U} \hat{\mathbf{\Lambda}} \mathbf{U}^T)^2 \right] \right) \quad (6.34)$$

Si, lors d'une étude, seules les valeurs propres des matrices $\mathbf{L}_i$ sont conservées (on se place dans la base canonique), on choisit $\mathbf{U} = \mathbf{I}_p$. Avec cette hypothèse, on remarque que le deuxième terme dans (6.34) s'annule. On arrive à l'expression du test sur les valeurs propres :

$$T_{LRV} = \frac{n}{\hat{\sigma}_1^2} \text{tr} \left[(\hat{\mathbf{\Lambda}} - \mathbf{\Lambda}_0)^2 \right] \quad (6.35)$$

avec :

$$\begin{aligned} \hat{\sigma}_1^2 &= s^2 \\ \hat{\mathbf{\Lambda}} &= \bar{\mathbf{L}} \end{aligned} \quad (6.36)$$

Intéressons nous à la distribution asymptotique de T_{LRV} donnée dans (6.35) [46]. La variance estimée peut être vue comme une variable aléatoire : $\frac{qn}{\sigma^2} s^2 \sim \mathcal{X}^2(q(n-1))$ et $\frac{n}{\sigma^2} \text{tr} \left[(\hat{\mathbf{\Lambda}} - \mathbf{\Lambda}_0)^2 \right] \sim \mathcal{X}^2(p)$. En combinant ces deux termes, on obtient le test T_{1VP} qui suit asymptotiquement une loi de Fisher notée F :

$$T_{1VP} = \frac{\frac{1}{p} \frac{n}{\sigma^2} \text{tr} \left[(\hat{\mathbf{\Lambda}} - \mathbf{\Lambda}_0)^2 \right]}{\frac{1}{q(n-1)} \frac{qn}{\sigma^2} s^2} = \frac{n-1}{ps^2} \text{tr} \left[(\hat{\mathbf{\Lambda}} - \mathbf{\Lambda}_0)^2 \right] \sim F(p, q(n-1)) \quad (6.37)$$

6.2.2 Le test du rapport de vraisemblance à deux échantillons (RV2)

Soient $\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}$ et $\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}$ deux échantillons i.i.d. suivant respectivement les lois $\mathcal{N}_{pp}(\mathbf{M}_1, \sigma_1^2 \mathbf{I}_p)$ et $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma_2^2 \mathbf{I}_p)$. Trois tests sont envisagés :

- Le premier teste si $\mathbf{M}_1$ et $\mathbf{M}_2$ ont les mêmes valeurs propres, sans contrainte sur les vecteurs propres qui sont traités comme des paramètres de nuisance :

$$H_0 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,\Lambda} \text{ vs. } H_1 : (\mathbf{M}_1, \mathbf{M}_2) \notin \mathcal{M}_{2,\Lambda} \quad (6.38)$$

où :

$$\mathcal{M}_{2,\Lambda} = \{(\mathbf{M}_1, \mathbf{M}_2) : \mathbf{M}_1 = \mathbf{U}_1 \mathbf{\Lambda} \mathbf{U}_1^T, \mathbf{M}_2 = \mathbf{U}_2 \mathbf{\Lambda} \mathbf{U}_2^T\} \quad (6.39)$$

pour une matrice $\mathbf{\Lambda} \in \text{diag}(p)$ et $\mathbf{U}_1, \mathbf{U}_2 \in O(p)$.

- Un autre test sur les vecteurs propres permet de savoir si $\mathbf{M}_1$ et $\mathbf{M}_2$ ont les mêmes orientations. Cette fois, les valeurs propres sont supposées égales dans les deux échantillons :

$$H_0 : \mathbf{M}_1 = \mathbf{M}_2 \text{ vs. } H_1 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,\Lambda} \quad (6.40)$$

- Un troisième test sur la structure matricielle complète du tenseur de diffusion est considéré :

$$H_0 : \mathbf{M}_1 = \mathbf{M}_2 \text{ vs. } H_1 : \mathbf{M}_1 \neq \mathbf{M}_2 \quad (6.41)$$

Dans ce cas particulier, les $\mathbf{L}_{1i}$ et les $\mathbf{L}_{2i}$ peuvent être vus comme des échantillons multivariés à q dimensions :

$$\mathbf{L} \sim \mathcal{N}_{pp}(\mathbf{M}, \sigma^2 \mathbf{I}_p) \Leftrightarrow \text{vec}(\mathbf{L}) \sim \mathcal{N}_p(\text{vec}(\mathbf{M}), \sigma^2 \mathbf{I}_q) \quad (6.42)$$

Le rapport de vraisemblance de deux échantillons

Soient deux échantillons i.i.d. $\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I}_p)$ et $\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I}_p)$. La log-vraisemblance l_{12} s'écrit comme la somme de la log-vraisemblance de l'échantillon 1 et de l'échantillon 2 (6.22). Soit $n = n_1 + n_2$, on obtient :

$$\begin{aligned} l_{12}(\mathbf{M}_1, \mathbf{M}_2, \sigma^2 \mathbf{I}_p) &= l(\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}; \mathbf{M}_1, \sigma^2 \mathbf{I}_p) + l(\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}; \mathbf{M}_2, \sigma^2 \mathbf{I}_p) \\ &= -\frac{nq}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (nqs_{12}^2 + n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \mathbf{M}_1)^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \mathbf{M}_2)^2]) \end{aligned} \quad (6.43)$$

où :

$$\begin{aligned} \bar{\mathbf{L}}_1 &= \sum_{i=1}^{n_1} \mathbf{L}_{1,i} \quad \bar{\mathbf{L}}_2 = \sum_{i=1}^{n_2} \mathbf{L}_{2,i} \\ s_{12}^2 &= \frac{n_1 s_1^2 + n_2 s_2^2}{n} = \frac{1}{qn} \left(\sum_{i=1}^{n_1} \text{tr} [(\mathbf{L}_{1,i} - \bar{\mathbf{L}}_1)^2] + \sum_{i=1}^{n_2} \text{tr} [(\mathbf{L}_{2,i} - \bar{\mathbf{L}}_2)^2] \right) \end{aligned} \quad (6.44)$$

Le problème d'estimation des moyennes $(\mathbf{M}_1, \mathbf{M}_2)$ au sens du maximum de vraisemblance s'écrit :

$$(\hat{\mathbf{M}}_1, \hat{\mathbf{M}}_2) = \underset{(\mathbf{M}_1, \mathbf{M}_2)}{\text{argmin}} (n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \mathbf{M}_1)^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \mathbf{M}_2)^2]) \quad (6.45)$$

Sans contrainte sur $\mathbf{M}_1$ et $\mathbf{M}_2$, les estimées associées sont $\bar{\mathbf{L}}_1$ et $\bar{\mathbf{L}}_2$. Par contre, avec des contraintes sur $\mathbf{M}_1$ ou $\mathbf{M}_2$ (sur les valeurs propres et/ou sur les vecteurs propres), les expressions des estimées peuvent changer. En dérivant la log-vraisemblance l_{12} par rapport à σ^2 et en remplaçant M_1 et M_2 par leurs estimées, on obtient :

$$\frac{\partial l_{12}(\hat{\mathbf{M}}_1, \hat{\mathbf{M}}_2, \sigma^2 \mathbf{I}_p)}{\partial \sigma^2} = -\frac{nq}{2\sigma^2} + \frac{1}{2\sigma^4} \left(nqs_{12}^2 + n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_1)^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_2)^2] \right) \quad (6.46)$$

Une estimation au sens du maximum de vraisemblance de σ^2 est obtenue en annulant cette dérivée :

$$\hat{\sigma}^2 = s_{12}^2 + \frac{1}{qn} \left(n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_1)^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_2)^2] \right) \quad (6.47)$$

Finalement, la vraisemblance sous H_0 est donnée par $V(\mathbf{L}_{i,1}, \mathbf{L}_{i,2}, \dots, \mathbf{L}_{i,j}; \mathbf{M}_{i,0}, \sigma_0^2 \mathbf{I}_p)$, $(i, j) \in \{(1, n_1), (2, n_2)\}$, et la vraisemblance sous H_1 est donnée par $V(\mathbf{L}_{i,1}, \mathbf{L}_{i,2}, \dots, \mathbf{L}_{i,j}; \mathbf{M}_{i,1}, \sigma_1^2 \mathbf{I}_p)$, $(i, j) \in \{(1, n_1), (2, n_2)\}$. L'expression du RV s'écrit alors :

$$RV = \frac{V(\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}; \mathbf{M}_{1,1}, \sigma_1^2 \mathbf{I}_p) V(\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}; \mathbf{M}_{2,1}, \sigma_1^2 \mathbf{I}_p)}{V(\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}; \mathbf{M}_{1,0}, \sigma_0^2 \mathbf{I}_p) V(\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}; \mathbf{M}_{2,0}, \sigma_0^2 \mathbf{I}_p)} \quad (6.48)$$

En général, les moyennes $\mathbf{M}_{1,0}$, $\mathbf{M}_{2,0}$, $\mathbf{M}_{1,1}$ et $\mathbf{M}_{2,1}$ et les variances σ_0^2 et σ_1^2 ne sont pas connues. Il faut considérer leurs estimées (par exemple au sens du maximum de vraisemblance (6.45) et (6.47)). On obtient ainsi les estimées de la variance sous H_0 et H_1 au sens du maximum de vraisemblance :

$$\begin{aligned} \hat{\sigma}_0^2 &= s_{12}^2 + \frac{1}{qn} \left(n_1 tr \left[(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_{1,0})^2 \right] + n_2 tr \left[(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_{2,0})^2 \right] \right) \\ \hat{\sigma}_1^2 &= s_{12}^2 + \frac{1}{qn} \left(n_1 tr \left[(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_{1,1})^2 \right] + n_2 tr \left[(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_{2,1})^2 \right] \right) \end{aligned} \quad (6.49)$$

En remplaçant ces estimées dans (6.43), les log-vraisemblances maximisées associées ont pour expression :

$$\begin{aligned} l_{12}^{H_0} &= l(\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}; \hat{\mathbf{M}}_{1,0}, \hat{\sigma}_0^2 \mathbf{I}_p) + l(\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}; \hat{\mathbf{M}}_{2,0}, \hat{\sigma}_0^2 \mathbf{I}_p) \\ &= -\frac{nq}{2} - \frac{nq}{2} \log(2\pi \hat{\sigma}_0^2) \\ l_{12}^{H_1} &= l(\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}; \hat{\mathbf{M}}_{1,1}, \hat{\sigma}_1^2 \mathbf{I}_p) + l(\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}; \hat{\mathbf{M}}_{2,1}, \hat{\sigma}_1^2 \mathbf{I}_p) \\ &= -\frac{nq}{2} - \frac{nq}{2} \log(2\pi \hat{\sigma}_1^2) \end{aligned} \quad (6.50)$$

Finalement, en considérant le logarithme (L) du rapport de vraisemblance (RV), on obtient :

$$LRV = l_{12}^{H_1} - l_{12}^{H_0} = \frac{nq}{2} \log \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} \right) = \frac{nq}{2} \log \left(1 + \frac{\hat{\sigma}_0^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2} \right) \quad (6.51)$$

En remarquant que cette expression a le même sens de variation que $nq \frac{\hat{\sigma}_0^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2}$, on obtient le test suivant :

$$\begin{aligned} T_{LRV} &= nq \frac{\hat{\sigma}_0^2 - \hat{\sigma}_1^2}{\hat{\sigma}_1^2} = \frac{1}{\hat{\sigma}_1^2} \left(n_1 tr \left[(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_{1,0})^2 \right] + n_2 tr \left[(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_{2,0})^2 \right] \right. \\ &\quad \left. - n_1 tr \left[(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_{1,1})^2 \right] - n_2 tr \left[(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_{2,1})^2 \right] \right) \end{aligned} \quad (6.52)$$

On remarque que l'expression du test ne dépend plus de σ_0^2 , et nécessite d'estimer uniquement $\mathbf{M}_1$, $\mathbf{M}_2$ sous H_0 et H_1 et σ_1^2 .

Le test sur le tenseur de diffusion

Soient deux échantillons i.i.d. $\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I}_p)$ et $\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I}_p)$. Intéressons nous au test d'hypothèses défini par l'équation

(6.41) : $H_0 : \mathbf{M}_1 = \mathbf{M}_2$ vs. $H_1 : \mathbf{M}_1 \neq \mathbf{M}_2$. On teste si les moyennes des deux échantillons sont identiques. Tout d'abord, il faut déterminer les estimations au sens du maximum de vraisemblance de $\mathbf{M}_1$ et $\mathbf{M}_2$ sous les hypothèses H_0 et H_1 et de σ sous H_1 . Sous H_0 , les données peuvent être regroupées dans un échantillon de taille n et de moyenne $\mathbf{M}_0 = \mathbf{M}_{1,0} = \mathbf{M}_{2,0}$. L'estimation de $\mathbf{M}_0$ se fait sans contrainte et minimise $tr [(\bar{\mathbf{L}} - \mathbf{M}_0)^2]$ (6.23), soit $\hat{\mathbf{M}}_0 = \bar{\mathbf{L}} = \frac{1}{n}(n_1\bar{\mathbf{L}}_1 + n_2\bar{\mathbf{L}}_2)$. Sous H_1 , les estimations de $\mathbf{M}_{1,1}$ et $\mathbf{M}_{2,1}$ se font sans contrainte et minimisent $n_1 tr [(\bar{\mathbf{L}}_1 - \mathbf{M}_1)^2] + n_2 tr [(\bar{\mathbf{L}}_2 - \mathbf{M}_2)^2]$ (6.45), soient $\hat{\mathbf{M}}_{1,1} = \bar{\mathbf{L}}_1$ et $\hat{\mathbf{M}}_{2,1} = \bar{\mathbf{L}}_2$. En remplaçant $\hat{\mathbf{M}}_{1,1}$ et $\hat{\mathbf{M}}_{2,1}$ dans (6.49), on obtient une estimée de la variance sous H_1 $\hat{\sigma}_1^2 = s_{12}^2$. De manière synthétique, on a :

$$\begin{aligned} H_0 &: \hat{\mathbf{M}}_{1,0} = \hat{\mathbf{M}}_{2,0} = \bar{\mathbf{L}} \\ H_1 &: \hat{\mathbf{M}}_{1,1} = \bar{\mathbf{L}}_1, \hat{\mathbf{M}}_{2,1} = \bar{\mathbf{L}}_2 \\ &\hat{\sigma}_1^2 = s_{12}^2 \end{aligned} \quad (6.53)$$

En injectant ces estimées dans (6.52), on obtient l'expression du test suivante :

$$T_{LRV} = \frac{1}{s_{12}^2} (n_1 tr [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}})^2] + n_2 tr [(\bar{\mathbf{L}}_2 - \bar{\mathbf{L}})^2]) \quad (6.54)$$

Finalement, en remplaçant $\bar{\mathbf{L}}$ par son expression, on arrive au test sur le tenseur de diffusion :

$$\begin{aligned} T_{LRV} &= \frac{1}{s_{12}^2} (n_1 tr [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}})^2] + n_2 tr [(\bar{\mathbf{L}}_2 - \bar{\mathbf{L}})^2]) \\ &= \frac{1}{s_{12}^2} \left(n_1 tr \left[\left(\frac{n_2(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)}{n} \right)^2 \right] + n_2 tr \left[\left(\frac{n_1(\bar{\mathbf{L}}_2 - \bar{\mathbf{L}}_1)}{n} \right)^2 \right] \right) \\ &= \frac{n_1 n_2}{n s_{12}^2} tr [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2] \end{aligned} \quad (6.55)$$

Intéressons nous à la distribution asymptotique de T_{LRV} [46]. La variance estimée peut être vue comme une variable aléatoire : $\frac{qn}{\sigma^2} s_{12}^2 \sim \mathcal{X}^2(q(n-2))$ et $\frac{n_1 n_2}{n \sigma^2} tr [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2] \sim \mathcal{X}^2(q)$. En combinant ces deux termes, on obtient le test T_{2TD} qui suit asymptotiquement une loi de Fisher notée F :

$$T_{2TD} = \frac{\frac{1}{q} \frac{n_1 n_2}{n \sigma^2} tr [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2]}{\frac{1}{q(n-2)} \frac{qn}{\sigma^2} s_{12}^2} = \frac{(n-2)n_1 n_2}{qn^2 s_{12}^2} tr [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2] \sim F(q, q(n-2)) \quad (6.56)$$

Le test sur les valeurs propres

Soient deux échantillons i.i.d. $\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I}_p)$ et $\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I}_p)$. Intéressons nous au test d'hypothèses défini par l'équation (6.38) : $H_0 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,\Lambda}$ vs. $H_1 : (\mathbf{M}_1, \mathbf{M}_2) \notin \mathcal{M}_{2,\Lambda}$. On teste si $\mathbf{M}_1$ et $\mathbf{M}_2$ ont les mêmes valeurs propres sans contrainte sur les vecteurs propres. Pour ce faire, on considère

les décompositions spectrales suivantes : $\mathbf{M}_1 = \mathbf{U}_1 \mathbf{D}_1 \mathbf{U}_1^T$, $\mathbf{M}_2 = \mathbf{U}_2 \mathbf{D}_2 \mathbf{U}_2^T$, $\bar{\mathbf{L}}_1 = \mathbf{V}_1 \mathbf{\Lambda}_1 \mathbf{V}_1^T$ et $\bar{\mathbf{L}}_2 = \mathbf{V}_2 \mathbf{\Lambda}_2 \mathbf{V}_2^T$. Tout d'abord, il faut déterminer les estimations au sens du maximum de vraisemblance de $\mathbf{M}_1$ et $\mathbf{M}_2$ sous les hypothèses H_0 et H_1 et de σ sous H_1 . Sous H_0 , les moyennes $\mathbf{M}_{1,0}$ et $\mathbf{M}_{2,0}$ ont la même matrice des valeurs propres $\mathbf{\Lambda}$ et des matrices des vecteurs propres différentes $\mathbf{U}_1$ et $\mathbf{U}_2$. D'après (6.45), l'estimation au sens du maximum de vraisemblance des paramètres $\mathbf{\Lambda}$, $\mathbf{U}_1$ et $\mathbf{U}_2$ doit minimiser la quantité $g(\mathbf{U}_1, \mathbf{U}_2, \mathbf{\Lambda})$ suivante :

$$\begin{aligned} g(\mathbf{U}_1, \mathbf{U}_2, \mathbf{\Lambda}) &= n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \mathbf{U}_1 \mathbf{\Lambda} \mathbf{U}_1^T)^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \mathbf{U}_2 \mathbf{\Lambda} \mathbf{U}_2^T)^2] \\ &= n_1 \text{tr} [(\mathbf{U}_1^T \bar{\mathbf{L}}_1 \mathbf{U}_1 - \mathbf{\Lambda})^2] + n_2 \text{tr} [(\mathbf{U}_2^T \bar{\mathbf{L}}_2 \mathbf{U}_2 - \mathbf{\Lambda})^2] \\ &= \text{tr} (n_1 \bar{\mathbf{L}}_1^2 + n_2 \bar{\mathbf{L}}_2^2) - 2 \text{tr} (n_1 \mathbf{U}_1^T \bar{\mathbf{L}}_1 \mathbf{U}_1 \mathbf{\Lambda} + n_2 \mathbf{U}_2^T \bar{\mathbf{L}}_2 \mathbf{U}_2 \mathbf{\Lambda}) + n \text{tr} (\mathbf{\Lambda}^2) \end{aligned} \quad (6.57)$$

Cette dernière relation est obtenue en développant le terme dans la trace. En considérant la matrice $\mathbf{\Lambda}$ connue, les points critiques de $g(\mathbf{U}_1, \mathbf{U}_2, \mathbf{\Lambda})$ qui respectent les contraintes $\mathbf{U}_1^T \mathbf{U}_1 = \mathbf{I}_p$ et $\mathbf{U}_2^T \mathbf{U}_2 = \mathbf{I}_p$ sont $\hat{\mathbf{U}}_1 = \mathbf{V}_1$ et $\hat{\mathbf{U}}_2 = \mathbf{V}_2$ [46, 49]. Maintenant, si $\mathbf{\Lambda}$ est une inconnue, une estimée de ce paramètre est obtenue en dérivant (6.57) par rapport à $\mathbf{\Lambda}$ et en remplaçant $\mathbf{U}_1$ et $\mathbf{U}_2$ par leurs estimées :

$$\begin{aligned} \frac{\partial g(\hat{\mathbf{U}}_1, \hat{\mathbf{U}}_2, \mathbf{\Lambda})}{\partial \mathbf{\Lambda}} &= -2 \left(n_1 \hat{\mathbf{U}}_1^T \bar{\mathbf{L}}_1 \hat{\mathbf{U}}_1 + n_2 \hat{\mathbf{U}}_2^T \bar{\mathbf{L}}_2 \hat{\mathbf{U}}_2 \right) + 2n \mathbf{\Lambda} \\ &= -2 (n_1 \mathbf{\Lambda}_1 + n_2 \mathbf{\Lambda}_2) + 2n \mathbf{\Lambda} \end{aligned} \quad (6.58)$$

L'estimation au sens du maximum de vraisemblance de $\mathbf{\Lambda}$ se calcule en annulant cette dérivée, soit :

$$\hat{\mathbf{\Lambda}} = \frac{n_1 \mathbf{\Lambda}_1 + n_2 \mathbf{\Lambda}_2}{n} \quad (6.59)$$

Sous H_1 , les estimations de $\mathbf{M}_{1,1}$ et $\mathbf{M}_{2,1}$ se font sans contrainte et minimisent $n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \mathbf{M}_1)^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \mathbf{M}_2)^2]$ (6.45), soient $\hat{\mathbf{M}}_{1,1} = \bar{\mathbf{L}}_1$ et $\hat{\mathbf{M}}_{2,1} = \bar{\mathbf{L}}_2$. En remplaçant $\hat{\mathbf{M}}_{1,1}$ et $\hat{\mathbf{M}}_{2,1}$ dans (6.49), on obtient une estimée de la variance sous H_1 $\hat{\sigma}_1^2 = s_{12}^2$. De manière synthétique, on a :

$$\begin{aligned} H_0 &: \hat{\mathbf{M}}_{1,0} = \mathbf{V}_1 \hat{\mathbf{\Lambda}} \mathbf{V}_1^T, \hat{\mathbf{M}}_{2,0} = \mathbf{V}_2 \hat{\mathbf{\Lambda}} \mathbf{V}_2^T \\ H_1 &: \hat{\mathbf{M}}_{1,1} = \bar{\mathbf{L}}_1, \hat{\mathbf{M}}_{2,1} = \bar{\mathbf{L}}_2 \\ &\hat{\sigma}_1^2 = s_{12}^2 \end{aligned} \quad (6.60)$$

Comme précédemment, on peut simplifier le logarithme du rapport de vraisemblance dans (6.52) :

$$T_{LRV} = \frac{1}{s_{12}^2} \left(n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \hat{\mathbf{M}}_{1,0})^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \hat{\mathbf{M}}_{2,0})^2] \right) \quad (6.61)$$

En remplaçant les estimées de $\mathbf{M}_1$ et $\mathbf{M}_2$ sous H_0 (6.60) et en utilisant la propriété d'invariance de la trace par rotation, on obtient :

$$\begin{aligned} T_{LRV} &= \frac{1}{s_{12}^2} \left(n_1 \text{tr} \left[(\mathbf{V}_1 \mathbf{\Lambda}_1 \mathbf{V}_1^T - \mathbf{V}_1 \hat{\mathbf{\Lambda}} \mathbf{V}_1^T)^2 \right] + n_2 \text{tr} \left[(\mathbf{V}_2 \mathbf{\Lambda}_2 \mathbf{V}_2^T - \mathbf{V}_2 \hat{\mathbf{\Lambda}} \mathbf{V}_2^T)^2 \right] \right) \\ &= \frac{1}{s_{12}^2} \left(n_1 \text{tr} \left[(\mathbf{\Lambda}_1 - \hat{\mathbf{\Lambda}})^2 \right] + n_2 \text{tr} \left[(\mathbf{\Lambda}_2 - \hat{\mathbf{\Lambda}})^2 \right] \right) \end{aligned} \quad (6.62)$$

Finalement, en remplaçant $\hat{\mathbf{\Lambda}} = \frac{1}{n}(n_1 \mathbf{\Lambda}_1 + n_2 \mathbf{\Lambda}_2)$, on arrive au test sur les valeurs propres :

$$T_{LRV} = \frac{n_1 n_2}{n s_{12}^2} \text{tr} \left[(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2 \right] \quad (6.63)$$

Intéressons nous à la distribution asymptotique de T_{LRV} [46]. La variance estimée peut être vue comme une variable aléatoire : $\frac{qn}{\sigma^2} s_{12}^2 \sim \mathcal{X}^2(q(n-2))$ et $\frac{n_1 n_2}{n \sigma^2} \text{tr} \left[(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2 \right] \sim \mathcal{X}^2(p)$. En combinant ces deux termes, on obtient le test T_{2VP} qui suit asymptotiquement une loi de Fisher notée F :

$$T_{2VP} = \frac{\frac{1}{p} \frac{n_1 n_2}{n \sigma^2} \text{tr} \left[(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2 \right]}{\frac{1}{q(n-2)} \frac{qn}{\sigma^2} s_{12}^2} = \frac{(n-2) n_1 n_2}{p n^2 s_{12}^2} \text{tr} \left[(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2 \right] \sim F(p, q(n-2)) \quad (6.64)$$

Le test sur les vecteurs propres

Soient deux échantillons i.i.d. $\mathbf{L}_{1,1}, \mathbf{L}_{1,2}, \dots, \mathbf{L}_{1,n_1}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_1, \sigma^2 \mathbf{I}_p)$ et $\mathbf{L}_{2,1}, \mathbf{L}_{2,2}, \dots, \mathbf{L}_{2,n_2}$ suivant la loi $\mathcal{N}_{pp}(\mathbf{M}_2, \sigma^2 \mathbf{I}_p)$. Intéressons nous au test d'hypothèses défini par l'équation (6.40) : $H_0 : \mathbf{M}_1 = \mathbf{M}_2$ vs. $H_1 : (\mathbf{M}_1, \mathbf{M}_2) \in \mathcal{M}_{2,\Lambda}$. On teste si $\mathbf{M}_1$ et $\mathbf{M}_2$ ont les mêmes orientations. Pour ce faire, on considère les décompositions spectrales suivantes : $\mathbf{M}_1 = \mathbf{U}_1 \mathbf{D} \mathbf{U}_1^T$, $\mathbf{M}_2 = \mathbf{U}_2 \mathbf{D} \mathbf{U}_2^T$, $\bar{\mathbf{L}}_1 = \mathbf{V}_1 \mathbf{\Lambda}_1 \mathbf{V}_1^T$ et $\bar{\mathbf{L}}_2 = \mathbf{V}_2 \mathbf{\Lambda}_2 \mathbf{V}_2^T$. Tout d'abord, il faut déterminer les estimations au sens du maximum de vraisemblance de $\mathbf{M}_1$ et $\mathbf{M}_2$ sous les hypothèses H_0 et H_1 et de σ sous H_1 . Sous H_0 , les données peuvent être regroupées dans un échantillon de taille n et de moyenne $\mathbf{M}_0 = \mathbf{M}_{1,0} = \mathbf{M}_{2,0}$. L'estimation de $\mathbf{M}_0$ se fait sans contrainte et minimise $\text{tr} \left[(\bar{\mathbf{L}} - \mathbf{M}_0)^2 \right]$ (6.23), soit $\hat{\mathbf{M}}_0 = \bar{\mathbf{L}} = \frac{1}{n}(n_1 \bar{\mathbf{L}}_1 + n_2 \bar{\mathbf{L}}_2)$. Sous H_1 , les moyennes $\mathbf{M}_{1,1}$ et $\mathbf{M}_{2,1}$ ont par hypothèse la même matrice des valeurs propres $\mathbf{\Lambda}$ et des matrices des vecteurs propres différentes $\mathbf{U}_1$ et $\mathbf{U}_2$. D'après (6.45), l'estimation au sens du maximum de vraisemblance des paramètres $\mathbf{\Lambda}$, $\mathbf{U}_1$ et $\mathbf{U}_2$ doit minimiser la quantité $g(\mathbf{U}_1, \mathbf{U}_2, \mathbf{\Lambda})$ donnée en (6.57). Les estimées au sens du maximum de vraisemblance sont donc $\hat{\mathbf{U}}_1 = \mathbf{V}_1$, $\hat{\mathbf{U}}_2 = \mathbf{V}_2$ et $\hat{\mathbf{\Lambda}} = \frac{n_1 \mathbf{\Lambda}_1 + n_2 \mathbf{\Lambda}_2}{n}$. L'estimateur de la variance au sens du maximum de vraisemblance sous H_1 s'écrit alors (6.49) :

$$\begin{aligned} \hat{\sigma}_1^2 &= s_{12}^2 + \frac{1}{qn} \left(n_1 \text{tr} \left[(\bar{\mathbf{L}}_1 - \mathbf{V}_1 \hat{\mathbf{\Lambda}} \mathbf{V}_1^T)^2 \right] + n_2 \text{tr} \left[(\bar{\mathbf{L}}_2 - \mathbf{V}_2 \hat{\mathbf{\Lambda}} \mathbf{V}_2^T)^2 \right] \right) \\ &= s_{12}^2 + \frac{1}{qn} \left(n_1 \text{tr} \left[(\mathbf{\Lambda}_1 - \hat{\mathbf{\Lambda}})^2 \right] + n_2 \text{tr} \left[(\mathbf{\Lambda}_2 - \hat{\mathbf{\Lambda}})^2 \right] \right) \\ &= s_{12}^2 + \frac{n_1 n_2}{qn} \text{tr} \left[(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2 \right] \end{aligned} \quad (6.65)$$

La dernière relation est obtenue en remplaçant $\hat{\mathbf{\Lambda}}$ par son expression. De manière synthétique, on a :

$$\begin{aligned} H_0 &: \hat{\mathbf{M}}_{1,0} = \hat{\mathbf{M}}_{2,0} = \bar{\mathbf{L}} \\ H_1 &: \hat{\mathbf{M}}_{1,1} = \mathbf{V}_1 \hat{\mathbf{\Lambda}} \mathbf{V}_1^T, \hat{\mathbf{M}}_{2,1} = \mathbf{V}_2 \hat{\mathbf{\Lambda}} \mathbf{V}_2^T \\ &\hat{\sigma}_1^2 = s_{12}^2 + \frac{n_1 n_2}{qn} \text{tr} [(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2] \end{aligned} \quad (6.66)$$

Pour établir le test, on réécrit l'expression du logarithme du rapport de vraisemblance (6.52) :

$$\begin{aligned} T_{LRV} &= \frac{1}{\hat{\sigma}_1^2} \left(n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}})^2] + n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \bar{\mathbf{L}})^2] \right. \\ &\quad \left. - n_1 \text{tr} [(\bar{\mathbf{L}}_1 - \mathbf{V}_1 \hat{\mathbf{\Lambda}} \mathbf{V}_1^T)^2] - n_2 \text{tr} [(\bar{\mathbf{L}}_2 - \mathbf{V}_2 \hat{\mathbf{\Lambda}} \mathbf{V}_2^T)^2] \right) \end{aligned} \quad (6.67)$$

Pour deux ensembles $X = (\mathbf{X}_1, \mathbf{X}_2)$ et $Y = (\mathbf{Y}_1, \mathbf{Y}_2)$ contenant $n = n_1 + n_2$ éléments, il est alors possible d'écrire [49] :

$$\begin{aligned} n_1 \|\mathbf{X}_1 - \mathbf{Y}_1\|^2 + n_2 \|\mathbf{X}_2 - \mathbf{Y}_2\|^2 &= n \left\| \frac{1}{n} (n_1 \mathbf{X}_1 + n_2 \mathbf{X}_2) - \frac{1}{n} (n_1 \mathbf{Y}_1 + n_2 \mathbf{Y}_2) \right\|^2 \\ &\quad + \frac{n_1 n_2}{n} \|\mathbf{X}_1 - \mathbf{X}_2 - \mathbf{Y}_1 + \mathbf{Y}_2\|^2 \end{aligned} \quad (6.68)$$

En appliquant deux fois cette relation sur l'équation (6.52), on arrive à la simplification suivante :

$$\begin{aligned} T_{LRV} &= \frac{1}{\hat{\sigma}_1^2} \left(\frac{n_1 n_2}{n} \text{tr} [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2] - n_1 \text{tr} [(\mathbf{\Lambda}_1 - \hat{\mathbf{\Lambda}})^2] - n_2 \text{tr} [(\mathbf{\Lambda}_2 - \hat{\mathbf{\Lambda}})^2] \right) \\ &= \frac{n_1 n_2}{n \hat{\sigma}_1^2} \left(\text{tr} [(\bar{\mathbf{L}}_1 - \bar{\mathbf{L}}_2)^2] - \text{tr} [(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2)^2] \right) \end{aligned} \quad (6.69)$$

En développant le terme dans le trace et en utilisant la propriété d'invariance de la trace par rotation, on obtient le test sur les vecteurs propres :

$$T_{LRV} = \frac{2n_1 n_2}{n \hat{\sigma}_1^2} \left(\text{tr}(\mathbf{\Lambda}_1 \mathbf{\Lambda}_2) - \text{tr}(\bar{\mathbf{L}}_1 \bar{\mathbf{L}}_2) \right) \quad (6.70)$$

La limitation de ce test dans le cadre d'un suivi longitudinal réside dans le fait que les valeurs propres sont supposées égales pour les deux échantillons. En effet, on cherche à savoir si $\mathbf{M}_1$ et $\mathbf{M}_2$ ont les mêmes vecteurs propres en faisant l'hypothèse que les valeurs propres sont les mêmes. Si les matrices de vecteurs propres $\mathbf{V}_1$ et $\mathbf{V}_2$ sont les mêmes alors $T_{LRV} = 0$. Lorsque l'angle augmente entre $\mathbf{V}_1$ et $\mathbf{V}_2$, le produit interne entre $\bar{\mathbf{L}}_1$ et $\bar{\mathbf{L}}_2$ diminue ce qui a pour conséquence une augmentation de T_{LRV} pour des valeurs propres $\mathbf{\Lambda}_1$ et $\mathbf{\Lambda}_2$ fixées.

Intéressons nous à la distribution asymptotique de T_{LRV} [46]. La variance estimée peut être vue comme une variable aléatoire : $\frac{qn}{\sigma^2} \hat{\sigma}^2 \sim \chi^2(q(n-2)+p)$ et $\frac{2n_1 n_2}{n \sigma^2} \left(\text{tr}(\mathbf{\Lambda}_1 \mathbf{\Lambda}_2) - \text{tr}(\bar{\mathbf{L}}_1 \bar{\mathbf{L}}_2) \right) \sim$

$\chi^2(q-p)$. En combinant ces deux termes, on obtient le test T_{2VEC} qui suit asymptotiquement une loi de Fisher notée F :

$$\begin{aligned} T_{2VEC} &= \frac{\frac{1}{q-p} \frac{2n_1 n_2}{n\sigma^2} (tr(\mathbf{\Lambda}_1 \mathbf{\Lambda}_2) - tr(\bar{\mathbf{L}}_1 \bar{\mathbf{L}}_2))}{\frac{1}{q(n-2)+2p} \frac{qn}{\sigma^2} \hat{\sigma}_1^2} \\ &= \frac{2n_1 n_2 (q(n-2) + 2p)}{q(q-p)n\hat{\sigma}_1^2} (tr(\mathbf{\Lambda}_1 \mathbf{\Lambda}_2) - tr(\bar{\mathbf{L}}_1 \bar{\mathbf{L}}_2)) \sim F(q-p, q(n-2) + p) \end{aligned} \quad (6.71)$$

6.2.3 La distribution empirique sous l'hypothèse nulle

Pour seuiller une carte de statistique à un certain seuil de signification, il est nécessaire de connaître sa distribution sous l'hypothèse nulle. Les distributions théoriques ont été présentées dans les sections précédentes, mais elles résultent d'une modélisation purement théorique. En pratique, les données ne suivent pas exactement cette distribution théorique. Par exemple, l'hypothèse d'indépendance entre les éléments des échantillons n'est pas toujours respectée. Ainsi, les p-valeurs calculées à partir de la distribution sous l'hypothèse nulle théoriques ne sont pas toujours pertinentes.

Dans [20], l'auteur soutient l'idée qu'il est possible de confirmer les hypothèses faites dans la conception théorique du test statistique lorsque l'on effectue de nombreuses comparaisons et que la plupart des tests statistiques vérifient effectivement l'hypothèse nulle. En effet, un simple histogramme construit à partir des valeurs obtenues pour l'ensemble des tests permet de vérifier si les hypothèses faites dans la théorie sont raisonnables. L'auteur propose également d'utiliser cet histogramme pour ajuster les paramètres théoriques de la distribution sous l'hypothèse nulle afin que cette distribution reflète plus fidèlement la distribution des données sous H_0 . Dans le cadre du suivi longitudinal, il est possible d'estimer cette distribution empirique puisque les tests d'hypothèses sont calculés sur tout le cerveau et que les changements (*i.e.*, lésions ou modifications de la substance blanche apparemment normale) affectent une très petite partie de ces voxels.

Une distribution du χ^2 à l'échelle est ajustée sur les données (cette distribution est plus facile à ajuster que la loi de Fisher). Ainsi, la première étape consiste à transformer les statistiques des tests T_{1VP} , T_{2TD} , T_{2VP} et T_{2VEC} supposés suivre une loi $F(d_1, d_2)$ en une nouvelle statistique qui suit une loi $\chi^2(d_1)$. Pour ce faire, les p-valeurs sont tout d'abord calculées. Puis, une statistique de test suivant une loi $\chi^2(d_1)$ est associée à chaque p-valeur. À partir de ces nouvelles valeurs de la statistique, un histogramme est construit. Un nouveau degré de liberté et un facteur d'échelle sont estimés en ajustant la distribution du χ^2 à l'échelle sur l'histogramme seuillé à 90% (on peut raisonnablement supposer dans le cadre des pathologies étudiées que moins de 10% des voxels du cerveau subissent un changement au cours du temps). Ce seuillage permet d'écarter les voxels où l'hypothèse nulle est rejetée. Finalement, un seuil est sélectionné à partir des nouvelles p-valeurs non corrigées ou en appliquant le FDR pour tenir compte des comparaisons multiples (cf. section 4.3.2).

Conclusion partielle

Nous venons de faire un travail de synthèse bibliographique qui a permis de présenter la problématique de détection de changements en IRM de diffusion. Cet état de l'art a permis de soulever la problématique du recalage d'images en présence de zones en évolution et de s'intéresser aux différentes modélisations statistiques développées pour la détection de changements en IRM de diffusion. Ces méthodes tiennent compte de la nature multivariée des données, ainsi que de la géométrie de l'espace définie par le tenseur de diffusion. Afin de détecter les zones d'évolution significatives, différentes solutions au problème des comparaisons multiples ont également été présentées. La partie suivante propose deux chaînes de traitement pour le suivi longitudinal de patients atteints de pathologies neurodégénératives. Ces développements sont basés sur le modèle du tenseur d'ordre 2.

Bibliographie

- [1] H. ABDI : *Bonferroni and Sidak corrections for multiple comparisons*, p. 103 – 107. Sage, Thousand Oaks, CA, 2007.
- [2] D. ALEXANDER et J. GEE : Elastic matching of diffusion tensor images. *Computer Vision and Image Understanding*, 77(9):233 – 250, 2000.
- [3] D. ALEXANDER, C. PIERPAOLI, P. BASSER et J. GEE : Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Transaction on Medical Imaging*, 20(11):1131 – 1139, 2001.
- [4] E. ANGELINI, J. ATIF, J. DELON, E. MANDONNET, H. DUFFAU et L. CAPELLE : Detection of glioma evolution on longitudinal MRI studies. *In IEEE Internaional Symposium on Biomedical Imaging, Washington, ISBI '07*, p. 49 – 52, 2007.
- [5] E. ANGELINI, O. CLATZ, E. MANDONNET, E. KONUKOGLU, L. CAPELLE et H. DUFFAU : Glioma dynamics and computational models : A review of segmentation, registration and in silico growth algorithms and their clinical validations. *Current Medical Imaging Review*, 3(4):262 – 276, 2007.
- [6] V. ARSIGNY, P. FILLARD, X. PENNEC et N. AYACHE : Fast and simple calculus on tensors in the log-euclidean framework. *In Medical Image Computing and Computer Assisted Intervention, Palm Springs*, vol. 3749 de LNCS, p. 115 – 122, 2005.
- [7] V. ARSIGNY, P. FILLARD, X. PENNEC et N. AYACHE : Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411 – 421, 2006.
- [8] P. BASSER et S. PAJEVIC : A normal distribution for tensor-valued random variables : Applications to diffusion tensor MRI. *IEEE Transaction on Medical Imaging*, 22(7):785 – 794, 2003.
- [9] Y. BENJAMINI et Y. HOCHBERG : Controlling the false discovery rate : a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society series B-Methodological*, 57(1):289 – 300, 1995.
- [10] Y. BENJAMINI et D. YEKUTIELI : The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4):1165 – 1188, 2001.
- [11] H. BOISGONTIER, V. NOBLET, F. HEITZ, L. RUMBACH et J. ARMSPACH : An automatic method for change detection in serial DTI-derived scalar images. *In Medical*

- Image Computing and Computer Assisted Intervention, Workshop on MIAMS, New York, MICCAI '08*, 2008.
- [12] H. BOISGONTIER, V. NOBLET, F. HEITZ, L. RUMBACH et J. ARMSPACH : Generalized likelihood ratio tests for change detection in diffusion tensor images : Application to multiple sclerosis. *Medical Image Analysis*, 16(1):325 – 338, 2012.
 - [13] M. BOSC, F. HEITZ, J. ARMSPACH, I. NAMER, D. GOUNOT et L. RUMBACH : Automatic change detection in multimodal serial MRI : application to multiple sclerosis lesion evolution. *Neuroimage*, 20(2):643 – 656, 2003.
 - [14] Y. CAO, M. MILLER, S. MORI, R. WINSLOW et L. YOUNES : Diffeomorphic matching of diffusion tensor images. In *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, CVPRW '06, p. 67, 2006.
 - [15] J. CAREW, C. KOAY, G. WAHBA, A. ALEXANDER, P. BASSER et M. MEYERAND : The asymptotic distribution of diffusion tensor and fractional anisotropy estimates. *Magnetic Resonance in Medicine*, 16(1):2006 – 2006, 2006.
 - [16] Y. CHIKUSE : *Statistics on special manifolds*. Lecture notes in statistics. Springer, 2003.
 - [17] S. CHUNG, D. PELLETIER, M. SDIKA, Y. LU, J. BERMAN et R. HENRY : Whole brain voxel-wise analysis of single-subject serial DTI by permutation testing. *Neuroimage*, 39(4):1693 – 1705, 2008.
 - [18] O. CICCARELLI, M. CATANI, H. JOHANSEN-BERG, C. CLARK et A. THOMPSON : Diffusion-based tractography in neurological disorders : concepts, applications, and future developments. *The Lancet Neurology*, 7(8):715 – 727, 2008.
 - [19] I. COROUGE, P. FLETCHER, S. JOSHI, S. GOUTTARD et G. GERIG : Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Medical Image Analysis*, 10(5):786 – 798, 2006.
 - [20] B. EFRON : Large-scale simultaneous hypothesis testing : The choice of a null hypothesis. *Journal of the American Statistical Association*, 99(465):96 – 104, 2004.
 - [21] J. FARRELL, B. LANDMAN, C. JONES, S. SMITH, J. PRINCE, P. van ZIJL et S. MORI : Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T. *Journal of Magnetic Resonance Imaging*, 26(3):756 – 767, 2007.
 - [22] P. FLETCHER et S. JOSHI : Principal geodesic analysis on symmetric spaces : Statistics of diffusion tensors. In *European Conference on Computer Vision, Workshops CVAMIA and MMBIA, Prague*, vol. 3117 de LNCS, p. 87 – 98, 2004.
 - [23] P. FREEBOROUGH et N. FOX : Modeling brain deformations in Alzheimer disease by fluid registration of serial 3D MR images. *Journal of Computer Assisted Tomography*, 22(5):838 – 843, 1998.
 - [24] A. GUPTA et D. NAGAR : *Matrix variate distributions*. Chapman & Hall/CRC monographs and surveys in pure and applied mathematics. Chapman & Hall, 2000.
 - [25] K. HUA, J. ZHANG, S. WAKANA, H. JIANG, X. LI, D. REICH, P. CALABRESI, J. PEKAR, P. van ZIJL et S. MORI : Tract probability maps in stereotaxic spaces : Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, 39(1):336 – 347, 2008.

- [26] P. KHURD, R. VERMA et C. DAVATZIKOS : Kernel-based manifold learning for statistical analysis of diffusion tensor images. *In International Conference on Information Processing in Medical Imaging, Kerkrade, IPMI '07*, p. 581 – 593, 2007.
- [27] L. LEMIEUX, U. WIESHMANN, F. NICHOLAS, D. FISH et S. SHORVON : The detection and significance of subtle changes in mixed-signal brain lesions by serial MRI scan matching and spatial normalization. *Medical Image Analysis*, 2(3):227 – 242, 1998.
- [28] S. LOPE-PIEDRAFITA, M. GARCIA-MARTIN, J. GALONS, R. GILLIES et T. TROUARD : Longitudinal diffusion tensor imaging in a rat brain glioma model. *NMR Biomedicine*, 21(8):799 – 808, 2008.
- [29] M. MADDAH, L. ZÖLLEI, W. GRIMSON, C. WESTIN et W. WELLS : A mathematical framework for incorporating anatomical knowledge in DT-MRI analysis. *In IEEE International Symposium on Biomedical Imaging, Paris, ISBI '08*, p. 105 – 108, 2008.
- [30] T. NICHOLS et A. HOLMES : Nonparametric permutation tests for functional neuroimaging : A primer with examples. *Human Brain Mapping*, 15(1):1 – 25, 2002.
- [31] C. NIKOU, F. HEITZ et J. ARMSPACH : Robust voxel similarity metrics for the registration of dissimilar single and multimodal images. *Pattern Recognition*, 32(8):1349 – 1366, 1999.
- [32] V. NOBLET : *Recalage non rigide d'images cérébrales 3D avec contrainte de conservation de la topologie*. Thèse de doctorat, Université de Strasbourg, 2006.
- [33] L. O'DONNELL et C. WESTIN : Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Transactions on Medical Imaging*, 26(11):1562 – 1575, 2007.
- [34] J. OLLINGER, J. S.C. et A. ALEXANDER : Multivariate analysis of diffusion tensor data using hotelling T2 statistic. *In International Society for Magnetic Resonance in Medicine, Miami, ISMRM '05*, p. 1324, 2005.
- [35] J. O'RIORDAN, M. CAIN, A. COLES, L. WANG, D. COMPSTON, P. TOFTS et D. MILLER : T1 hypointense lesion load in secondary progressive multiple sclerosis : A comparison of pre versus post contrast loads and of manual versus semi automated threshold techniques for lesion segmentation. *Multiple Sclerosis*, 4(5):408 – 412, 1998.
- [36] S. PAJEVIC et P. BASSER : Parametric and non-parametric statistical analysis of DT-MRI data. *Journal of Magnetic Resonance*, 161(1):1 – 14, 2002.
- [37] J. PATRIARCHE et B. ERICKSON : Part 1. Automated change detection and characterization in serial MR studies of brain-tumor patients. *Journal of Digital Imaging*, 20(3):203 – 222, 2007.
- [38] X. PENNEC, P. FILLARD et N. AYACHE : A riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41 – 66, 2006.
- [39] S. PRICE, N. BURNET, T. DONOVAN, H. GREEN, A. PENA, N. ANTOUN, J. PICKARD, T. CARPENTER et J. GILLARD : Diffusion tensor imaging of brain tumours at 3T : a potential tool for assessing white matter tract invasion. *Clinical Radiology*, 58(6):455 – 462, 2003.

-
- [40] W. PRICE : Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion, part 1 : basic theory. *Magnetic Resonance : an Educational Journal*, 9(5):299 – 336, 1997.
 - [41] D. REY, G. SUBSOL, H. DELINGETTE et N. AYACHE : Automatic detection and segmentation of evolving processes in 3D medical images : Application to multiple sclerosis. *Medical Image Analysis*, 6:163 – 179, 2002.
 - [42] R. R.J., S. ANDRA, O. AL-KOFAHI et B. ROYSAM : Image change detection algorithms : a systematic survey. *IEEE Transactions on Image Process*, 14(3):294 – 307, 2005.
 - [43] F. ROUSSEAU : Comparaison d’images : analyse longitudinale. In *Ecole de printemps, ANG D CNRS, Traitement des images médicales : du voxel aux atlas*. <http://lsit-miv.u-strasbg.fr/ecoleTIM/download.html>, 2008.
 - [44] F. ROUSSEAU, S. FAISAN, F. HEITZ, J.-P. ARMSPACH, Y. CHEVALIER, F. BLANC, J. D. SEZE et L. RUMBACH : An a contrario approach for change detection in 3D multimodal images : Application to multiple sclerosis in MRI. In *IEEE Engineering in Medicine and Biology Society, Lyon, EMBC ’07*, p. 2069 – 2072, 2007.
 - [45] M. SCHLUTER, B. STIELTJES, H. HAHN, J. REXILIUS, O. KONRAD-VERSE et H. P. HO : Detection of tumour infiltration in axonal fibre bundles using diffusion tensor imaging. *The International Journal of Medical Robotics*, 1(3):80 – 86, 2005.
 - [46] A. SCHWARTZMAN : *Random ellipsoids and false discovery rates : statistics for diffusion tensor imaging data*. Thèse de doctorat, Université de Stanford, 2006.
 - [47] A. SCHWARTZMAN, R. DOUGHERTY et J. TAYLOR : Cross-subject comparison of principal diffusion direction maps. *Magnetic Resonance in Medicine*, 53(6):1423 – 1431, 2005.
 - [48] A. SCHWARTZMAN, R. DOUGHERTY et J. TAYLOR : Group comparison of eigenvalues and eigenvectors of diffusion tensors. *Journal of the American Statistical Association*, 105(490):588 – 599, 2010.
 - [49] A. SCHWARTZMAN, W. MASCARENHAS et J. TAYLOR : Inference for eigenvalues and eigenvectors of gaussian symmetric matrices. *Annals of Statistics*, 36(6):2886 – 2919, 2009.
 - [50] R. SIMOES et C. SLUMP : Change detection and classification in brain MR images using Change Vector Analysis. In *IEEE Engineering in Medicine and Biology Society, Boston, EMBC ’11*, p. 7803 – 7807, 2011.
 - [51] R. STEFANESCU, O. COMMOWICK, G. MALANDAIN, P. BONDIAU, N. AYACHE et X. PENNEC : Non-rigid atlas to subject registration with pathologies for conformal brain radiotherapy. In *Medical Image Computing and Computer Assisted Intervention, Palm Springs*, vol. 3216 de *LNCS*, p. 704 – 711, 2005.
 - [52] B. STIELTJES, M. SCHLÄ $\frac{1}{4}$ TER, B. DIDINGER, M. WEBER, H. HAHN, P. PARZER, J. REXILIUS, O. KONRAD-VERSE, H. PEITGEN et M. ESSIG : Diffusion tensor imaging in primary brain tumors : reproducible quantitative analysis of corpus callosum infiltration and contralateral involvement using a probabilistic mixture model. *Neuroimage*, 31(2):531 – 542, 2006.

- [53] J. STOREY : A direct approach to false discovery rates. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 64(3):479 – 498, 2002.
- [54] C. STUDHOLME : Dense feature deformation morphometry : Incorporating DTI data into conventional MRI morphometry. *Medical Image Analysis*, 12(6):742 – 751, 2008.
- [55] J. THIRION et G. CALMON : Deformation analysis to detect and quantify active lesions in three-dimensional medical image sequences. *IEEE Transactions on Medical Imaging*, 18(5):429 – 441, 1999.
- [56] A. TROPINE, G. VUCUREVIC, P. DELANI, S. BOOR, N. HOPF, J. BOHL et P. STOETER : Contribution of diffusion tensor imaging to delineation of gliomas and glioblastomas. *Journal of Magnetic Resonance Imaging*, 20(6):905 – 912, 2004.
- [57] R. VERMA, P. KHURD et C. DAVATZIKOS : On analyzing diffusion tensor images by identifying manifold structure using isomaps. *IEEE Transaction on Medical Imaging*, 26(6):772 – 778, 2007.
- [58] R. VERMA, P. KHURD, J. LOUGHEAD, R. GUR, R. GUR et C. DAVSATZIKOS : Manifold Based Morphometry applied to schizophrenia. In *IEEE International Symposium on Biomedical Imaging, Paris, ISBI '08*, p. 704 – 707, 2008.
- [59] B. WHITCHER, J. WISCO, N. HADJIKHANI et D. TUCH : Statistical group comparison of diffusion tensors via multivariate hypothesis testing. *Magnetic Resonance in Medicine*, 57(6):1065 – 1074, 2007.
- [60] K. WORSLEY, S. MARRETT, P. NEELIN, A. VANDAL, K. FRISTON et A. EVANS : A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4(1):58 – 73, 1996.
- [61] B. YEO, T. VERCAUTEREN, P. FILLARD, X. PENNEC, P. GOLLAND, N. AYACHE et O. CLATZ : Dti registration with exact finite-strain differential. In *IEEE International Symposium on Biomedical Imaging, Paris, ISBI '08*, p. 700 – 703, 2008.
- [62] E. ZACHARAKI, C. HOGEA, G. BIROS et C. DAVATZIKOS : A comparative study of biomechanical simulators in deformable registration of brain tumor images. *IEEE Transactions on Biomedical Engineering*, 55(3):1233 – 1236, 2008.
- [63] E. ZACHARAKI, D. SHEN, S. LEE et C. DAVATZIKOS : ORBIT : Multiresolution framework for deformable registration of brain tumor images. *IEEE Transactions on Medical Imaging*, 27(8):1003 – 1017, 2008.

PARTIE III

Suivi longitudinal des tenseurs d'ordre 2 : prétraitements et outils statistiques

Introduction

LA troisième partie décrit l'ensemble des développements et des contributions méthodologiques apportés par ce travail de thèse dans le domaine de la détection de changement en imagerie du tenseur de diffusion d'ordre 2. En chaque point d'une acquisition d'IRMd, la distribution des directions de diffusion des molécules d'eau est modélisée par un tenseur d'ordre 2. Pour détecter des changements, de nouvelles méthodes de détection statistique adaptées aux images du tenseur d'ordre 2 sont développées. Notons que cette problématique a fait l'objet d'un nombre limité de travaux dans la littérature (cf. chapitre 5).

Tout d'abord, le premier chapitre décrit les prétraitements communs aux différentes chaînes de traitement proposées pour le suivi longitudinal en imagerie de diffusion. Les prétraitements permettent de corriger certains artefacts d'acquisition, et de recalibrer les images de diffusion avant d'effectuer une détection des différences inter-images. Ensuite, dans le deuxième chapitre, différentes méthodes d'échantillonnage sont présentées pour constituer deux populations de tenseurs de diffusion à comparer en chaque voxel. Ces techniques exploitent l'information contenue dans le signal de diffusion, le voisinage spatial ou une combinaison de ces deux. De plus, lors de cette étape, la possibilité d'utiliser un *a priori* sur les faisceaux de fibres de la substance blanche est envisagée. Les populations ainsi extraites permettent de refléter la neuro-anatomie sous-jacente avec précision. Finalement, dans le troisième chapitre, deux chaînes de traitement sont proposées pour répondre à la problématique du suivi longitudinal de patients en IRMd. Ces techniques sont élaborées autour d'une modélisation de la diffusion par des tenseurs d'ordre 2 et s'appuient sur les prétraitements et les concepts présentés précédemment.

Les prétraitements des images

Sommaire

7.1	Les artefacts en imagerie de diffusion	121
7.1.1	Les courants de Foucault	122
7.1.2	L’effet de susceptibilité magnétique	122
7.2	Le recalage des images	122
7.3	L’interpolation et la réorientation des images de tenseurs . . .	123

AFIN de pouvoir quantifier les différences inter-images dans le cadre d’une modélisation statistique, différents prétraitements sont requis. En effet, il est nécessaire de corriger certains artefacts liés à l’acquisition : les distorsions géométriques induites par les courants de Foucault et les effets de susceptibilité magnétique. Ensuite, pour effectuer une détection statistique des différences inter-images, une étape de recalage non-linéaire est requise. Un champ de déformation est estimé. Ce dernier décrit en chaque voxel la déformation à appliquer pour mettre en correspondance les voxels des deux images de tenseurs de diffusion à comparer. Appliquer un champ de déformation en IRMd est une opération délicate. Tout d’abord, l’interpolation du modèle de diffusion en chaque voxel (représenté par une forme géométrique, ici un tenseur) requiert certaine précaution, puis la nature multivariée des données et le fait qu’elles représentent une information directionnelle imposent une réorientation des tenseurs en chaque voxel de l’image.

7.1 Les artefacts en imagerie de diffusion

En imagerie de diffusion, obtenir des données fiables et reproductibles pour réaliser des inférences statistiques est un challenge. Une revue des différents écueils que l’on peut rencontrer est proposée dans [17]. Toute chaîne de traitement commence par l’acquisition de données brutes, suivie par une chaîne de traitement complexe. Chaque bloc peut entraîner un biais plus ou moins significatif, pouvant altérer la précision des résultats. Puisque l’acquisition des séquences de diffusion par des techniques d’imagerie écho-planaire est connue pour être génératrice de nombreux artefacts, nous devons corriger au mieux les données.

7.1.1 Les courants de Foucault

Les courants de Foucault sont des courants électriques induits par des changements rapides du champ magnétique. Ils sont à l'origine de variations locales des gradients de champ magnétique qui perturbent le signal recueilli au temps d'écho (la position et la phase du signal au temps d'écho sont décalées). Lors de l'acquisition des images de diffusion, les courants de Foucault sont responsables de distorsions géométriques dans les images. Ainsi, chaque coupe d'une image peut être compressée, décalée ou étirée.

L'existence de ces distorsions géométriques peut être gênante lors de la modélisation du signal de diffusion (*i.e.*, si les coupes ne sont pas alignées, on mélange des signaux provenant de structures différentes, en conséquence, le tenseur estimé n'est en rien pertinent). Ces artefacts peuvent être minimisés à l'acquisition ou être en partie corrigés après l'acquisition. Puisque les images de diffusion disponibles présentent cet artefact, ces dernières sont traitées *a posteriori* afin d'en minimiser les effets. Pour ce faire, il faudrait effectuer un recalage affine coupe à coupe par rapport à une image de référence (l'image sans gradient de diffusion) [17]. Dans notre cas, cette correction est réalisée avec FSL¹ qui effectue un recalage affine global entre une image pondérée en diffusion et l'image de référence de la séquence de diffusion.

7.1.2 L'effet de susceptibilité magnétique

L'acquisition écho-planaire est très sensible aux effets de susceptibilité magnétique. Chaque élément constitutif de la région étudiée (tissu, air, os, ...) est caractérisé par une susceptibilité magnétique spécifique. Le champ magnétique B_0 est localement modifié ce qui provoque, dans les images, des distorsions au niveau des zones d'interface entre les différents tissus. Plus la différence de susceptibilité magnétique entre les tissus est importante et plus la distorsion est prononcée.

Cet artefact engendre de légères déformations de l'image au niveau des interfaces aire/tissus ou os/tissus. Ces déformations sont induites par les inhomogénéités de champ magnétique à l'interface des tissus qui provoquent des déphasages. Une correction de l'effet de susceptibilité magnétique nécessite l'acquisition d'une carte du champ statique B_0 [17]. Généralement, en routine clinique, aucune carte de B_0 n'est acquise. De ce fait, aucune correction de ce type n'est appliquée dans nos prétraitements. Néanmoins, le recours à une séquence en écho de spin limite les effets de cet artefact grâce aux impulsions 180° de repasage.

7.2 Le recalage des images

La comparaison voxel à voxel de deux examens requiert une mise en correspondance des structures anatomiques par un recalage préalable des images. Les acquisitions sont tout d'abord recalées grâce à une transformation affine estimée en maximisant l'information mutuelle entre les images de fraction d'anisotropie (FA). Par rapport au recalage rigide, le

¹<http://www.fmrib.ox.ac.uk/fsl/>

recalage affine permet également de compenser certaines distorsions géométriques globales.

Après cette étape, d'autres distorsions géométriques globales peuvent subsister. En effet, une transformation affine ne permet pas par exemple de prendre en compte l'atrophie cérébrale. Un tel changement, dû au vieillissement du cerveau ou à l'évolution globale d'une maladie neurodégénérative, ne correspond pas à un changement d'intérêt dans notre cas. Le recalage est donc amélioré avec une méthode de recalage déformable [22]. Cette méthode est basée sur une représentation paramétrique multi-échelle du champ de déformation dans une base de fonctions B-splines. L'estimation de la transformation est effectuée au sens des moindres carrés à partir des images de FA grâce à une approche hiérarchique et en utilisant la méthode de Levenberg-Marquardt. Une telle approche, autorisant un grand nombre de degrés de liberté, permet de compenser presque toutes les différences entre les deux images. Or nous ne souhaitons pas compenser les modifications pathologiques d'intérêt, qui sont supposées être des changements focaux (ex : lésions de sclérose en plaques). Pour ce faire, la taille finale des supports des fonctions B-splines est imposée à être sensiblement plus grande que la taille des changements à détecter (dans notre cas, $7 \times 7 \times 7$ points de contrôle sur l'image). La transformation estimée à partir des images de FA est ensuite appliquée sur les images de tenseurs de diffusion préalablement estimées.

7.3 L'interpolation et la réorientation des images de tenseurs

L'interpolation

Une méthode d'interpolation est nécessaire pour rééchantillonner l'image de tenseur déformée. La métrique utilisée lors de cette étape va engendrer des résultats très différents [2]. Dans [2], il apparaît clairement qu'une interpolation linéaire dans le domaine log-euclidien donne de meilleurs résultats que dans le domaine euclidien. La figure 11.3 présente le résultat d'une interpolation linéaire dans le domaine log-euclidien et le résultat d'une interpolation linéaire des valeurs propres couplée à une interpolation appropriée (qui conserve les propriétés) de la matrice de rotation [15, 27] pour un exemple synthétique. L'interpolation séparée de la matrice des valeurs propres et des vecteurs propres permet de conserver la forme des tenseurs. Néanmoins, dans certains cas, des comportements étranges ont été observés (par exemple, la colonne de tenseurs encadrés dans la figure 11.3-b ne semble pas représenter le chemin le plus court pour interpoler les tenseurs). Nous utiliserons donc une méthode de la littérature, à savoir une interpolation B-spline d'ordre 3 (plus précise) dans le domaine log-euclidien. L'interpolation séparée des valeurs propres et vecteurs propres est laissée comme une perspective de notre travail.

La réorientation des tenseurs d'ordre 2

Appliquer une transformation géométrique sur une image de tenseurs nécessite une étape de réorientation qui conserve une orientation relative cohérente des différentes structures anatomiques (on préserve l'orientation relative des tenseurs par rapport à un référentiel 3D lié

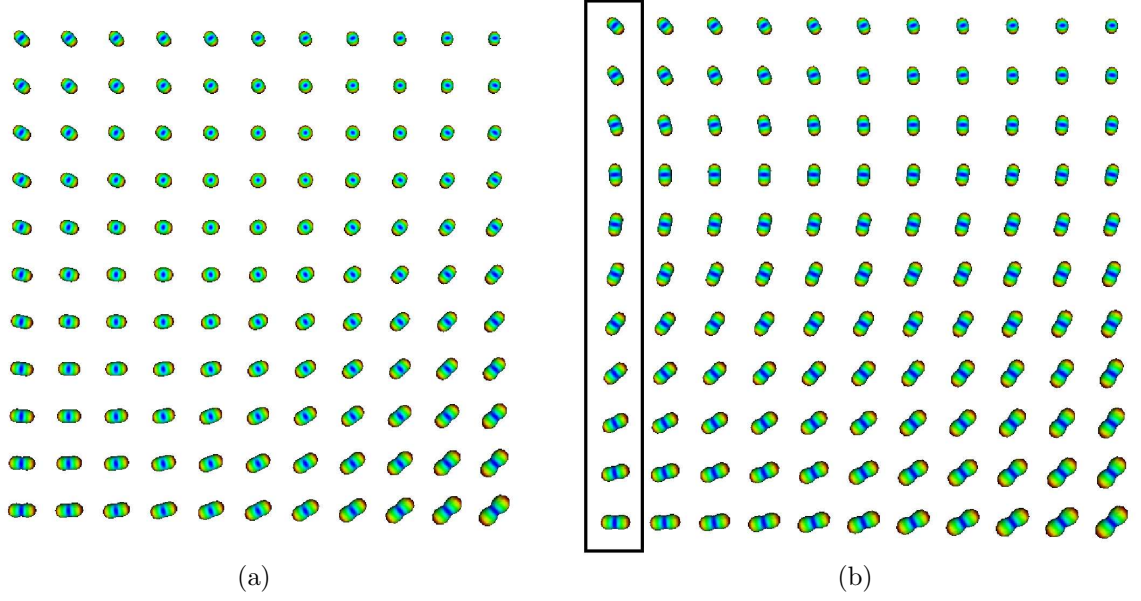


FIG. 7.1 – (a) interpolation linéaire dans le domaine log-euclidien et (b) interpolation linéaire des valeurs propres couplée avec l'interpolation de la matrice de rotation représentée comme un quaternion.

au volume de l'image). Soient $H : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un champ de déformation et $I : \mathbb{R}^3 \rightarrow \text{Sym}(3)^+$ un champ de tenseurs avec $\text{Sym}(3)^+$ l'espace des tenseurs d'ordre 2 symétriques définis positifs. L'objectif est d'estimer le champ de tenseur déformé I' après application de la transformation géométrique. La prise en compte des modifications d'orientation induites par la transformation H se fait par une transformation locale. Elle est appliquée sur chaque tenseur $\mathbf{D}$ d'ordre 2 afin d'obtenir un tenseur réorienté $\mathbf{D}'$:

$$\mathbf{D}' = \mathbf{R}^T \mathbf{D} \mathbf{R}, \quad (7.1)$$

où $\mathbf{R}$ est une matrice de rotation qui permet de conserver au mieux les propriétés d'orientation. Idéalement, cette étape de réorientation doit préserver l'information d'orientation et de forme contenues dans le tenseur. La matrice Jacobienne $\mathbf{F}(x, y, z)$ de la transformation H est calculée en chaque voxel pour approcher localement le déplacement non-linéaire par une transformation affine. L'expression de $\mathbf{F}$ pour un déplacement $\vec{d}(x, y, z) = (u(x, y, z), v(x, y, z), w(x, y, z))$ est donnée par :

$$\begin{aligned} \mathbf{F}(x, y, z) &= \begin{pmatrix} \frac{\partial(x+u(x,y,z))}{\partial x} & \frac{\partial(x+u(x,y,z))}{\partial y} & \frac{\partial(x+u(x,y,z))}{\partial z} \\ \frac{\partial(y+v(x,y,z))}{\partial x} & \frac{\partial(y+v(x,y,z))}{\partial y} & \frac{\partial(y+v(x,y,z))}{\partial z} \\ \frac{\partial(z+w(x,y,z))}{\partial x} & \frac{\partial(z+w(x,y,z))}{\partial y} & \frac{\partial(z+w(x,y,z))}{\partial z} \end{pmatrix} \\ &= \begin{pmatrix} 1 + \frac{\partial u(x,y,z)}{\partial x} & \frac{\partial u(x,y,z)}{\partial y} & \frac{\partial u(x,y,z)}{\partial z} \\ \frac{\partial v(x,y,z)}{\partial x} & 1 + \frac{\partial v(x,y,z)}{\partial y} & \frac{\partial v(x,y,z)}{\partial z} \\ \frac{\partial w(x,y,z)}{\partial x} & \frac{\partial w(x,y,z)}{\partial y} & 1 + \frac{\partial w(x,y,z)}{\partial z} \end{pmatrix} \end{aligned} \quad (7.2)$$

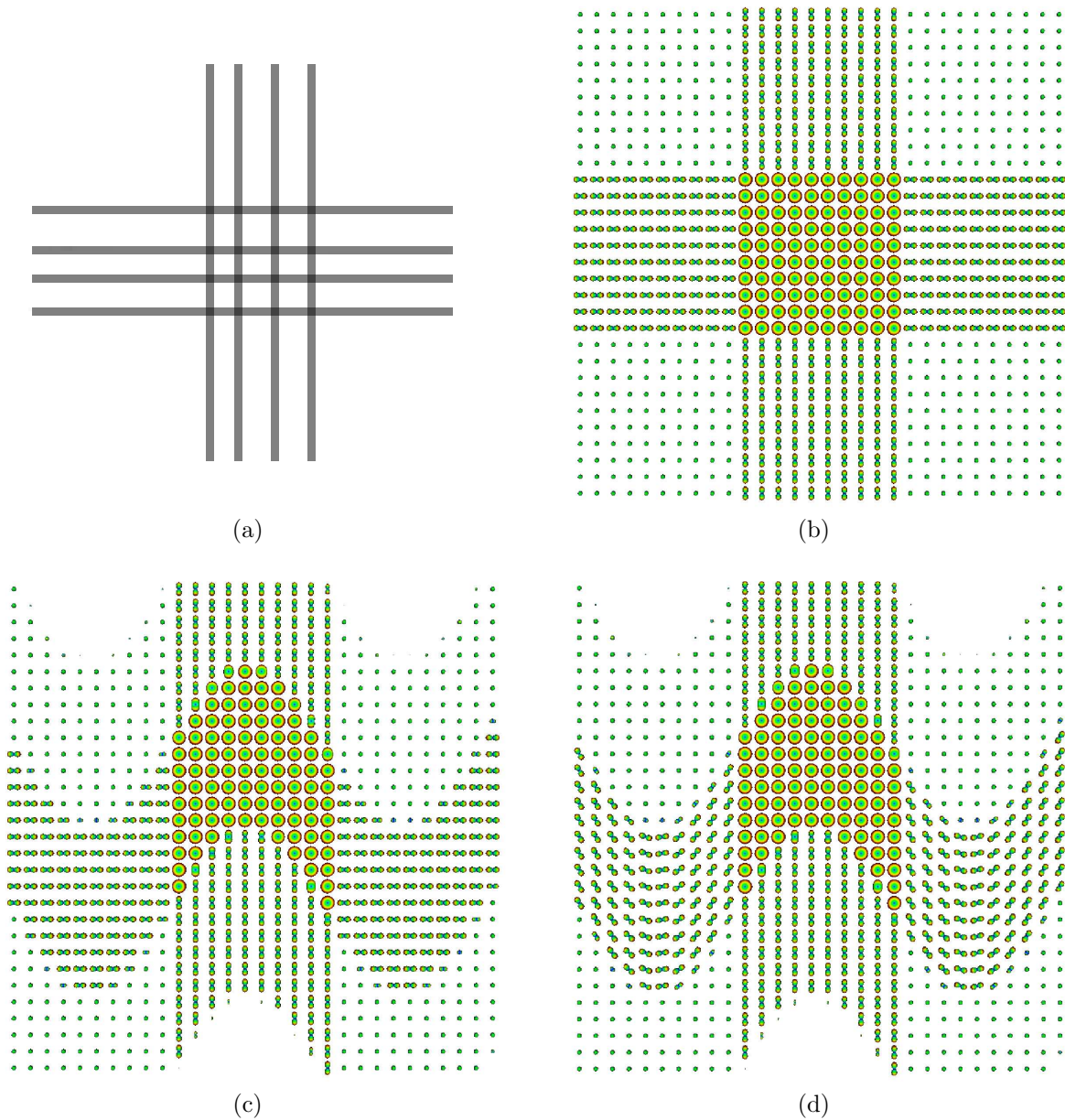


FIG. 7.2 – Illustration de la réorientation PPD : (a) le gabarit du croisement de fibres synthétique, (b) les tenseurs d'ordre 2 associés à ce croisement de fibres synthétique, puis application du champs de déformation, (c) image de tenseurs interpolés et (d) image de tenseurs interpolés et réorientés.

$\mathbf{F}$ peut être exprimée comme le produit d'une matrice de rotation $\mathbf{R}$ et d'une composante déformable $\mathbf{U}$: $\mathbf{F} = \mathbf{R}\mathbf{U}$. En ne considérant que la matrice de rotation $\mathbf{R}$ et en l'injectant dans l'équation 7.1, on obtient la méthode de réorientation rigide ou, en anglais, *Finite Strain* (FS) [1]. Toutefois, cette stratégie ne prend en compte que la transformation estimée à partir du champ de déformation et pas les tenseurs observés. Une stratégie alternative, la préservation des directions principales ou, en anglais, *Preservation of Principal Direction*

(PPD) [1], permet de lever cette limitation. Soient $\vec{e}_1, \vec{e}_2, \vec{e}_3$ les vecteurs propres de $\mathbf{D}$ associés aux trois valeurs propres triées par ordre croissant. La réorientation PPD consiste à estimer une rotation permettant d'aligner le premier vecteur propre $\vec{e}_1$ avec $\mathbf{F}\vec{e}_1$ et de basculer le deuxième vecteur propre $\vec{e}_2$ dans le plan défini par $\mathbf{F}\vec{e}_1$ et $\mathbf{F}\vec{e}_2$. La figure 7.2 illustre les résultats obtenus avec la réorientation PPD pour un champ de déformation synthétique de forme sinusoïdale.

Nous venons de présenter les prétraitements permettant de corriger des images de diffusion et d'aligner les tenseurs d'ordre 2 associés à deux image de diffusion à comparer. Ces outils seront utilisés par nos chaînes de traitement lors du suivi longitudinal de patients (cf. chapitre 9). Avant de présenter ces chaînes de traitement, des stratégies permettant d'engendrer en chaque voxel deux populations de tenseurs d'ordre 2 doivent être imaginées. Cette étape est très importante et n'est pas immédiate lors d'un suivi longitudinal. Le prochain chapitre présente les différentes stratégies proposées pour répondre à cette problématique.

Extraction de populations de tenseurs

Sommaire

8.1	Prise en compte du voisinage spatial à l'échelle du voxel	127
8.2	Méthode de rééchantillonnage : le bootstrap	128
8.2.1	Le Bootstrap Local (BL)	129
8.2.2	Le Bootstrap Spatial (BS)	132
8.3	Prise en compte d'un voisinage basé sur une modélisation des faisceaux de fibres de la substance blanche	133
8.3.1	Modélisation des fibres de la substance blanche : la tractographie .	133
8.3.2	Regroupement des fibres en faisceaux	135
8.3.3	Estimation des paramètres dans une section de faisceau	139

ON se propose de comparer différentes stratégies pour engendrer une population de tenseurs caractérisant la variabilité en chaque voxel de l'image. Dans le cadre d'études de groupes, en un voxel, les populations à comparer sont directement construites en considérant les tenseurs dans chacun des groupes ramené vers un espace commun. Cette étape est moins évidente lors du suivi dans le temps d'un même patient. Pour répondre à ce problème, nous proposons d'exploiter l'information contenue dans le signal de diffusion (par des techniques de bootstrap), le voisinage spatial ou une combinaison de ces deux. De plus, lors de cette étape, un *a priori* sur les faisceaux de fibres de la substance blanche est utilisé afin d'extraire des populations reflétant la neuro-anatomie sous-jacente.

8.1 Prise en compte du voisinage spatial à l'échelle du voxel

La prise en compte d'un voisinage spatial cubique, de taille fixée par l'utilisateur, est une façon immédiate de construire une population de tenseurs (cf. figure 8.1). Ainsi, on fait l'hypothèse d'un modèle constant par morceaux et la variabilité observée dans le voisinage est supposée provenir d'une perturbation aléatoire [5]. Les limitations d'un modèle constant par

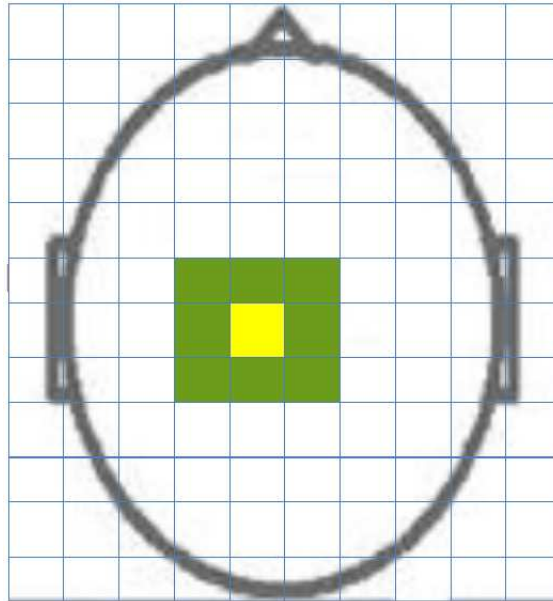


FIG. 8.1 – Illustration d'un voisinage spatial 2D de taille 3×3 .

morceaux sont bien connues, notamment à l'interface entre différents tissus. Cette stratégie reste néanmoins la plus utilisée dans la littérature, le modèle étant connu pour être robuste et efficace. De plus, cette stratégie est peu coûteuse en temps de calcul mais n'exploite pas toute l'information contenue dans les images de diffusion. Notons que le choix de la taille (voisinage non nécessairement isotrope en fonction de l'anisotropie des images) et de la géométrie du voisinage peut permettre de générer une population de tenseurs respectant plus fidèlement les hypothèses des tests statistiques. On a fait le choix d'un voisinage cubique de taille fixée par l'utilisateur qui ne tient pas compte de l'information de l'image. Cependant d'autres voisinages peuvent être utilisés (cf. [16] ou section 8.3).

8.2 Méthode de rééchantillonnage : le bootstrap

Lors de la prise en compte d'un voisinage spatial à l'échelle du voxel, les tenseurs de diffusion sont traités comme des observations. Or, un tenseur de diffusion modélise le processus physique de diffusion dont les données mesurées sont contenues dans les images de diffusion. La variabilité ainsi observée dépend du niveau de bruit dans ces données et de l'inadéquation des données avec le modèle de diffusion qui est supposé gaussien. Une idée alternative consiste à apprendre la variabilité inhérente aux images de diffusion en utilisant une procédure de bootstrap (cf. figure 8.2). Dans la littérature, différentes procédures de bootstrap ont été proposées pour générer un groupe de tenseurs [25, 9, 30, 12, 8]. Ces procédures permettent une estimation non-paramétrique de la distribution d'un estimateur (un paramètre d'intérêt) en utilisant un Processus Générateur de Données (PGD) simple, bien souvent un rééchantillonnage des données. La mise en œuvre du bootstrap consiste à construire un PGD qui génère des données qui ressemblent le plus possible aux données réelles. Une procédure de Bootstrap Local (BL) permet de modéliser la probabilité d'ob-

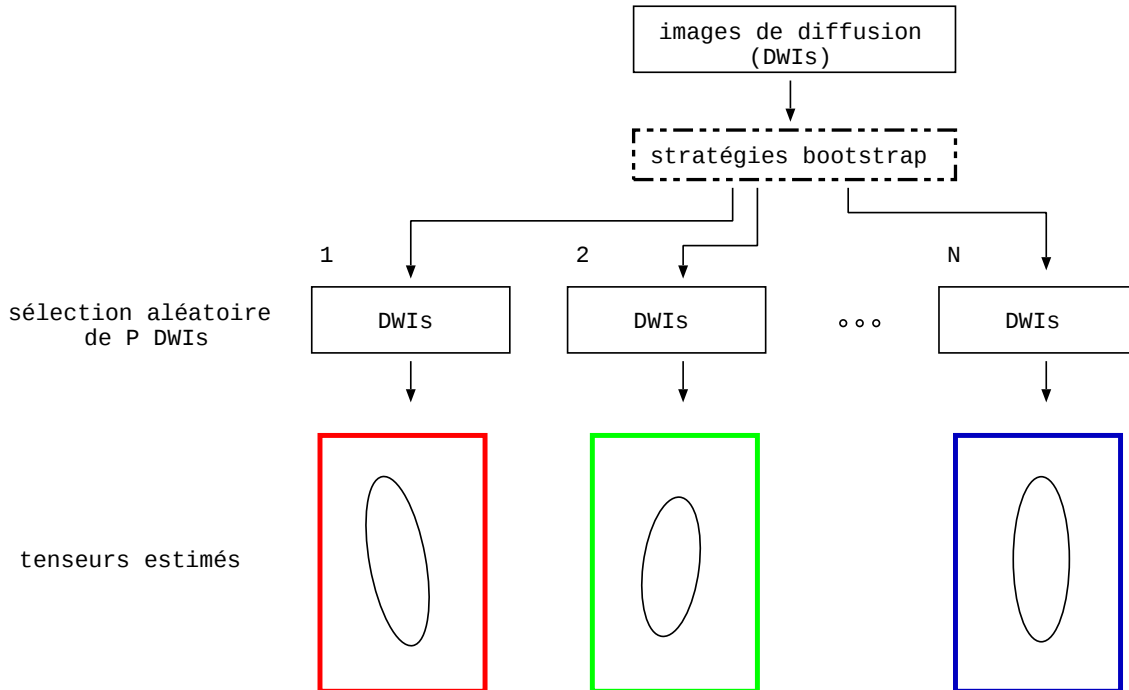


FIG. 8.2 – Illustration de l'idée sous-jacente à une procédure de bootstrap.

server un tenseur $\mathbf{D}$ en un voxel v sachant les mesures de diffusion à ce voxel : $S_i(v) : P(\mathbf{D}(v)|S_1(v), \dots, S_P(v))$. Une procédure de Bootstrap Spatial (BS) permet de modéliser la probabilité d'observer un tenseur en un voxel v étant données les mesures de diffusion d'un voisinage spatial $\mathcal{N}(v) : P(\mathbf{D}(v)|S_1(u), \dots, S_P(u); u \in \mathcal{N}(v))$.

Rappelons que l'estimation d'un tenseur de diffusion est faite à partir des mesures de diffusion en un voxel v et revient à résoudre un système linéaire de la forme (cf. section 2.2) : $\vec{y} = \mathbf{X}\vec{d} + \vec{\epsilon}$. Soit $\vec{g}_i = (g_{i,1}, g_{i,2}, g_{i,3})$ la direction de gradient de l'observation i ($i = 1, \dots, P$) et b la b-valeur. $\vec{y}$ est un vecteur de taille P contenant le logarithme du signal normalisé ($\log(S_i/S_0)$), $\mathbf{X}$ est la matrice de dessin (à l'ordre 2, une matrice de taille $P \times 6$ avec pour expression pour la i -ème ligne : $(-bg_{i,1}^2, -bg_{i,2}^2, -bg_{i,3}^2, -2bg_{i,1}g_{i,2}, -2bg_{i,1}g_{i,3}, -2bg_{i,2}g_{i,3})$), $\vec{d}$ est un vecteur contenant les N_l (à l'ordre 2, les six) paramètres du tenseur à estimer et $\vec{\epsilon}$ contient les résidus de l'estimation.

8.2.1 Le Bootstrap Local (BL)

Les méthodes de bootstrap dans les modèles de régression sont nombreuses. En ITD, il a été proposé de construire un PGD à partir d'acquisitions répétées de chaque direction de gradient [25]. Malheureusement, cette procédure n'est pas applicable dans de nombreuses situations en routine clinique, puisque, lors de l'acquisition des images de diffusion, seulement une acquisition est effectuée pour un nombre relativement important de directions d'étude. D'autres méthodes ont été développées pour un modèle de régression (linéaire dans le cadre de notre application) avec des résidus i.i.d. ou hétéroscédastiques.

Résidus i.i.d.

$\vec{\epsilon}$ contient les résidus de l'estimation supposés indépendants, dont la distribution est notée F . Cette loi est supposée d'espérance nulle et de variance σ^2 . Dans notre cas, F est supposé gaussien puisque les tenseurs sont estimés au sens des moindres carrés, mais en pratique, la distribution de F est différente et inconnue. Il faut donc une estimation convergente de cette loi des résidus $\bar{F}$, appelée fonction de distribution empirique ou loi de rééchantillonnage. On parle alors de bootstrap non-paramétrique. La génération de données bootstrap consiste alors à effectuer des tirages aléatoires à partir de $\bar{F}$. Pour que le bootstrap soit valide, la loi de rééchantillonnage doit être d'espérance nulle et les résidus doivent donc être centrés. Puisque le modèle de régression linéaire est résolu avec une méthode des moindres carrés, les résidus sont par construction centrés. Ensuite, comme dans [9], il est possible de standardiser les résidus pour que la variance de la fonction de rééchantillonnage soit un estimateur sans biais de la variance des résidus du modèle. Supposons que la variance des résidus est connue et de la forme [9] :

$$Var\{\epsilon_i\} = \frac{\sigma^2}{S_i^2} \quad (8.1)$$

où σ est l'écart type du bruit dans le signal de diffusion et S_i est la valeur du signal dans la direction g_i . σ est supposé constant quelle que soit l'intensité du signal de diffusion mesurée. Les résidus standardisés r_i s'écrivent alors [9] :

$$r_i = \frac{\epsilon_i}{w_i^{-1/2}(1 - h_i)^{1/2}} \quad (8.2)$$

où h_i est le i -ème terme diagonal de la matrice $\mathbf{H}$ définie par $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ et w_i est le i -ème élément diagonal de la matrice de pondération $\mathbf{W}$. L'expression de $\mathbf{W}$ est donnée par $\mathbf{W} = \exp(\mathbf{X}\vec{d})$, où $\vec{d}$ contient l'estimée des éléments du tenseur de diffusion. Les résidus standardisés et centrés $\vec{\epsilon}^*$ sont obtenus en retranchant la moyenne des résidus standardisés à tous les résidus standardisés : $\vec{\epsilon}^* = \vec{r} - \vec{\bar{r}}$ [9]. Finalement, le PGD du bootstrap basé sur les résidus, ou, en anglais, *Residual Bootstrap* (RB) s'écrit :

$$y_i^* = (\mathbf{X}\vec{d})_i + w_i^{-1/2}\epsilon_i^* \quad (8.3)$$

où ϵ_i^* est tiré aléatoirement dans la liste des résidus standardisés et centrés ϵ^* .

Résidus hétéroscédastiques

Lors de la régression linéaire, la variance des résidus peut ne pas être la même suivant chaque direction de gradient. On parle alors d'hétéroscédasticité. La forme de l'hétéroscédasticité peut dépendre de la matrice de dessin (ou matrices des régresseurs) et dans ce cas, on ne peut plus faire de tirage aléatoire des résidus indépendamment des régresseurs. Dans la littérature, il existe deux méthodes pour contourner cette difficulté, le bootstrap par paires (que l'on utilise dans la suite comme stratégie de bootstrap local) [12] et le wild bootstrap [9, 30].

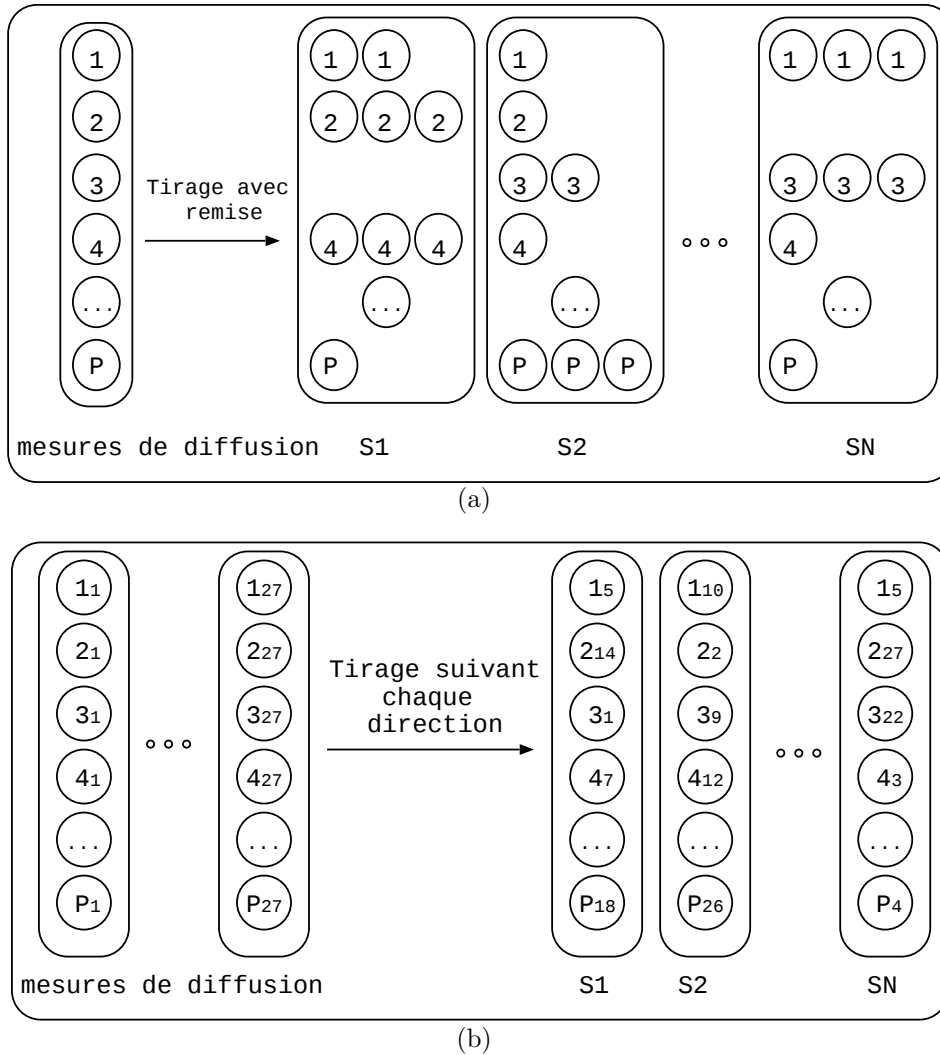


FIG. 8.3 – Représentation schématique des méthodes (a) bootstrap local (BL) et (b) bootstrap spatial (BS) (pour un voisinage $3 \times 3 \times 3$). 1₂₇ correspond au signal de diffusion du voxel 27 du voisinage dans la première direction.

Le Bootstrap par Paires (BP) : dans sa forme initiale, le bootstrap par paires consiste à rééchantillonner directement le couple observations/régresseurs à partir des mesures de diffusion : un échantillon bootstrap $(\bar{y}^*, \mathbf{X}^*)$ est obtenu en tirant aléatoirement avec remise les lignes de la matrice $(\bar{y}, \mathbf{X})$. En ITD, dans le cadre d'un suivi longitudinal, la méthode de bootstrap par paires consiste à générer en chaque voxel N échantillons (un échantillon contient P mesures de diffusion) en tirant aléatoirement avec remise P mesures de diffusion ainsi que les directions de gradient associées (cf. figure 8.3-a). Cette procédure revient à associer à chaque direction de gradient un poids différent lors de l'estimation du tenseur de diffusion.

Le Wild Bootstrap (WB) : au lieu de regrouper tous les résidus standardisés et centrés

pour effectuer un tirage, le wild bootstrap rééchantillonne les résidus associés à chaque direction de gradient en les multipliant par une fonction aléatoire. On tient ainsi compte de l'hétéroscédasticité des résidus (*i.e.*, $E(\epsilon_i^2|\mathbf{X}) = \sigma_i^2$) tout en conservant l'hypothèse d'indépendance des résidus dans l'échantillon bootstrap. Le PGD du wild bootstrap s'écrit :

$$y_i^* = (\mathbf{X}\vec{d})_i + \epsilon_i^* \quad (8.4)$$

où l'erreur rééchantillonnée ϵ_i^* est de la forme :

$$\epsilon_i^* = \frac{\epsilon_i}{(1 - h_i)^{1/2}} t \quad (8.5)$$

où t est la réalisation d'une fonction de distribution quelconque d'espérance $E(t) = 0$ et de variance $E(t^2) = 1$. En pratique, la distribution de Redemacher F_2 est utilisée :

$$F_2 : \quad t = \begin{cases} 1 & \text{avec comme probabilité associée } 0,5 \\ -1 & \text{avec comme probabilité associée } 0,5 \end{cases} \quad (8.6)$$

Il a été montré que parmi toutes les fonctions de distribution proposées dans la littérature, la distribution de Redemacher donne des performances toujours supérieures [9].

Nous choisissons le bootstrap par paires comme méthode de bootstrap local. Le principal inconvénient de cette méthode est qu'elle considère les variables explicatives comme des variables aléatoires alors qu'elles devraient être fixées lors de l'expérience [8]. De plus, la condition d'indépendance des résidus dans l'échantillon bootstrap n'est pas respectée puisqu'une même direction peut être tirée plusieurs fois. L'avantage du bootstrap par paires est qu'il permet de prendre en compte l'inadéquation des données au modèle du tenseur de diffusion d'ordre 2 (*i.e.*, ce modèle fait l'hypothèse que le processus de diffusion est gaussien et qu'un seul faisceau dominant est imagé en chaque voxel de l'image). Or aux résolutions d'étude actuelles en ITD, de nombreux faisceaux de fibres de la substance blanche se « croisent ».

8.2.2 Le Bootstrap Spatial (BS)

Le bootstrap spatial est une méthode qui se situe à mi-chemin entre le bootstrap local et la méthode du voisinage spatial à l'échelle du voxel. A partir d'un voisinage spatial de taille fixée par l'utilisateur, N échantillons bootstrap sont générés en tirant aléatoirement et pour chaque direction de gradient une mesure de diffusion (cf. figure 8.3-b). C'est une implémentation un peu différente de la procédure de bootstrap proposée dans [25] qui nécessite plusieurs répétitions de la mesure pour chaque direction de gradient. La méthode de bootstrap spatial est basée sur l'hypothèse d'un modèle constant par morceaux sur la région d'étude, ce qui permet de considérer toutes les mesures de diffusion du voisinage comme des répétitions de la mesure du voxel courant.

Avec la méthode de bootstrap spatial, on ne tient pas seulement compte de l'effet du bruit sur la distribution des tenseurs, mais également du fait que les données ne correspondent pas nécessairement à un modèle de diffusion gaussien, ainsi que des possibles erreurs induites par le recalage des images de diffusion lors de la correction des courants de Foucault.

8.3 Prise en compte d'un voisinage basé sur une modélisation des faisceaux de fibres de la substance blanche

Lors d'un test statistique, on fait l'hypothèse que chaque échantillon provient d'une même distribution. Dans la plupart des cas, considérer un voisinage spatial cubique pour extraire une population de tenseurs ne permet pas de vérifier cette hypothèse, par exemple à l'interface entre différents tissus. Pour éviter ce problème, nous proposons d'utiliser la géométrie locale des fibres de la substance blanche pour extraire une population de tenseurs issue d'une même distribution. Plusieurs outils sont nécessaires pour effectuer cette opération. Tout d'abord, un algorithme de tractographie permet de créer une modélisation des fibres de la substance blanche. Ensuite, un algorithme de *clustering* permet de regrouper les fibres « similaires » ensemble pour créer un modèle de l'organisation des faisceaux de fibres dans la substance blanche. Pour finir, une fibre dite de référence est sélectionnée pour chaque faisceau afin d'extraire les populations de tenseurs dans les sections transversales au faisceau.

8.3.1 Modélisation des fibres de la substance blanche : la tractographie

En neuroimagerie, l'imagerie du tenseur de diffusion permet d'extraire un modèle des fibres de la substance blanche, également appelé tractographie (en référence au nom de l'algorithme utilisé pour le calcul de ce modèle). La tractographie offre la possibilité d'étudier, *in vivo* et de manière non invasive, les grands faisceaux anatomiques qui composent la substance blanche cérébrale et ainsi d'inférer la connectivité anatomique cérébrale. L'algorithme de tractographie exploite l'effet des inhomogénéités du milieu sur le mouvement des molécules d'eau afin de déterminer pour chaque voxel la ou les directions principales de diffusion. Rappelons que le diamètre des fibres myélinisées (axones) varie de 1 à 30 μm et que le diamètre des faisceaux de fibres varie de 1 à 20 mm . Or, la résolution actuelle de l'imagerie de diffusion est de l'ordre de 8 mm^3 (pour une résolution spatiale de $2 \times 2 \times 2 \text{ mm}^3$ fréquemment utilisée en routine clinique). Ainsi, il s'agit d'étudier des micro-structures dont l'échelle est plus petite que la résolution autorisée par l'imagerie. Dans la littérature, il existe deux grandes familles d'algorithmes de tractographie : les algorithmes dits déterministes [21, 3, 29, 7] ou probabilistes [26, 18, 13, 4, 28]. La plupart de ces techniques permettent de reconstruire les principaux faisceaux de fibres : corps calleux, faisceau cortico-spinal, faisceau longitudinal supérieur, faisceau longitudinal inférieur, faisceau unciné, faisceau cingulaire, capsule externe, capsule extrême (cf. figure 8.4). L'enjeu des innovations actuelles en tractographie porte sur la modélisation des faisceaux d'association courts dits en « U » (cf. figure 8.4) qui relient différentes zones fonctionnelles sous-corticales.

Les méthodes déterministes sont simples à mettre en œuvre et nécessitent peu de ressources, mais elles sont très sensibles au bruit (c'est pour cette raison que le trajet des fibres extraites est parfois interrompu). Les méthodes probabilistes sont robustes au bruit et aux effets de volume partiel (ce type de méthode permet notamment de traverser un croisement de fibres).

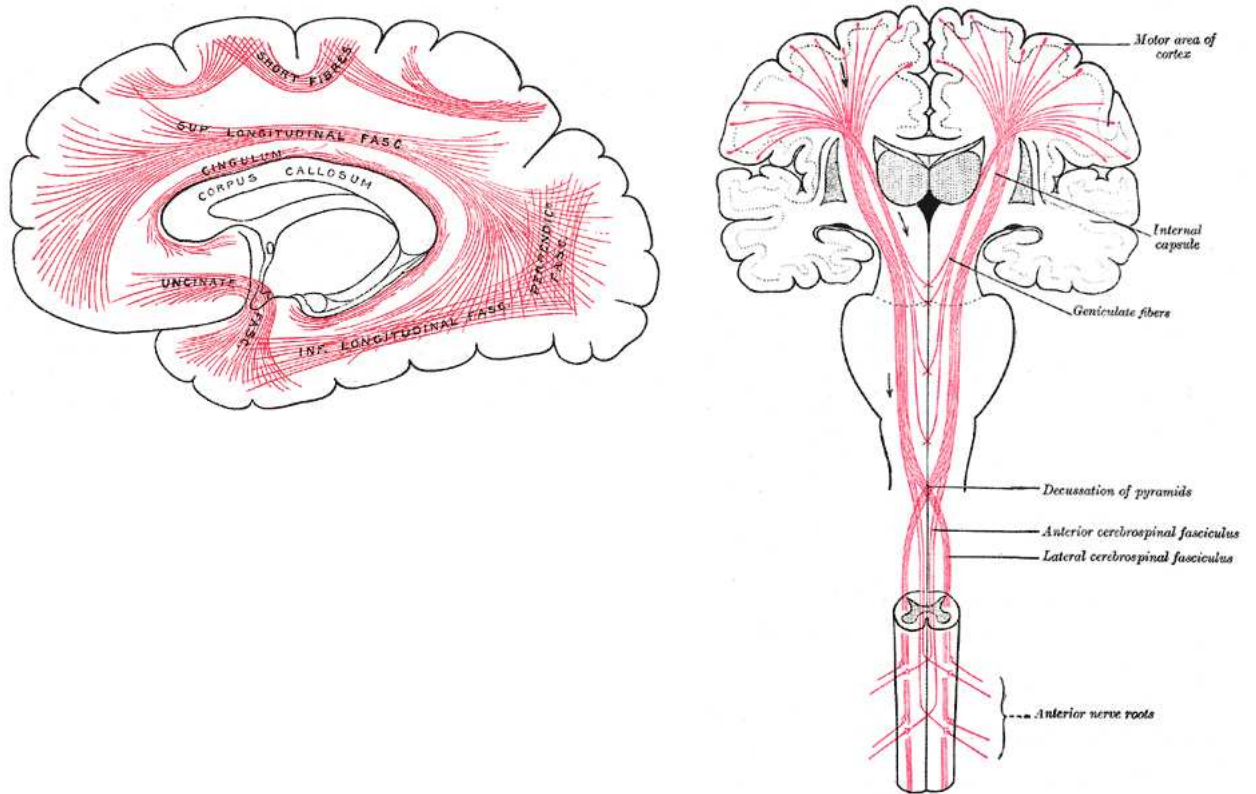


FIG. 8.4 – Principaux faisceaux de la substance blanche (source : wikipedia).

De plus, un avantage certain de ces méthodes est de donner directement une probabilité de connexion entre différentes régions du cerveau. Pour ces méthodes, l'inconvénient majeur concerne le temps de calcul qui peut être très important.

Dans le cadre de ce travail de thèse, une tractographie déterministe par streamline basée sur un modèle d'ordre 2 est utilisée [3]. L'algorithme disponible dans Slicer¹ est utilisé. Cette méthode permet de ne pas ralentir excessivement la chaîne de traitement et permet de reconstruire la plupart des faisceaux, ce qui est suffisant dans le contexte de notre application. Plusieurs contraintes sont à respecter lors de la construction d'une fibre de la substance blanche. Tout d'abord, les fibres doivent avoir leurs extrémités situées dans la substance grise (cette contrainte n'est souvent pas respectée) et effectuer un parcours dans la substance blanche. De plus, une fibre ne peut pas traverser le liquide céphalo-rachidien.

L'algorithme de tractographie est initialisé avec un champ de tenseurs de diffusion préalablement calculés en chaque voxel de l'image et une liste de points germes. A partir de chaque point germe, une croissance de la fibre dans la direction la plus probable du modèle local $\vec{e}_1$ est initiée. Ce processus est répété de proche en proche. Les voxels identifiés sont réunis par une courbe qui matérialise le modèle de fibre. Les conditions d'arrêt sont déterminées par rapport à la valeur de FA et à l'angle formé entre trois points consécutifs. Le seuil sur la

¹<http://www.slicer.org/>

FA garantit au modèle de fibre de rester dans la substance blanche. Le seuil sur la courbure est très dépendant du pas choisi pour la propagation et permet de générer des modèles de fibres réalistes (ceci permet d'éviter que le modèle de fibre ne se retourne sur lui même par exemple). Généralement, aucune contrainte n'est imposée sur la position des extrémités du modèle de fibre. La tractographie *streamline* peut être calculée avec un modèle d'ordre 2 ou d'ordre 4 (cf. figure 8.5). Lorsqu'un modèle d'ordre 4 est utilisé, il faut ajouter une étape permettant de trouver la direction la plus probable du modèle local (classiquement la direction courante qui maximise la valeur absolue du produit scalaire avec la direction précédente).

Dans le cadre du suivi longitudinal, nous proposons d'utiliser l'information contenue dans le modèle mathématique des fibres de la substance blanche (une information anatomique) comme un *a priori* lors de la génération d'une population de tenseurs. Introduire une telle information permet de respecter l'hypothèse du modèle constant par morceaux, mais également d'obtenir des informations sur le retentissement de certaines pathologies sur l'organisation de la substance blanche.

8.3.2 Regroupement des fibres en faisceaux

Le résultat de la tractographie est un ensemble de courbes tridimensionnelles donnant une idée de l'organisation des fibres de la substance blanche. Ici, l'idée est de regrouper de manière automatique les fibres en faisceaux. Dans la littérature, de nombreuses techniques de classification ont été proposées pour résoudre ce problème [31, 14, 11, 10, 6, 23, 24, 19, 20]. Nous allons présenter deux de ces stratégies. La méthode hiérarchique proposée dans [31] est présentée dans un premier temps. Cette stratégie donne de bons résultats pour un temps de calcul relativement élevé. Cette étape ralentit significativement la vitesse de notre chaîne de traitement (cf. section 9.2). En s'inspirant de cette technique, avec pour objectif toutefois d'accélérer le procédé, Garyfallidis [14] propose une méthode d'agglomération plus grossière qui sera utilisée dans notre chaîne de traitement (cf. section 9.2).

Méthode hiérarchique

La méthode proposée dans [31] est explicitée dans cette partie. Elle repose sur la définition d'une distance entre deux fibres Q et R . Supposons que la fibre Q est échantillonnée avec M points. Cette distance est alors calculée en faisant la moyenne de toutes les distances d_i ($i \in [1, M]$) entre les points des deux fibres les plus proches, chaque distance d_i étant conservée si et seulement si $d_i > t$ (cf. figure 8.6) :

$$d_t(Q, R, t) = \text{moyenne}_{a \in Q, \|a-b\| > t} \min_{b \in R} \|a - b\| \quad (8.7)$$

où t est un seuil. Par exemple, ce seuil permet de faire la distinction entre deux fibres qui ont un trajet proche pendant longtemps et qui divergent fortement à la fin de leurs trajets. Il est également intéressant de considérer une métrique symétrique. Pour la classification, le maximum de la métrique directe ($Q \rightarrow R$) et de la métrique indirecte ($R \rightarrow Q$) (cf. figure 8.6) est conservé et permet de distinguer au mieux deux faisceaux :

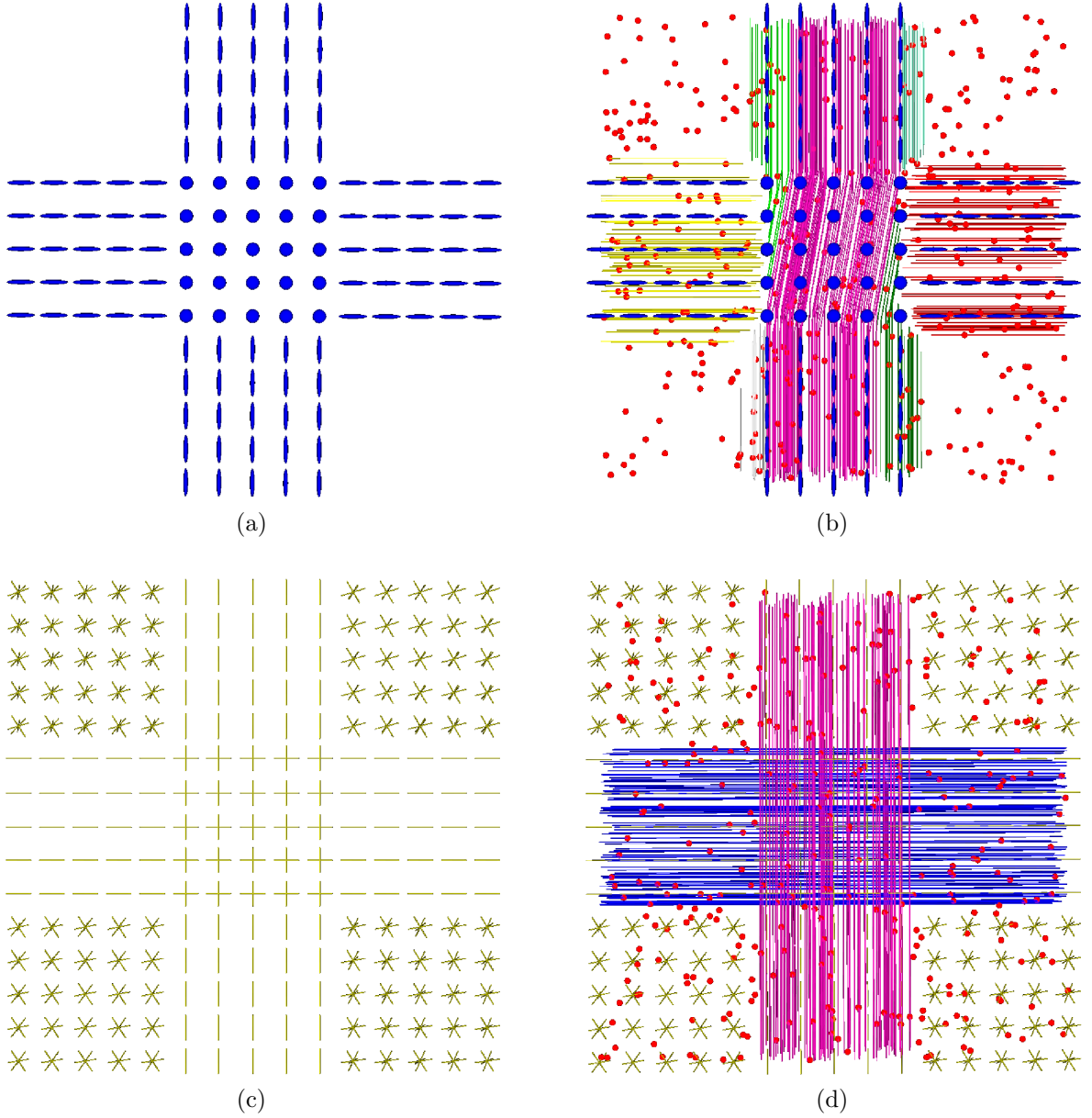


FIG. 8.5 – Tractographie avec un modèle d'ordre 2 (a) et (b) et tractographie avec un modèle d'ordre 4 (c) et (d). Les points rouges sur les figures (b) et (d) représentent les points germes à partir desquels un modèle de fibre a été calculé. Les fibres obtenues en figures b- et d- correspondant à une même structure ont été colorées de la même couleur. Dans la figure (b), les tenseurs centraux (issus d'un signal modélisant un croisement de fibres) sont planaires, avec une FA suffisante pour prolonger la tractographie.

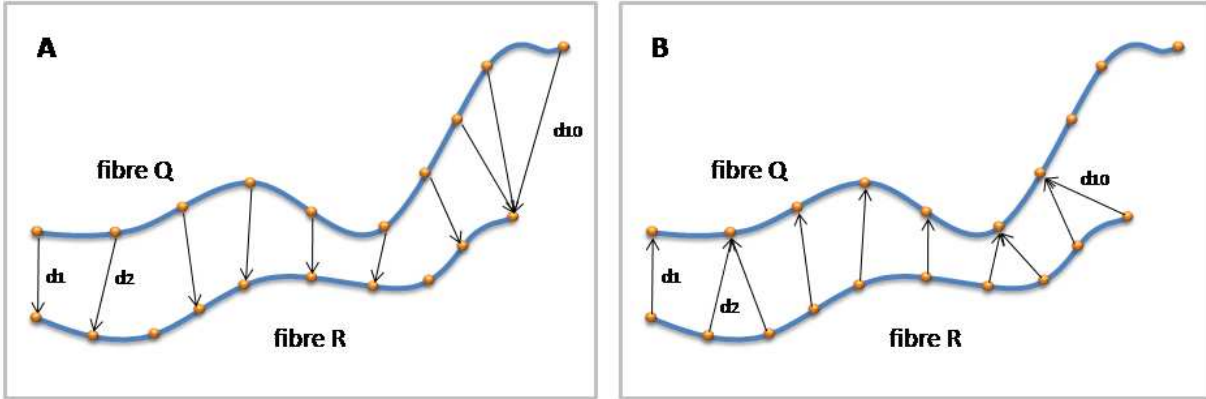


FIG. 8.6 – Mesure de similarité entre deux courbes Q et R : A- illustration de la métrique directe ($Q \rightarrow R$) et B- illustration de la métrique indirecte ($R \rightarrow Q$).

$$d_{Lt}(Q, R, t) = \max(d_t(Q, R, t), d_t(R, Q, t)) \quad (8.8)$$

La métrique ainsi définie permet d'identifier et de regrouper des fibres cohérentes appartenant à un même faisceau [31].

Cette métrique est utilisée comme critère de similarité dans un algorithme de classification automatique des fibres de la substance blanche. Pour ce faire, une méthode hiérarchique est utilisée. On appellera un groupe (*cluster* en anglais) un ensemble de une ou plusieurs fibres regroupées par la méthode de classification. Au début de la classification, il y a autant de groupes que de fibres. Ensuite, l'algorithme se déroule de la manière suivante :

- s'il ne reste plus qu'un groupe, arrêter la procédure.
- calculer le critère de similarité pour toutes les paires des différents groupes.
- fusionner les deux groupes montrant la plus grande ressemblance (*i.e.*, qui ont la métrique de similarité la plus faible).

En spécifiant un seuil d_f , la classification peut être arrêtée lorsque le critère de similarité dépasse ce seuil. Le nombre de groupes formés varie inversement avec le seuil d_f . Il existe différentes méthodes hiérarchiques qui diffèrent par leurs façons de mesurer la similarité entre deux groupes :

- **La méthode du plus proche voisin, ou en anglais, *single linkage*** : soient deux groupes G_1 et G_2 et d_{Lt} la métrique de l'équation (8.8). Soit T_G^i la i -ème fibre du groupe G . On calcule la similarité de G_1 avec G_2 , notée $D(G_1, G_2)$, en considérant la plus petite valeur de similarité pour toutes les paires possibles entre les deux groupes, soit :

$$D(G_1, G_2) = \min [d_{Lt}(T_{G_1}^i, T_{G_2}^j, t), i \in \{1, \dots, \#\{G_1\}\} \text{ et } j \in \{1, \dots, \#\{G_2\}\}] \quad (8.9)$$

- **La méthode du voisin le plus éloigné, ou en anglais, *complete linkage*** : toujours avec les mêmes notations, on calcule la similarité de G_1 avec G_2 en considérant la plus grande valeur de similarité pour toutes les paires possibles entre les deux groupes, soit :

$$D(G_1, G_2) = \max [d_{Lt}(T_{G_1}^i, T_{G_2}^j, t), i \in \{1, \dots, \#\{G_1\}\} \text{ et } j \in \{1, \dots, \#\{G_2\}\}] \quad (8.10)$$

- **La méthode de la moyenne, ou en anglais, *average*** : toujours avec les mêmes notations, on calcule la similarité de G_1 avec G_2 en considérant la similarité moyenne entre groupes, soit :

$$D(G_1, G_2) = \frac{1}{\#\{G_1\}\#\{G_2\}} \sum_{i=1}^{\#\{G_1\}} \sum_{j=1}^{\#\{G_2\}} d_{Lt}(T_{G_1}^i, T_{G_2}^j, t) \quad (8.11)$$

Méthode d'agglomération

Le problème avec les méthodes hiérarchiques de classification est le temps de calcul nécessaire pour mettre à jour la matrice de proximité contenant les distances entre les paires des différents groupes. La majeure partie de ce calcul intensif provient de la comparaison géométrique fine mise en œuvre pour chaque paire de fibres. En gardant la même philosophie, avec pour objectif toutefois d'accélérer le procédé, Garyfallidis [14] propose une classification grossière pouvant être éventuellement raffinée avec la méthode hiérarchique présentée précédemment. Pour traiter nos données, nous utiliserons cette approche qui est très efficace en temps de calcul. Par exemple, une tractographie constituée de 11 737 fibres est classifiée en 237 faisceaux en seulement 20,48 secondes (un CPU de 2,67 GHz). Elle est basée sur les simplifications suivantes :

- chaque fibre est représentée par seulement trois points. Ces trois points sont choisis pour créer un modèle de fibre constitué de deux segments de même longueur attachés aux extrémités de la fibre d'origine.
- chaque groupe est représenté par une fibre moyenne composée de trois points.

Avec ces simplifications, le critère de similarité direct entre deux courbes Q_3 et R_3 (qui correspondent respectivement aux versions sous-échantillonnées en trois points de Q et R) s'écrit :

$$d_{direct} = \frac{1}{3} \sum_{i=1}^3 ||Q_3[i] - R_3[i]|| \quad (8.12)$$

Il faut également considérer la distance miroir pour permettre de tester les deux associations possibles, soit :

$$d_{miroir} = \frac{1}{3} \sum_{i=1}^3 ||Q_3[i] - R_3[4-i]|| \quad (8.13)$$

Soit un ensemble de fibres moyennes composées de trois points $A_{3,i}, i = (1, \dots, n)$, chacune d'entre elles représentant un *cluster*. Soit T_3 la fibre composée de trois points à classifier. Pour un *cluster* i donné, le minimum des distances $d_{direct}(T_3, A_{3,i})$ et $d_{miroir}(T_3, A_{3,i})$, noté d_i , définit le degré de similarité du *cluster* i avec la fibre T_3 . Ainsi, la procédure de fusion est donnée par :

- si la distance $d_{min} = \min_i(d_i)$, $i = (1, \dots, n)$ est plus petite qu'un seuil de fusion d_f , la fibre T_3 est ajoutée au *cluster* correspondant (le plus proche), et la fibre moyenne du *cluster* est mise à jour.
- sinon, un nouveau *cluster* est créé.

Le nombre de groupes formés varie inversement avec le seuil d_f . Une illustration de cette méthode d'agglomération des fibres en faisceaux est présentée en figure 8.7.

8.3.3 Estimation des paramètres dans une section de faisceau

Grâce à la méthode d'agglomération, des paquets de fibres représentatifs de faisceaux de la substance blanche ont été construits. L'objectif ici est d'extraire des populations de tenseurs caractéristiques de ces faisceaux. Pour cela, une fibre représentative pour chaque faisceau est tout d'abord extraite (cette fibre sera également rééchantillonnée régulièrement). Puis, les populations de tenseurs sont extraites à partir des tenseurs observés dans des plans transversaux des fibres représentatives.

Sélection d'une fibre représentative

La fibre moyenne composée de trois points associée à chaque *cluster* n'est pas utilisée pour représenter le faisceau. En effet, cette fibre offre une représentation géométrique trop grossière de l'organisation anatomique sous-jacente. Afin d'obtenir une description géométrique plus riche d'un faisceau composé de fibres similaires, une d'entre elles est sélectionnée. Soient $T_{i,j}$ les fibres du faisceau i et n_i le nombre de fibres dans ce faisceau : $j \in \{1, \dots, n_i\}$. La fibre de référence ou centroïde du faisceau i , notée $T_{i,ref}$, est définie comme étant la fibre du faisceau minimisant la distance avec toutes les autres fibres du faisceau. Pour créer une métrique symétrique, la moyenne des deux distances est calculée :

$$d_{MC}(Q, R) = \frac{d_t(Q, R, 0) + d_t(R, Q, 0)}{2} \quad (8.14)$$

Le paramètre t est fixé à zéro puisque le faisceau est formé par définition avec des fibres cohérentes. Le critère de sélection du centroïde peut alors se résumer de la manière suivante :

$$T_{i,ref} = \arg \min_{T_{i,k}, k \in \{1, \dots, n_i\}} \left[\sum_{l \in \{1, \dots, n_i\}} d_{MC}(T_{i,k}, T_{i,l}) \right] \quad (8.15)$$

Une illustration de ce procédé est présentée en figure 8.8. Pour lisser les courbes et/ou avoir un échantillonnage constant, chaque centroïde $T_{i,ref}$ peut être rééchantillonné avec A points

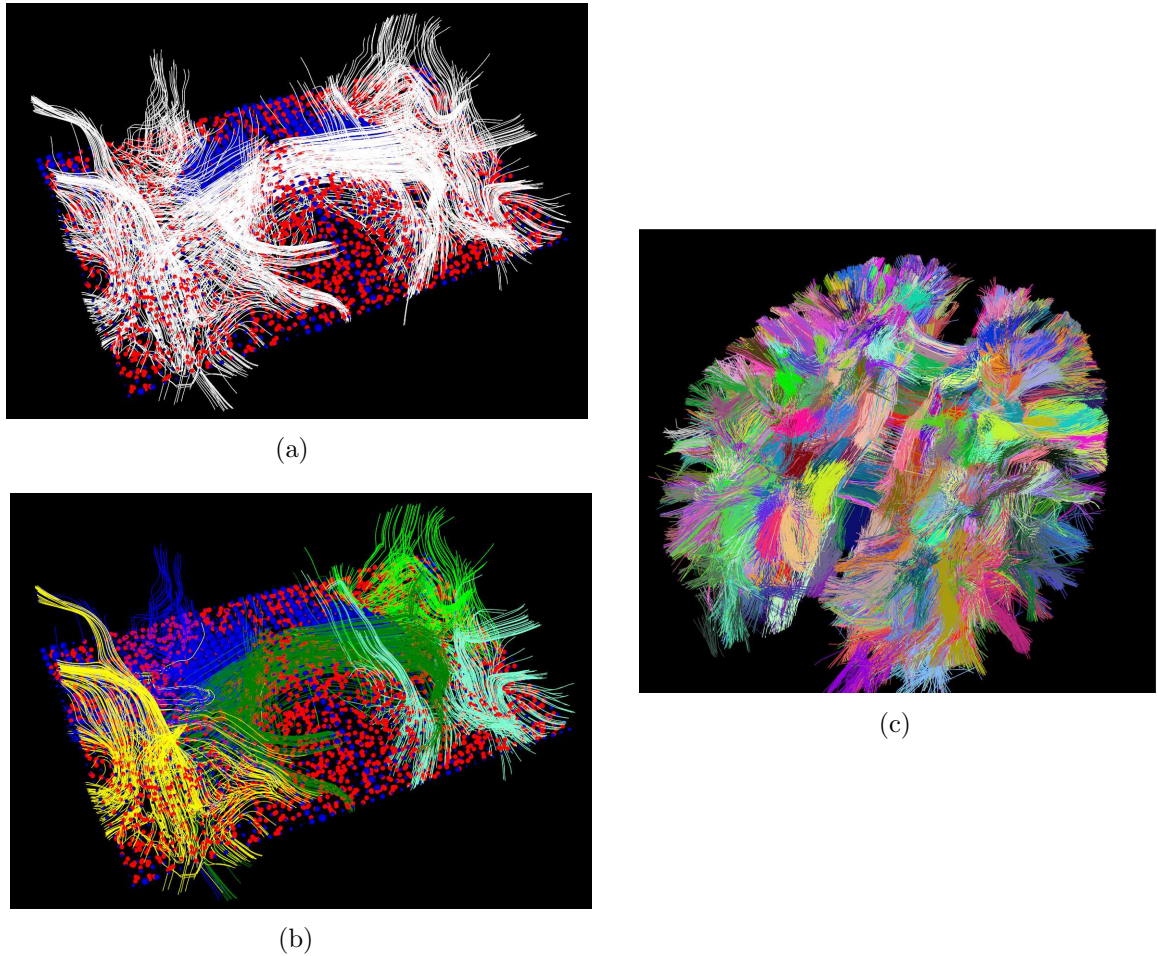


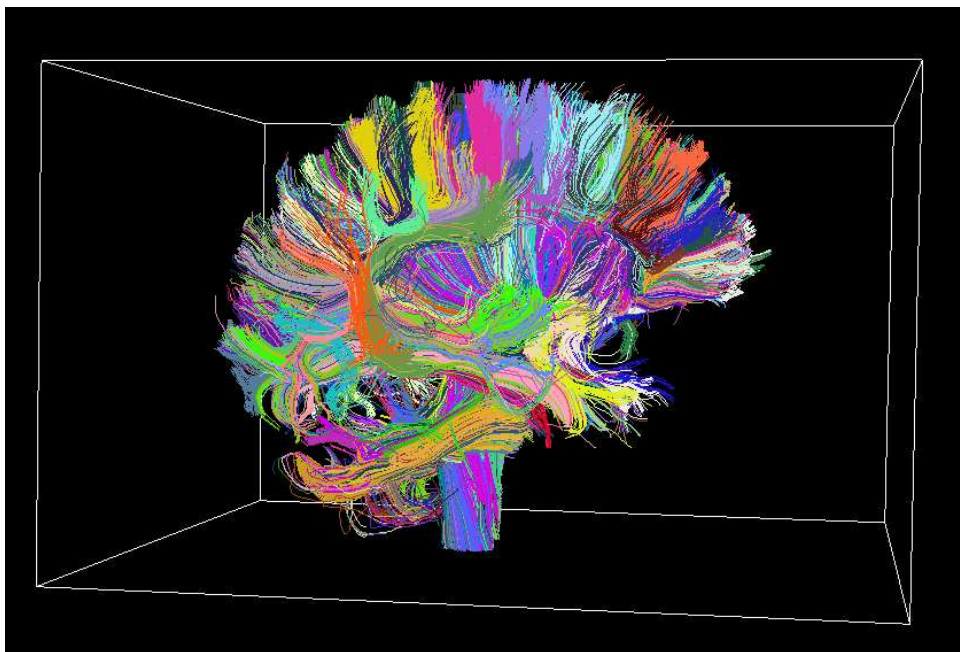
FIG. 8.7 – (a) tractographie avec un modèle de tenseur d'ordre 2 dans une région d'intérêt proche du corps calleux postérieur, (b) résultat obtenu avec la méthode d'agglomération sur cette région, et (c) résultat obtenu avec la méthode d'agglomération sur la totalité du cerveau : $t = 0,5 \text{ mm}$ et $d_f = 5 \text{ mm}$ (la méthode hiérarchique raffine légèrement ce partitionnement, mais globalement de manière peu significative).

en utilisant des splines cardinales implémentées dans VTK² pour interpoler les courbes. Pour l'élaboration des différentes chaînes de traitement, la fibre représentative $T_{i,ref}$ et la fibre rééchantillonnée associée $\tilde{T}_{i,ref}^A$ sont stockées.

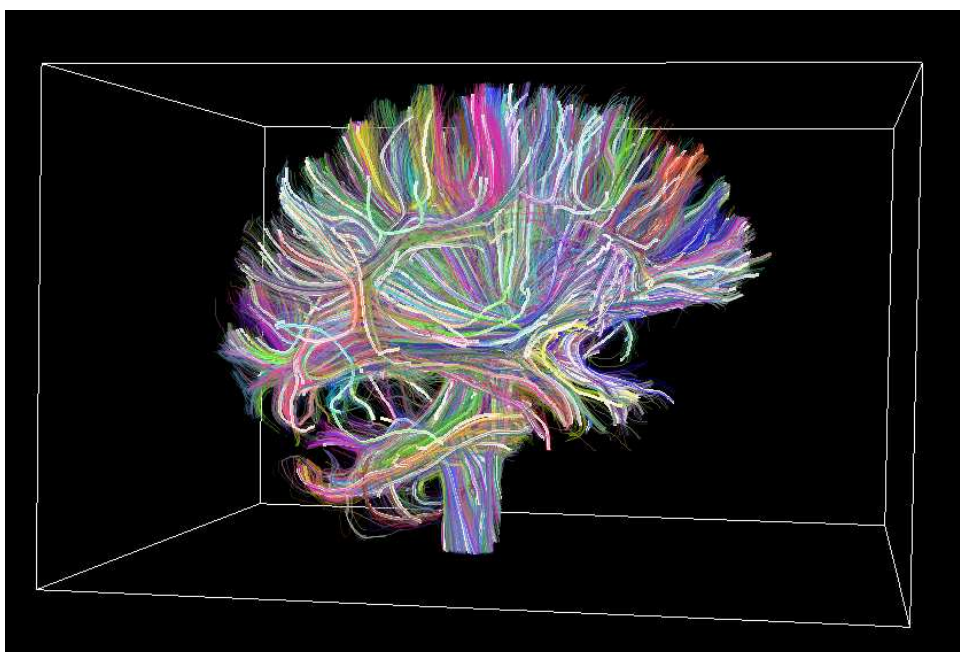
Extraction d'une population basée sur l'organisation de la substance blanche

La figure 8.9-a propose une illustration de la méthode de construction du voisinage basée sur l'organisation de la substance blanche décrite dans cette section. Soit P_k la section transversale définie comme le plan perpendiculaire à la fibre représentative $T_{i,ref}$ au point d'échantillonnage l_k . A chaque point d'échantillonnage de la fibre de référence $T_{i,ref}$, les 2τ

²[http ://www.vtk.org/](http://www.vtk.org/)



(a)



(b)

FIG. 8.8 – (a) résultat du clustering direct et (b) résultat du calcul des représentants de chaque groupe (fibres blanches plus larges).

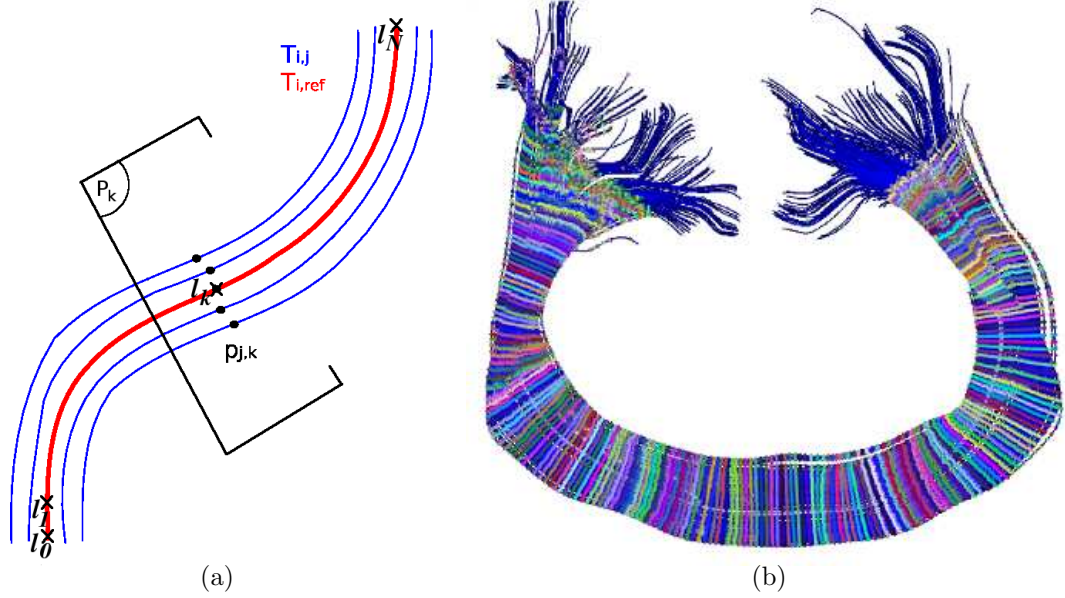


FIG. 8.9 – (a) représentation schématique de la procédure d'extraction d'une population basée sur l'organisation de la substance blanche et (b) résultat de cette méthode d'extraction sur un faisceau (une couleur aléatoire a été associée pour chaque section de faisceau).

sections transversales symétriques (P_{k-l}, P_{k+l}) , $l \in (1, \dots, \tau)$ sont considérées pour le calcul de la population. τ est un paramètre de décalage, fixé par l'utilisateur, qui permet de contrôler l'extension spatiale du voisinage. Toutes les intersections $p_{j,l}$, $l \in (k - \tau, \dots, k + \tau)$ entre les sections transversales P_l et les fibres $T_{i,j}$ du faisceau i sont calculées. Dans certains cas, les sections transversales peuvent couper une fibre du faisceau plusieurs fois (par exemple pour une fibre en « U »). Pour éviter cela, seule l'intersection la plus proche de la fibre de référence est considérée.

A partir des points $p_{j,l}$, une population de N tenseurs, représentative de la variabilité locale, est extraite (N est fixé pour garantir que les statistiques ont les mêmes degrés de liberté et ainsi peuvent être comparées directement). Pour les mêmes raisons, le calcul de la statistique se fait uniquement si assez de tenseurs sont disponibles dans la population. La méthode la plus directe pour construire cette population consiste à interpoler tous les tenseurs aux points $p_{j,l}$ et d'en sélectionner N . Néanmoins, l'hypothèse d'indépendance des tenseurs de la population (imposée par le test statistique) risque de ne pas être vérifiée puisque plusieurs $p_{j,l}$ peuvent appartenir à un même voxel. Ainsi, la méthode proposée consiste à sélectionner N tenseurs à partir des voxels contenant au moins un $p_{j,l}$ en appliquant les règles suivantes :

- les intersections avec la section transversale P_k sont considérées en premier lieu.
- ensuite, les intersections avec les sections transversales symétriques (P_{k-l}, P_{k+l}) les plus proches (l croissant) sont considérées.
- dans les deux cas précédent, les points les plus proches de l_k (par rapport à la distance euclidienne) sont considérés en premier.

Dans ce chapitre nous venons de présenter trois méthodes d'extraction de populations de tenseurs. Les stratégies de prise en compte du voisinage spatial à l'échelle du voxel et de rééchantillonnage bootstrap (local et spatial) permettront la détection de changement de diffusion à l'échelle du voxel. La prise en compte d'un voisinage basé sur une modélisation des faisceaux de la substance blanche permettra la détection de changements de diffusion sur une modélisation des faisceaux de fibres de la substance blanche. Ces faisceaux proviennent de fibres issues d'un algorithme de tractographie déterministe sur le modèle du tenseur d'ordre 2 suivi par une classification de chaque fibre grâce à la méthode d'agglomération. L'extraction des populations à comparer se fait alors le long de chaque faisceau en considérant les plans transversaux successifs d'une fibre représentative. Les deux chaînes de traitement correspondantes sont présentées dans le chapitre suivant.

Les deux chaînes de traitement proposées pour le suivi longitudinal

Sommaire

9.1	Détection de changement de diffusion à l'échelle du voxel . . .	146
9.2	Détection de changements de diffusion sur une modélisation des faisceaux de fibres	147
9.2.1	Statistique locale	147
9.2.2	Statistique globale	148

DANS cette partie, nous décrivons deux chaînes de traitement permettant le suivi longitudinal de différentes pathologies neurodégénératives (telles que la sclérose en plaques et la neuromyéélite optique) en ITD. La diffusion des molécules d'eau est modélisée par des tenseurs d'ordre 2. Les points clés dans ces chaînes de traitement sont la manière d'échantillonner une population de tenseurs à comparer en chaque voxel et le test statistique associé. Deux chaînes de traitement sont envisagées dans ce chapitre : une détection de changement de diffusion à l'échelle du voxel et une détection de changements de diffusion sur une modélisation des faisceaux de fibres.

L'analyse des résultats de ces stratégies permettra de répondre à un certain nombre d'interrogations : tout d'abord faut-il faire une statistique locale, locale intégrant un voisinage spatial ou essayer de faire une statistique globale sur une région prédéfinie ? Une statistique locale risque d'être sensible aux artefacts et bruit d'acquisition. Ainsi, la stratégie la plus communément utilisée dans la littérature consiste à prendre en compte un voisinage spatial de taille fixée par l'utilisateur. Cependant, le changement ne concernant qu'un ou peu de voxels seront difficilement détectables par une telle approche. De tels changements sont souvent immédiatement rejetés par les neurologues. En intégrant un voisinage spatial, la qualité de la détection est améliorée (au moins visuellement). Mais d'un point de vue quantitatif, cette idée est-elle réellement fondée ? Ne peut-on pas mettre à défaut la méthode intégrant un voisinage spatial, au moins dans certains cas ? La perspective d'une statistique globale

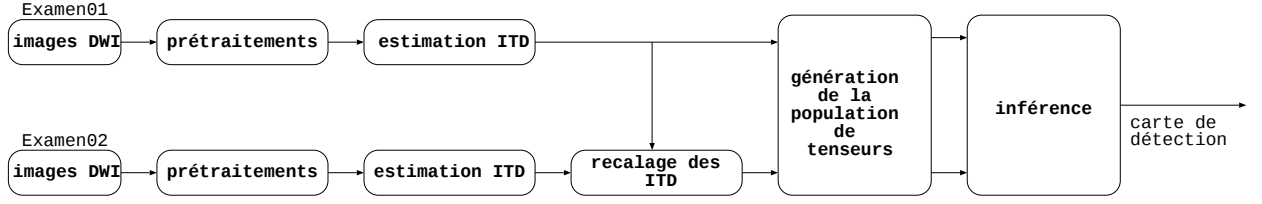


FIG. 9.1 – Etapes à suivre pour effectuer un suivi longitudinal lors d’une détection à l’échelle du voxel.

semble intéressante, mais comment définir une région d’intérêt pour effectuer cette statistique ? Nous avons vu qu’il était également possible de comparer directement les tenseurs entre eux (et donc de les traiter comme des observations) ou d’utiliser l’information contenue dans le signal de diffusion pour construire la statistique. En augmentant la quantité d’information exploitée, pourrait-on trouver d’autres changements, ou raffiner les changements déjà détectés ?

9.1 Détection de changement de diffusion à l’échelle du voxel

La détection de changement de diffusion à l’échelle du voxel suit le schéma général présenté en figure 9.1. En entrée, deux séquences de diffusion d’un même patient sont dans un premier temps prétraitées afin de corriger les artefacts causés par l’acquisition (avec les méthodes présentées dans la section 7.1). Puisque les cartes de champ B_0 permettant la correction des effets de susceptibilité magnétique n’étaient pas disponibles, seuls les courants de Foucault ont été corrigés avec FSL¹. Ensuite, les ITD sont estimées en utilisant la méthode classique des moindres carrés. Pour garantir la positivité des tenseurs de diffusion (et utiliser la métrique log-euclidienne), les valeurs propres négatives ou nulles sont remplacées par une valeur positive arbitrairement petite. Un recalage non-rigide permet de recalculer l’ITD calculée à partir de l’examen #2 sur l’ITD de l’examen #1 (cf. sections 7.2 et 7.3). Ensuite, différentes méthodes permettent de générer les populations de tenseurs à comparer : prise en compte du voisinage à l’échelle du voxel (cf. section 8.1) ou deux méthodes de rééchantillonnage bootstrap (cf. section 8.2 : bootstrap local ou spatial). Finalement, le test statistique sur les valeurs propres du tenseur de diffusion, noté T_{e-val} , est calculé afin de mettre en évidence des changements entre deux images de diffusion (cf. section 6.2.2) :

$$T_{e-val} = \frac{(n-2)n_1n_2}{3n^2\hat{\sigma}^2} \text{tr} [(\bar{\mathbf{A}}_1 - \bar{\mathbf{A}}_2)^2] \sim F(3, 6(n-2)) \quad (9.1)$$

où n_1, n_2 sont le nombre de tenseurs dans la première et deuxième population, $n = n_1 + n_2$ et $\bar{\mathbf{A}}_1, \bar{\mathbf{A}}_2$ sont les valeurs propres des matrices moyennes des deux populations de log-tenseurs. L’estimation des moyennes $\bar{\mathbf{L}}_1$ et $\bar{\mathbf{L}}_2$ de chaque population de tenseurs s’écrit (cf. section 6.2.2) :

¹<http://www.fmrib.ox.ac.uk/fsl/>

Suivi longitudinal des tenseurs d'ordre 2 : prétraitements et outils statistiques

Les deux chaînes de traitement proposées pour le suivi longitudinal

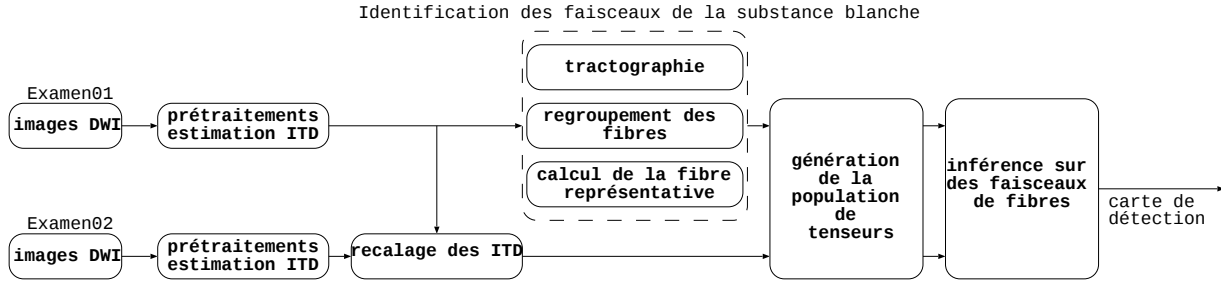


FIG. 9.2 – Etapes à suivre pour effectuer un suivi longitudinal lors d’une détection basée sur une modélisation des faisceaux de fibres.

$$\bar{\mathbf{L}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \text{Log}(\mathbf{D}_{i,j}) \quad (9.2)$$

et l’estimation de la variance s’écrit (cf. section 6.2.2) :

$$\hat{\sigma}^2 = \frac{1}{6n} \left[\sum_{j=1}^{n_1} \text{tr}(\mathbf{L}_{1,j} - \bar{\mathbf{L}}_1)^2 + \sum_{j=1}^{n_2} \text{tr}(\mathbf{L}_{2,j} - \bar{\mathbf{L}}_2)^2 \right] \quad (9.3)$$

La carte de détection contient les p-valeurs corrigées avec la stratégie proposée en section 6.2.3. La chaîne de traitement met en œuvre un test sur les matrices de valeurs propres. Ce dernier permet de quantifier des changements de diffusion, sans se préoccuper de l’orientation des tenseurs. Ce test est très adapté aux pathologies neurodégénératives étudiées dans la partie suivante où une démyélinisation est attendue. Toutefois, cette chaîne de traitement reste générique et peut être utilisée pour d’autres pathologies en intégrant les tests sur la structure matricielle ou les vecteurs propres présentés en section 6.2.2. Cette intégration est immédiate.

9.2 Détection de changements de diffusion sur une modélisation des faisceaux de fibres

La détection de changement de diffusion basée sur une modélisation des faisceaux de fibres suit le schéma général présenté en figure 9.2. La procédure est identique à celle de la chaîne de traitements proposée dans la section précédente (section 9.1) jusqu’à l’étape de recalage. Ensuite, une statistique locale (autour de chaque échantillon des fibres représentatives $T_{i,ref}$) et une statistique globale (intégrée sur chaque fibre représentative rééchantillonnée avec A points $\tilde{T}_{i,ref}^A$) sont envisagées.

9.2.1 Statistique locale

Pour générer les populations de tenseurs à comparer, une modélisation des faisceaux de fibres de la substance blanche est intégrée dans la chaîne de traitements (cf. section 8.3). La

statistique locale est obtenue en comparant des échantillons de tenseurs dans des sections transversales aux faisceaux. Ensuite, puisque des changements de diffusion sont attendus, l'inférence sur les valeurs propres du tenseur de diffusion T_{e-val} est de nouveau calculée, mais cette fois-ci en chaque échantillon des fibres représentatives $T_{i,ref}$. Utiliser $T_{i,ref}$ lors du calcul du test statistique permet d'avoir un espacement régulier entre les échantillons. Cette propriété est intéressante, elle permet notamment de comparer équitablement d'un point de vue statistique deux faisceaux de taille très différente. La carte de détection contient les p-valeurs corrigées avec la stratégie proposée en section 6.2.3.

9.2.2 Statistique globale

Une statistique globale permet de connaître la/les fibres les plus touchées par la pathologie étudiée. Ainsi, ce test permet de construire une liste de faisceaux ordonnée par probabilité d'évolution. Aux A points d'échantillons de chaque fibre représentative $\tilde{T}_{i,ref}^A$, un log-tenseur moyen est calculé en considérant l'intersection des fibres avec la section transversale au faisceau au point courant dans chacun des deux examens recalés. Ensuite, les différences entre les log-tenseurs moyens associés (dans leurs représentations canoniques) sont calculées afin de réaliser un test apparié. Il est important de considérer les log-tenseurs moyens dans leurs représentations canoniques (*i.e.*, les matrices de valeurs propres issues de la décomposition spectrale) car les tenseurs le long d'un faisceau n'ont pas nécessairement la même orientation. Ainsi, une population de A différences de log-tenseurs est obtenue pour chaque faisceau. On veut tester si ces A différences de log-tenseurs ont une matrice de valeurs propres nulle (hypothèse H_0). Le test apparié T_{global} s'écrit alors (voir section 6.2.1) :

$$T_{global} = \frac{A-1}{3\hat{\sigma}^2} tr [(\bar{\mathbf{L}} - \mathbf{\Lambda}_0)^2] \sim F(3, 6(A-1)) \quad (9.4)$$

où $\bar{\mathbf{L}}$ représente les valeurs propres de la matrice moyenne de la population $\bar{\mathbf{L}}$ et $\mathbf{\Lambda}_0 = \mathbf{0}_3$. L'estimation de la moyenne $\bar{\mathbf{L}}$ pour chaque faisceau s'écrit (cf. section 6.2.1) :

$$\bar{\mathbf{L}} = \frac{1}{A} \sum_{i=1}^A Log(\mathbf{D}_i) \quad (9.5)$$

et l'estimation de la variance s'écrit (cf. section 6.2.1) :

$$\hat{\sigma}^2 = \frac{1}{6A} \left[\sum_{i=1}^A tr (Log(\mathbf{D}_i) - \bar{\mathbf{L}})^2 \right] \quad (9.6)$$

Utiliser $\tilde{T}_{i,ref}^A$ lors du calcul de la statistique globale permet d'avoir le même nombre de degrés de liberté pour chaque faisceau. Il est alors aisé de comparer les faisceaux entre eux. La carte de détection contient les p-valeurs corrigées avec la stratégie proposée en section 6.2.3

Conclusion partielle

Pour détecter automatiquement des changements dans des séquences temporelles d'images de diffusion, nous venons de proposer différentes stratégies pour construire des échantillons de tenseurs reflétant la variabilité existante en chaque voxel. Ces échantillons sont générés en exploitant soit le voisinage spatial, soit le signal de diffusion (bootstrap local), soit une combinaison des deux (bootstrap spatial). Un voisinage adaptatif défini sur les faisceaux de fibres de la substance blanche est également proposé.

Ces stratégies ont été intégrées dans deux chaînes de traitement permettant de détecter des modifications des propriétés de diffusion des tenseurs d'ordre 2 (*i.e.*, des modifications de leurs valeurs propres). Pour mesurer ces changements, un test statistique multivarié basé sur le rapport de vraisemblance à deux populations permet de comparer localement les images de tenseurs de diffusion, et un test statistique multivarié basé sur le rapport de vraisemblance à une population permet de comparer deux régions de tenseurs de diffusion appariées. Ces chaînes de traitements s'appuient sur des prétraitements communs pour corriger les distorsions géométriques des images de diffusion, recalcr les deux examens à comparer, et calculer des p-valeurs corrigées pour tenir compte des comparaisons multiples. La partie suivante présente les résultats expérimentaux obtenus avec ces deux chaînes de traitement.

Bibliographie

- [1] D. ALEXANDER, C. PIERPAOLI, P. BASSER et J. GEE : Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Transaction on Medical Imaging*, 20(11):1131 – 1139, 2001.
- [2] V. ARSIGNY, P. FILLARD, X. PENNEC et N. AYACHE : Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328 – 347, 2007.
- [3] P. BASSER, S. PAJEVIC, C. PIERPAOLI, J. DUDA et A. ALDROUBI : In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44(4):625 – 632, 2000.
- [4] T. E. J. BEHRENS, H. J. BERG, S. JBABDI, M. S. RUSHWORTH et M. W. WOOLRICH : Probabilistic diffusion tractography with multiple fibre orientations : What can we gain ? *Neuroimage*, 34(1):144 – 155, 2007.
- [5] M. BOSC, F. HEITZ, J. ARMSPACH, I. NAMER, D. GOUNOT et L. RUMBACH : Automatic change detection in multimodal serial MRI : application to multiple sclerosis lesion evolution. *NeuroImage*, 20(2):643 – 656, 2003.
- [6] A. BRUN, H. KNUTSSON, H. PARK, M. SHENTON et C. WESTIN : Clustering fiber traces using normalized cuts. In *Medical Image Computing and Computer Assisted Intervention, Saint-Malo*, vol. 3216 de *LNCS*, p. 368 – 375, 2004.
- [7] J. CAMPBELL, K. SIDDIQI, V. RYMAR, A. SADIKOT et G. PIKE : Flow-based fiber tracking with diffusion tensor and q-ball data : Validation and comparison to principal diffusion direction techniques. *Neuroimage*, 27(4):725 – 736, 2005.
- [8] M. CHERNICK : *Bootstrap methods : a guide for practitioners and researchers*. Wiley-Interscience, 2007.
- [9] S. CHUNG, Y. LU et R. HENRY : Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. *NeuroImage*, 33(2):531 – 541, 2006.
- [10] I. COROUGE, S. GOUTTARD et G. GERIG : Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *IEEE International Symposium on Biomedical Imaging, Arlington*, ISBI '04, p. 344 – 347, 2004.
- [11] Z. DING, J. GORE et A. ANDERSON : Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49(4):716 – 721, 2003.

- [12] D. FREEDMAN : Bootstrapping regression-models. *Annals of Statistics*, 9(6):1218 – 1228, 1981.
- [13] O. FRIMAN, G. FARNEBAECK et C. WESTIN : A Bayesian approach for stochastic white matter tractography. *IEEE Transactions on Medical Imaging*, 25(8):965 – 978, 2006.
- [14] E. GARYFALLIDIS, M. BRETT et I. NIMMO-SMITH : Fast dimensionality reduction for brain tractography clustering. 16th Annual Meeting of the Organization for Human Brain Mapping, 2010.
- [15] C. GRAMKOW : On averaging rotations. *International Journal of Computer Vision*, 42(1-2):7 – 16, 2001.
- [16] A. GRIGIS, V. NOBLET, F. HEITZ et J. ARMSPACH : On the integration of spatial neighborhood information for detecting longitudinal changes in mri sequences. In *LIVIM*, 2011.
- [17] D. JONES et M. CERCIGNANI : Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR in Biomedicine*, 23(7):803 – 820, 2010.
- [18] M. LAZAR et A. ALEXANDER : Bootstrap white matter tractography (BOOT-TRAC). *Neuroimage*, 24(2):524 – 532, 2005.
- [19] M. MADDAH, W. WELLS, S. WARFIELD, C. WESTIN et W. GRIMSON : Probabilistic clustering and quantitative analysis of white matter fiber tracts. In *Information Processing in Medical Imaging*, vol. 4584 de *LNCS*, p. 372 – 383, 2007.
- [20] B. MOBERTS, A. VILANOVA et J. van WIJK : Evaluation of fiber clustering methods for diffusion tensor imaging. In *IEEE International Conference on Information Visualization, London, VIS '05*, p. 65 – 72, 2005.
- [21] S. MORI, B. CRAIN, V. CHACKO et P. van ZIJL : Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, 45(2):265 – 269, 1999.
- [22] V. NOBLET, C. HEINRICH, F. HEITZ et J. ARMSPACH : Retrospective evaluation of a topology preserving non-rigid registration method. *Medical Image Analysis*, 10(3):366 – 384, 2006.
- [23] L. O'DONNELL et C. WESTIN : White matter tract clustering and correspondence in populations. In *Medical Image Computing and Computer Assisted Intervention, Palm Springs*, vol. 3749 de *LNCS*, p. 140 – 147, 2005.
- [24] L. O'DONNELL et C. WESTIN : High-dimensional white matter atlas generation and group analysis. In *Medical Image Computing and Computer Assisted Intervention, Copenhagen*, vol. 4191 de *LNCS*, p. 243 – 251, 2006.
- [25] S. PAJEVIC et P. BASSER : Parametric and non-parametric statistical analysis of DT-MRI data. *Journal of Magnetic Resonance*, 161(1):1 – 14, 2002.
- [26] G. PARKER et D. ALEXANDER : Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *Information Processing in Medical Imaging*, vol. 2732 de *LNCS*, p. 684 – 695, 2003.
- [27] X. PENNEC : Computing the mean of geometric features : Application to the mean rotation. Institut National de Recherche en Informatique et en Automatique, Rapport de Recherche 3371, 1998.

- [28] J. PONTABRY et F. ROUSSEAU : Probabilistic tractography using q-ball modeling and particle filtering. *In Medical Image Computing and Computer Assisted Intervention, Toronto*, vol. 6892 de *LNCIS*, p. 209 – 216, 2011.
- [29] D. TUCH, T. REESE, M. WIEGELL et V. WEDEEN : Diffusion MRI of complex neural architecture. *Neuron*, 40(5):885 – 895, 2003.
- [30] B. WHITCHER, D. TUCH, J. WISCO, A. SORENSEN et L. WANG : Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging. *Human Brain Mapping*, 29(3):346 – 362, 2008.
- [31] S. ZHANG, S. CORREIA et D. LAIDLAW : Identifying white-matter fiber bundles in DTI data using an automated proximity-based fiber-clustering method. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1044 – 1053, 2008.

PARTIE IV

Simulations et études de cas cliniques

Introduction

LA quatrième partie décrit les différentes étapes de validation mises en place afin d'évaluer les deux chaînes de traitement proposées pour le suivi longitudinal de patients atteints de maladies neurodégénératives et présente des résultats obtenus pour différents cas cliniques.

Le premier chapitre décrit les pathologies neurodégénératives étudiées lors de cette thèse : la sclérose en plaques et la neuromyéélite optique. Ces deux pathologies se caractérisent par une dégénérescence des neurones au niveau du système nerveux central. Le deuxième chapitre présente les résultats de validation obtenus pour la première chaîne de traitement : détection de changement de diffusion à l'échelle du voxel. Enfin, le troisième chapitre présente les résultats de validation obtenus pour la deuxième chaîne de traitement : détection de changements de diffusion sur une modélisation des faisceaux de fibres.

La validation sur données réelles est problématique puisque la construction d'une vérité terrain est quasiment impossible. La construction d'un fantôme de simulation est difficile et ne reflète pas la complexité cérébrale. Par conséquent, nous avons mis en place une validation en deux étapes : la première consiste à simuler différents types de modifications, et la seconde consiste à valider, par un expert neurologue, les résultats obtenus chez des sujets ayant une pathologie. Le recours à des simulations permet de comparer les performances des différentes stratégies et permet également d'effectuer des comparaisons avec des méthodes déjà publiées (cf. chapitre 5). Les simulations sont effectuées à partir de deux examens d'un sujet sain. Comme en routine clinique, aucune précaution n'est prise concernant la position de la tête du sujet dans l'IRM entre les deux acquisitions, ce qui permet de se placer, pour nos simulations, dans des conditions réalistes d'acquisition tout en gardant la maîtrise sur les changements introduits dans l'un des deux examens. Dans un deuxième temps, la validation sur des cas cliniques a été réalisée sur deux bases d'images : une base de patients atteints de sclérose en plaques et une base de patients atteints de neuromyéélite optique. La validation pour la base de sclérose en plaques a été faite visuellement, en comparant les détections obtenues par nos méthodes avec celles segmentées manuellement par un expert sur des IRM morphologiques (T_1 et *FLAIR*). La validation sur la base de neuromyéélite optique a été effectuée par un expert en confrontant les détections obtenues par les différentes méthodes

avec l'évolution des signes cliniques des patients. Dans ce cas précis, les méthodes proposées ont permis de détecter des modifications en imagerie de diffusion au sein de la substance blanche d'apparence normale en imagerie conventionnelle.

Chapitre 10

Les maladies neurodégénératives étudiées

Sommaire

10.1 La sclérose en plaques (SEP)	159
10.1.1 Description de la pathologie	159
10.1.2 Description de la base de patients utilisée	161
10.2 La neuromyéélite optique (NMO)	162
10.2.1 Description de la pathologie	162
10.2.2 Description de la base de patients utilisée	163

L'OBJECTIF de ces travaux est le suivi longitudinal de pathologies neurodégénératives. Une maladie neurodégénérative est caractérisée par la détérioration de certaines cellules nerveuses, en particulier les neurones. L'altération induite peut aller jusqu'à la mort neuronale. Ainsi, la sclérose en plaques (SEP), la neuromyéélite optique (NMO), les maladies d'Alzheimer et de Parkinson sont des maladies dues à une dégénérescence du système nerveux central. Cette thèse s'intéresse particulièrement à la SEP et à la NMO.

10.1 La sclérose en plaques (SEP)

10.1.1 Description de la pathologie

La sclérose en plaques (SEP) est une maladie auto-immune qui évolue par poussées et touche le système nerveux central (cerveau, nerfs optiques et moelle épinière). Le fonctionnement du système nerveux repose sur un grand réseau de neurones interconnectés entre eux. Les neurones sont responsables de l'émission et de la propagation de l'information nerveuse. La structure d'un neurone est présentée en figure 10.1. Un neurone comprend des dendrites, un corps cellulaire, un axone, des terminaisons axonales et des boutons terminaux. L'influx nerveux se déplace des dendrites vers les boutons synaptiques en se propageant le long de l'axone. La gaine de myéline joue un rôle déterminant dans ce transport de l'information nerveuse. En plus de protéger l'axone, elle accélère également la vitesse de transmission de l'influx nerveux. L'influx nerveux saute alors d'un nœud de Ranvier à un autre, c'est la

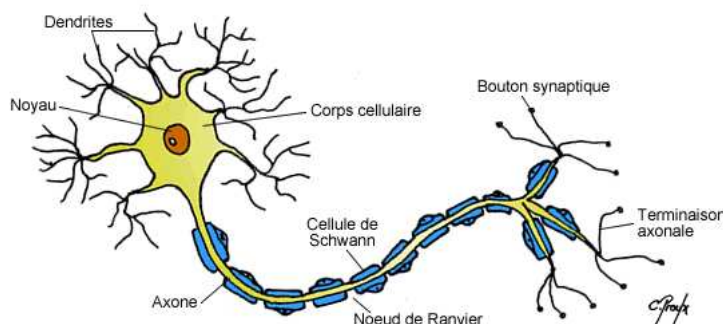


FIG. 10.1 – Le neurone (illustration issue du site de l'université de Nice : <http://www.unice.fr>).

conduction saltatoire.

Dans la suite, nous porterons un intérêt particulier à l'étude du cerveau et plus particulièrement à la substance blanche atteinte par le SEP [13, 3]. La SEP se caractérise par des poussées inflammatoires. Elle est provoquée par un dysfonctionnement du système immunitaire (maladie auto-immune). Dans notre corps, des cellules nous protègent contre des agents infectieux (ou autres qui peuvent nous nuire) en fabriquant des anticorps. Ces cellules sont essentiellement constituées par des lymphocytes T et B. Lors d'une attaque par un agent pathogène, ces cellules sont activées, provoquant l'activation en chaîne d'autres cellules (c'est ce qu'on appelle l'inflammation). La réaction se termine lorsque l'agent pathogène responsable de cette activation est détruit. Dans le SEP, un tel mécanisme d'inflammation se met en place pour des raisons encore inconnues par la communauté médicale. Les cellules qui sont censées nous protéger vont s'activer sans raison et se retourner contre des cellules saines. En particulier, dans la SEP, les lymphocytes T anormalement activés s'attaquent à la myéline qu'ils considèrent comme un agent pathogène et entraînent par endroits la destruction de la gaine de myéline (démýélinisation). C'est cette destruction qui est responsable des symptômes de la SEP. En effet, lorsque la gaine de myéline est altérée, la conduction électrique est moins bonne, les informations ne passent plus correctement ou plus du tout. Notons que l'on parle de « sclérose » car on observe un durcissement des tissus atteints et de « plaques » en raison de la cohérence spatiale des lésions.

Dans la SEP, deux processus compétitifs sont mis en jeu : poussées-rémissions. La rémission peut provoquer une régression des symptômes, mais avec le temps, la cicatrisation se fait moins bien et les lésions deviennent définitives provoquant des symptômes permanents. Le rythme des alternances poussées-rémissions varie d'une personne à l'autre. On distingue généralement trois formes principales de la maladie [13] :

- la forme rémittente se caractérise par des poussées entrecoupées de rémissions et concerne 80-90 % des cas ;
- la forme primaire progressive se caractérise par une évolution lente et constante de la maladie et concerne 10 % des cas ;

Patient	Temps entre les acquisitions	Forme de sclérose en plaques
SP_1	7,5 mois	secondaire-progressive
SP_2	8,5 mois	secondaire-progressive
SP_3	7,5 mois	secondaire-progressive
SP_4	8,5 mois	secondaire-progressive
SP_5	9 mois	secondaire-progressive
SP_6	9,5 mois	secondaire-progressive
PR_1	8,5 mois	poussée-rémission
PR_2	8,5 mois	poussée-rémission
PR_3	9,5 mois	poussée-rémission
PR_4	8,75 mois	poussée-rémission
PP_1	7,5 mois	primaire-progressive
PP_2	8 mois	primaire-progressive

TAB. 10.1 – Données sur les patients SEP du CHU de Besançon.

- la forme secondaire progressive est une forme rémittente qui évolue vers une forme primaire progressive.

A ce jour, aucun traitement n'est capable de stopper l'activation intempestive des lymphocytes T.

En radiologie, toute lésion cérébrale est susceptible de produire un hypersignal de la substance blanche en IRM pondérée T_2 ou en IRM FLAIR et un hyposignal de la substance blanche en en IRM pondérée T_1 . Il faut être vigilant avant de considérer la présence de lésions, des variations du signal pouvant provenir d'autres phénomènes (par exemple une hypertension artérielle). Le diagnostic de sclérose en plaques est donc basé sur des critères stricts aussi bien de diffusion des lésions que d'évolution des lésions en IRM de contrôle (critères de Barkhof) [6].

Finalement, la SEP est une maladie complexe dont l'évolution est à ce jour imprévisible. Ni le nombre de poussées, ni le type d'atteinte, ni l'âge du diagnostic ne permettent de prévoir ou d'envisager l'avenir de la personne qui est atteinte. L'enjeu de ce travail de thèse est de quantifier l'évolution de la pathologie et d'apporter aux neurologues un marqueur de l'évolution de la maladie à partir de l'imagerie de diffusion.

10.1.2 Description de la base de patients utilisée

Les images de patients atteints de SEP ont été acquises à Besançon (service de neurologie, CHU J. Minjoz, Besançon). Les informations sur ces patients sont résumées dans le tableau 10.1. Les acquisitions ont été réalisées à huit mois d'intervalles environ.

Les images des patients atteints de SEP ont été acquises sur un scanner IRM General Electric (GE) de 3 Tesla. Le protocole incluait l'acquisition d'une image pondérée en T_1 , d'une image

FLAIR et d'une série d'images de diffusion avec 33 directions de gradient et une b-valeur de 1000 s/mm^2 . Pour la diffusion, les images ont une résolution de $1 \times 1 \times 3 \text{ mm}^3$ (la résolution isotropique se situe au niveau des plans de coupe qui sont séparés les uns des autres de 3 mm).

10.2 La neuromyéélite optique (NMO)

10.2.1 Description de la pathologie

La neuromyéélite optique (NMO) est une maladie rare qui a longtemps été considérée comme une forme particulière de la SEP [1]. Comme la SEP, la NMO est une pathologie inflammatoire du système nerveux central qui se caractérise par l'association d'épisodes de myélite (inflammation de la moelle épinière) extensives et de névrites optiques (atteintes aiguës ou dégénératives du nerf optique). La découverte récente d'un anticorps spécifique de la NMO (immunoglobuline G ou IgG), appelé anticorps anti-NMO, a fait de cette maladie une maladie à part entière. Les anticorps anti-NMO sont utilisés comme marqueurs biologiques pour le diagnostic et sont dirigés contre l'aquaporine 4 (AQP4), une protéine spécialisée dans la perméabilité des membranes à l'eau. Les experts considèrent la NMO comme une pathologie de l'immunité humorale (liée aux lymphocytes B et à la production d'anticorps).

Chez les patients atteints de NMO, on observe une démyélinisation au niveau du nerf optique et de la moelle épinière [1]. Lors d'épisode de myélite, de longues lésions le long de la moelle épinière apparaissent et sont visibles en IRM. L'IRM cérébrale, quant à elle, est normale au début de la maladie, puis seul 10 % des patients présentent des lésions correspondant aux critères radiologiques de la SEP [6]. Dans un premier temps, l'enjeu de notre travail est de fournir aux neurologues une analyse détaillée de l'IRM de diffusion cérébrale qui pourrait permettre de mettre en évidence des évolutions locales liées à l'évolution de la maladie.

La NMO provoque également des troubles cognitifs et moteurs. Pour mesurer l'étendue des troubles cognitifs, une batterie de tests cognitifs, initialement proposée pour évaluer les patients souffrant de SEP (BCogSEP) a été utilisée dans une version modifiée [5]. Pour ce faire, trois tests ont été ajoutés. Cette batterie se décline en 14 sous-ensembles permettant d'évaluer la mémoire verbale, la mémoire visuo-spatiale, la mémoire immédiate, la mémoire de travail, les fonctions exécutives, l'attention et la vitesse de traitement de l'information. Les troubles moteurs ont été évalués avec l'échelle EDSS (Expanded Disability Status Scale) [10]. Cette échelle tient compte des capacités ambulatrices mais également des paramètres fonctionnels du patient. L'échelle varie de 0 (pas de handicap) à 10 (décès du patient). L'EDSS permet d'évaluer la fonction pyramidale (la marche), la fonction cérébelleuse (la coordination), la parole et la déglutition, les fonctions sensitives (toucher et douleur), les fonctions intestinale et urinaire, la vision et les capacités mentales. Dans un deuxième temps, nous chercherons à interpréter les évolutions détectées chez chaque patient en les corrélant avec les signes cliniques observés.

Patient	Δt	Anticorps anti-NMO	ES au temps t_1	ES au temps t_2
P_1	18 mois	-	2	2
P_2	41 mois	-	5	8
P_3	17 mois	-	-	3
P_4	10 mois	-	-	-
P_5	13 mois	-	-	-
P_6	19 mois	-	1	2
P_7	14 mois	+	2	2
P_8	22 mois	-	3	4
P_9	20 mois	+	5	7
P_{10}	21 mois	+	-	-
P_{11}	11 mois	+	3	3

TAB. 10.2 – Données sur les patients NMO du CHU de Strasbourg. Les extensions spatiales (ES) des lésions de la moelle épinière sont quantifiées en nombre de segments vertébraux et le temps entre les acquisitions est noté Δt .

10.2.2 Description de la base de patients utilisée

Les images de patients atteints de NMO ont été acquises à Strasbourg (service de neurologie, hôpital de Haute-pierre, Strasbourg). Les informations sur ces patients sont résumées dans le tableau 10.2.

Les images des patients atteints de NMO ont été acquises sur un scanner IRM Siemens de 1,5 Tesla. Le protocole incluait l'acquisition d'une image pondérée en T_1 et d'une série d'images de diffusion du cerveau avec 30 directions de gradient et une b-valeur de 1000 s/mm^2 et l'acquisition d'une image pondérée en T_2 de la moelle épinière. Pour la diffusion, les images ont une résolution de $1,8 \times 1,8 \times 3,5 \text{ mm}^3$. La résolution isotropique se situe au niveau des plans de coupe qui sont séparés les uns des autres de $3,5 \text{ mm}$. Tous les patients respectent les critères modifiés de diagnostic pour la neuromyéélite optique proposés en 2006 par Wingerchuck [14]. Tous les patients ont déjà eu des lésions au niveau de la moelle épinière, au moins au début de la maladie. Quatre patients sur onze sont séropositifs pour les anticorps anti-aquaporine 4 et huit patients ont des lésions au niveau de la moelle épinière au moins sur l'une des images.

On peut noter que dans notre base de donnée NMO, le taux de patients séropositifs pour les anticorps anti-aquaporine 4 est faible (4/11). Une étude menée par la clinique Mayo de Rochester annonce un taux de 73% pour des patients nord américains [11], alors qu'une étude plus récente menée sur une cohorte française de patients NMO annonce un taux de 54% [4]. De plus, cette étude a montré que le test de séropositivité aux anticorps anti-aquaporine 4 est nécessaire pour confirmer le diagnostic (d'après les critères de Wingerchuck) dans seulement 10% des cas.

Finalement, la SEP et la NMO mettent en jeu des processus de démyélinisation. Des travaux ont montré l'impact de la démyélinisation sur les images de diffusion [8]. Les méthodes de

détection proposées dans le chapitre 9 pour l'imagerie de diffusion sont ainsi des outils pertinents pour trouver des marqueurs de l'évolution de ces pathologies.

Chapitre 11

Evaluation des méthodes de détection de changement de diffusion à l'échelle du voxel

Sommaire

11.1 Simulations	166
11.1.1 Distributions sous l'hypothèse nulle	166
11.1.2 Résultats sur données simulées	167
11.1.3 Evaluation du taux de recouvrement	171
11.1.4 Influence de l'anisotropie des images	171
11.2 Application à des données cliniques	172
11.2.1 Etude détaillée de deux patients	172
11.2.2 Base de données NMO	176

LA chaîne de traitement « détection de changement de diffusion à l'échelle du voxel » est mise en place pour faire le suivi longitudinal de patients atteints de SEP et de NMO (cf. section 9.1). Rappelons que cette stratégie permet de détecter des modifications des propriétés de diffusion en s'intéressant aux valeurs propres des tenseurs de diffusion. La nature multivariée de ces données impose la mise en place d'un test statistique multivarié. Ce test repose sur une méthode de voisinage spatial, de bootstrap local ou de bootstrap spatial lors de la construction des échantillons de tenseurs à comparer, reflétant la variabilité en chaque point de l'image. Avant d'exploiter les données des patients, la méthode est évaluée sur des données simulées. Dans un premier temps, nous verrons que la distribution multinormale matricielle pour le logarithme des matrices symétriques définies positives rend bien compte des données de diffusion. Des simulations sont alors menées pour comparer les différentes stratégies de construction des échantillons, entre elles et avec une méthode de la littérature : le GLRT [2]. La performance des approches proposées en terme de localisation spatiale et l'influence de l'anisotropie dans la résolution des images sur le résultat final sont aussi évaluées. Ensuite, une étude détaillée d'un patient atteint de sclérose en plaques et d'un patient atteint de neuromyéélite optique est proposée. Finalement, les résultats sur l'ensemble de la base de patients atteints de NMO sont présentés.

11.1 Simulations

11.1.1 Distributions sous l'hypothèse nulle

Tout d'abord, nous souhaitons vérifier la validité d'une des hypothèses de base sur laquelle est construite le test statistique T_{e-val} (cf. section 9.1), à savoir que les éléments des log-tenseurs suivent une distribution multinormale $\mathbf{L} \sim \mathcal{N}_{33}(\mathbf{M}, \sigma^2 \mathbf{I}_3)$ (*i.e.*, les six éléments indépendants suivent une distribution normale). Le nombre N d'échantillons générés par les méthodes de bootstrap est choisi empiriquement et l'effet de ce paramètre sur la distribution, sous l'hypothèse nulle, est également étudié. Trois régions, situées au sein de la matière blanche, de la matière grise et du liquide céphalo-rachidien, ont été choisies pour effectuer cette expérience (cf. figure 11.1). A ces endroits, les histogrammes d'un élément appartenant à la diagonale (d_{33}) et à l'anti-diagonale (d_{23}) ont été tracés pour la méthode spatiale, le bootstrap local et le bootstrap spatial (cf. figure 11.2). Pour la méthode spatiale, nous utiliserons un voisinage $5 \times 5 \times 5$ (pour avoir assez d'éléments pour construire l'histogramme), soit 125 log-tenseurs. Pour le bootstrap spatial, un voisinage $3 \times 3 \times 3$ est utilisé. Pour les deux méthodes de bootstrap, deux échantillonnages différents ont été considérés, $N = 200$ ou $N = 1000$. Les distributions issues de la méthode de bootstrap local sont les plus proches d'une distribution gaussienne. Augmenter le nombre d'échantillons bootstrap renforce la validité de l'hypothèse gaussienne. Néanmoins, un échantillonnage bootstrap de $N = 200$ semble être un bon compromis entre temps de calcul et qualité des résultats obtenus lors des tests statistiques. Cette observation rend légitime l'utilisation des tests paramétriques développés dans le chapitre 6 sur des données de diffusion. Notons que le coût de calcul informatique d'une procédure bootstrap est important, mais semble engendrer des population de tenseurs respectant plus fidèlement les hypothèses du test statistique.

Ensuite, on se propose d'étudier la distribution de la statistique T_{e-val} sous l'hypothèse nulle. Pour ce faire, deux acquisitions d'un sujet sain ont été réalisées à deux jours d'intervalle sur un scanner IRM Siemens de 3 Tesla. La séquence de diffusion comporte 64 directions de gradient pour une b-valeur de 1000 s/mm^2 . Les images pondérées en diffusion ont une résolution isotropique de $2 \times 2 \times 2 \text{ mm}^3$. Puisque le sujet est sain et que le temps entre les deux examens est court, aucune modification n'est attendue. Les histogrammes des valeurs prises par la statistique de test T_{e-val} sur l'ensemble du cerveau pour les méthodes de voisinage spatial (S), de bootstrap spatial (BS) et de bootstrap local (BL) sont représentés en figure 11.3. La distribution théorique et la distribution empirique ajustée sur les données sont estimées en utilisant la méthodologie proposée en section 6.2.3 et sont tracées également sur la figure 11.3. On remarque que l'estimation de la distribution nulle empirique donne le meilleur ajustement des données, ce qui nous permettra dans la suite de se baser sur cette distribution pour calculer les p-valeurs et ainsi accepter ou rejeter l'hypothèse nulle en fonction d'un seuil de signification spécifié par l'utilisateur et en utilisant une procédure FDR pour tenir compte du problème lié aux comparaisons multiples.

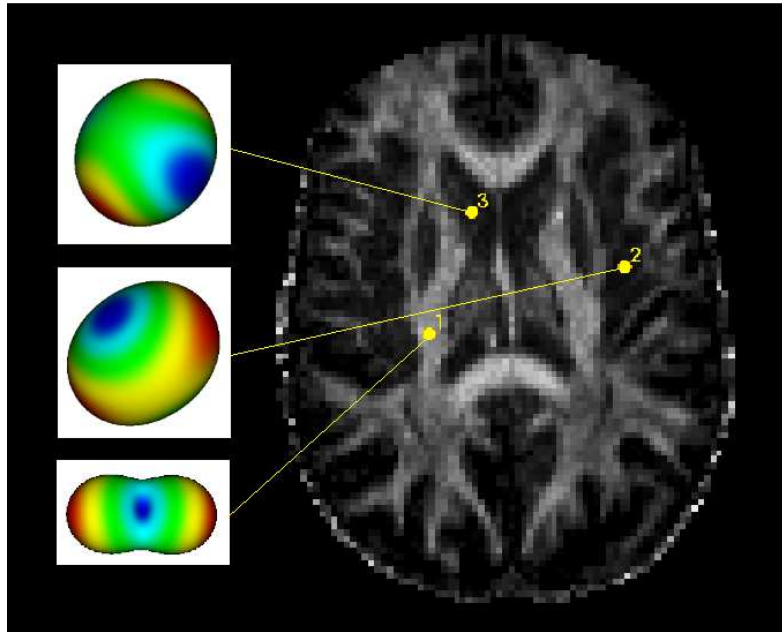


FIG. 11.1 – Visualisation des tenseurs dans trois régions d'intérêt situées dans la substance blanche (1), la substance grise (2) et le liquide céphalorachidien (3).

11.1.2 Résultats sur données simulées

Le recours à des simulations permet de comparer les performances des différentes approches proposées et permet également de se situer par rapport au GLRT proposé dans Boigontier *et coll.* [2]. L'évaluation des méthodes de détection de changements sur données réelles est un problème complexe puisque la construction d'une vérité terrain par les neurologues est quasiment impossible. Cette étape est encore plus difficile en IRMd où les données à analyser sont multivariées. C'est pour cette raison que les méthodes sont dans un premier temps validées sur données simulées.

Soient deux acquisitions d'un sujet sain, acquises à deux jours d'intervalle sur un IRM Siemens de 3,0 Tesla. La séquence de diffusion comporte 64 directions de gradient pour une b-valeur de 1000 s/mm^2 . De ce fait, les différences entre les deux examens sont uniquement dues au bruit et aux artefacts d'acquisition, ainsi qu'à un changement de position de la tête du patient dans l'imageur (on se place dans les conditions cliniques où aucun contrôle ne permet de corriger le mouvement possible de la tête du patient par rapport à un examen précédent). Les images pondérées en diffusion sont chacune de taille $96 \times 96 \times 55$ et la résolution spatiale est isotropique : $2 \times 2 \times 2 \text{ mm}^3$.

Une lésion est simulée dans un des deux examens de la manière suivante. Un masque binaire est délimité par un neurologue dans la substance blanche, avec une taille et une forme représentatives d'une lésion. Dans cette zone, la diffusion est augmentée uniformément dans la direction principale (modification de la valeur propre principale) en appliquant un facteur multiplicatif $k \in]1; 1,2]$.

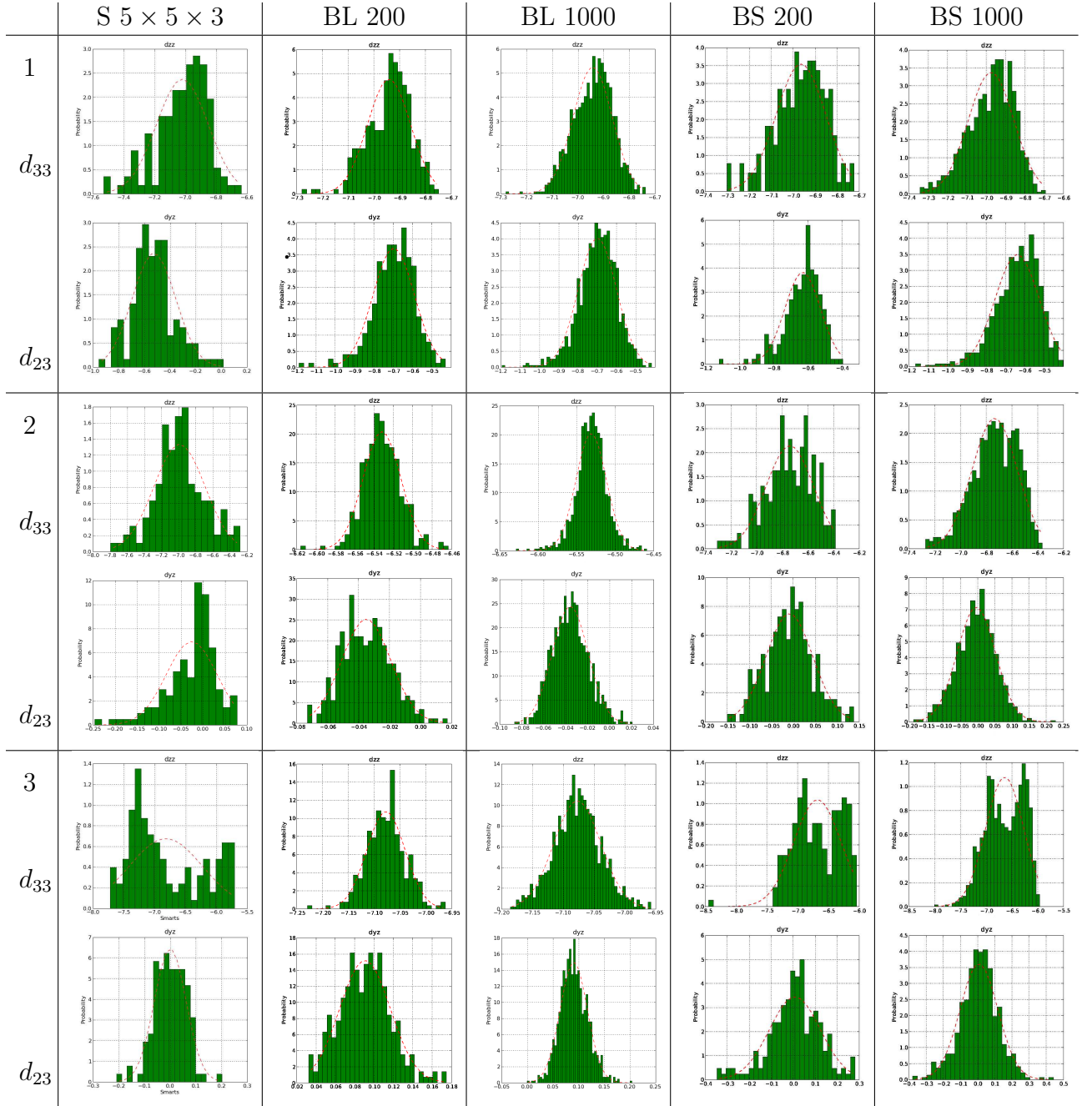


FIG. 11.2 – Distribution à la position 1, 2 et 3 des éléments d_{33} et d_{23} des log-tenseurs pour la méthode spatiale et les méthodes de bootstrap en considérant $N \in \{200, 1000\}$ échantillons. La courbe en pointillée représente une distribution gaussienne ajustée sur les données.

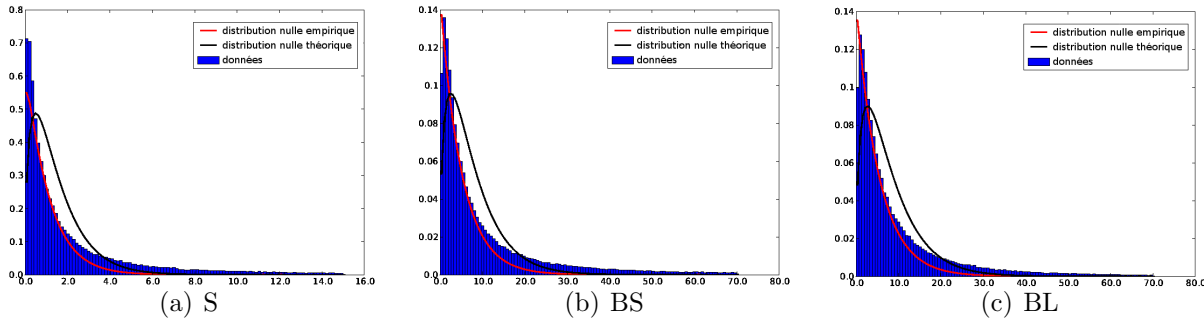


FIG. 11.3 – Histogrammes de la statistique de test T_{e-val} sous l'hypothèse nulle, la distribution nulle empirique ajustée sur les données et la distribution théorique pour les méthodes : a. S, b. BS et c. BL.

Pour classer les méthodes en terme de performance, un critère basé sur l'aire sous les courbes COR (Caractéristique Opérationnelle du Récepteur ou, en anglais, *Receiver Operating Characteristic*) est utilisé. Les courbes COR représentent le taux de vrais positifs (ou sensibilité) en fonction du taux de faux positifs (fausses alarmes ou 1-spécificité) [7]. Le test statistique produit une carte de statistiques sur laquelle différents seuils peuvent être appliqués pour classer les voxels (*i.e.*, les voxels sont détectés ou non, l'hypothèse nulle est acceptée ou rejetée). Seuiller la carte de statistiques permet de calculer à partir d'une vérité terrain (qui est connue en simulation : masque défini par le neurologue) un couple (taux de faux positifs, taux de vrais positifs). Ce couple correspond à un point sur la courbe COR. Une séparation parfaite des deux classes est caractérisée par une courbe COR qui vaut 0 en 0 puis 1 partout ailleurs, soit une aire sous cette courbe de 1. Les courbes COR ont été calculées à l'échelle du voxel sur l'intégralité du cerveau en utilisant les statistiques de test non corrigées pour les comparaisons multiples.

Les résultats sont comparés avec la méthode GLRT proposée par Boigontier *et coll.* dans [2]. Cette méthode est similaire à la méthode de Bootstrap Local dans le sens où une inférence statistique est effectuée en chaque voxel de l'image sur la distribution tensorielle à partir du signal de diffusion. Il existe toutefois une différence significative entre ces deux approches. Le GLRT n'utilise pas la même distribution sous l'hypothèse nulle et travaille sur des matrices symétriques. Les stratégies proposées permettent une meilleure prise en compte de la nature tensorielle des modèles d'ordre 2 (tests statistiques sur des matrices symétriques définies positives).

Le tableau 11.1 présente les résultats obtenus pour les quatre stratégies étudiées : S, BS, BL et GLRT. Pour tous les traitements réalisés, un voisinage spatial $3 \times 3 \times 3$ est considéré pour les stratégies S et BS. L'influence du nombre d'échantillons bootstrap N sur la statistique est également explorée (cf. tableau 11.1). Au vu des résultats présentés dans le tableau 11.1, un choix de $N = 200$ ou $N = 400$ échantillons bootstrap permet d'obtenir le meilleur compromis entre le temps de calcul et la qualité des résultats. Pour les différentes modifi-

k	S	BS					
	$3 \times 3 \times 3$	N=100	N=200	N=400	N=600	N=800	N=1000
1,02	0,559	0,550	0,573	0,585	0,579	0,579	0,580
1,04	0,592	0,691	0,706	0,711	0,716	0,714	0,714
1,06	0,641	0,775	0,793	0,804	0,803	0,808	0,805
1,08	0,683	0,839	0,854	0,859	0,860	0,866	0,863
1,1	0,730	0,882	0,885	0,898	0,900	0,900	0,901
1,2	0,867	0,967	0,965	0,969	0,972	0,972	0,973
1,4	0,957	0,981	0,983	0,985	0,987	0,988	0,989
1,6	0,981	0,983	0,986	0,988	0,990	0,990	0,991
1,8	0,990	0,983	0,986	0,989	0,991	0,991	0,991
2,0	0,993	0,984	0,986	0,989	0,991	0,991	0,991
k	GLRT	BL					
	-	N=100	N=200	N=400	N=600	N=800	N=1000
1,02	0,398	0,456	0,449	0,450	0,447	0,445	0,447
1,04	0,410	0,461	0,454	0,453	0,450	0,450	0,450
1,06	0,425	0,465	0,464	0,462	0,460	0,459	0,459
1,08	0,447	0,474	0,479	0,475	0,474	0,474	0,474
1,1	0,479	0,487	0,494	0,493	0,495	0,492	0,493
1,2	0,646	0,602	0,610	0,611	0,610	0,610	0,610
1,4	0,826	0,771	0,774	0,776	0,775	0,775	0,776
1,6	0,884	0,847	0,846	0,848	0,848	0,848	0,849
1,8	0,911	0,881	0,882	0,882	0,883	0,884	0,885
2,0	0,926	0,900	0,901	0,901	0,901	0,903	0,905

TAB. 11.1 – Aires sous les courbes COR pour le test sur les valeurs propres T_{e-val} couplé avec les différentes méthodes d'échantillonnage et pour le GLRT. k est un facteur multiplicatif utilisé pour modifier la diffusion dans la direction principale. Dans chaque ligne du tableau, l'échantillonnage qui donne les meilleurs résultats pour les méthodes de BS et de BL a été repéré en gras.

cations simulées, les deux méthodes basées sur l'information spatiale permettent d'obtenir les meilleurs résultats. Le voisinage spatial joue un rôle important et permet d'obtenir une détection de changements fiable. La méthode donnant les meilleurs performances est le BS. Cette méthode combine à la fois l'information spatiale et le rééchantillonnage bootstrap. Il est également possible d'utiliser un filtrage gaussien sur les images pondérées en diffusion pour introduire une cohérence spatiale dans les méthodes locales (BL et GLRT) et ainsi en améliorer leurs performances [2]. Ce filtrage n'est pas réalisé. Toutefois, les méthodes locales peuvent être comparées équitablement entre elles. Au vu des résultats simulés (cf., tableau 11.1), le BL semble moins efficace que le GLRT.

k	1,02	1,04	1,06	1,08	1,1	1,2	1,4	1,6	1,8	2,0
S	0,212	0,230	0,276	0,341	0,425	0,665	0,811	0,867	0,888	0,910
BS	0,275	0,546	0,725	0,793	0,831	0,898	0,923	0,935	0,936	0,942
BL	0,162	0,196	0,196	0,203	0,437	0,536	0,813	0,858	0,872	0,883

TAB. 11.2 – Taux de recouvrement maximal entre les détections et la vérité terrain. k est un facteur multiplicatif utilisé pour modifier la diffusion dans la direction principale. Le nombre d'échantillons bootstrap N vaut 200 et le voisinage spatial est de $3 \times 3 \times 3$.

11.1.3 Evaluation du taux de recouvrement

Dans cette section, nous présentons l'évaluation de la performance des différentes méthodes vis-à-vis de la localisation spatiale des modifications détectées. Le taux de recouvrement maximal entre la détection de la lésion simulée et la vérité terrain est calculé pour tous les seuils utilisés lors du calcul des courbes COR de la section précédente. Une analyse en composantes connexes permet d'obtenir un ensemble de groupes de voxels voisins. Les groupes qui ont une intersection non nulle avec la vérité terrain permettent le calcul du taux de recouvrement. Soient ROI_{vt} et ROI_i les régions d'intérêt associées à la vérité terrain et au groupe de voxels i . Le taux de recouvrement s'exprime de la manière suivante :

$$TR = \frac{2 \cup_i (ROI_i \cap ROI_{vt})}{\cup_i ROI_i + ROI_{vt}} \quad (11.1)$$

A partir de l'ensemble des taux de recouvrement (obtenus pour chaque seuil), le maximum est conservé. Les résultats sont présentés dans le tableau 11.2. Globalement, il apparaît clairement que le BS permet une meilleure localisation de la modification simulée. Les méthodes S et BL sont légèrement moins performantes que le BS, mais donnent des résultats satisfaisants.

11.1.4 Influence de l'anisotropie des images

Les images de diffusion acquises en routine clinique sont généralement anisotropiques (*i.e.*, l'épaisseur de coupe est plus importante que la résolution dans le plan d'acquisition). Des variations d'anisotropie peuvent avoir une influence directe sur la chaîne de traitement. En effet, le rééchantillonnage (l'interpolation) d'images anisotropiques lors de l'étape de recalage risque d'être sujet à davantage d'erreurs et d'artefacts. Nous proposons d'étudier l'influence de l'anisotropie sur la chaîne de traitement proposée, afin de s'assurer que les modifications observées lors du suivi longitudinal d'un patient ont un réel substrat anatomo-pathologique et ne sont pas de simples artefacts de détection induits par le recalage des deux examens.

Pour effectuer cette vérification, deux acquisitions d'un sujet sain ont été réalisées le même jour sur un scanner IRM Siemens de 3 Tesla (le patient est ressorti du scanner entre les deux acquisitions et le protocole d'acquisition a été recommencé). Comme en routine clinique, aucune précaution particulière n'a été prise concernant la position de la tête du sujet dans l'IRM entre les deux acquisitions. La séquence de diffusion comporte 31 directions de gradient pour une b-valeur de 1000 s/mm^2 . Un recalage rigide rétrospectif des deux séquences

a permis d'évaluer le mouvement du patient entre les deux acquisitions, soit une translation de $[-0,292; -4,378; 5,899]$ mm et une rotation autour des axes x, y, et z de $[2,596; 0,628; 0,226]$ degrés. Chaque acquisition comporte trois séquences de diffusion ayant chacune un ratio d'anisotropie différent : 1 (résolution spatiale de $2 \times 2 \times 2 \text{ mm}^3$), 2 (résolution spatiale de $1,8 \times 1,8 \times 3,5 \text{ mm}^3$) et 3 (résolution spatiale de $1 \times 1 \times 3 \text{ mm}^3$). La chaîne de traitement avec voisinage spatial (S) est lancée pour effectuer les suivis longitudinaux aux trois anisotropies (on choisit cette stratégie car elle est utilisée pour étudier la base de patients NMO en section 11.2.2). Les cartes de statistique obtenues ont été seuillées avec la procédure FDR basée sur la distribution sous l'hypothèse nulle empirique décrite en section 6.2.3. Un seuil de signification $p_{FDR} < 0,001$, identique au seuil utilisé lors des expérimentations cliniques, est utilisé. Puisqu'un sujet sain est utilisé lors de cette expérience, peu de changements sont attendus, le nombre de fausses alarmes étant directement lié au seuil de signification choisi.

Pour un facteur d'anisotropie de 1 (données isotropiques) et 3, aucune modification n'est détectée (la valeur maximale p_{FDR} pour chaque acquisition étant de 0,009 et de 0,004 respectivement). Pour un facteur d'anisotropie de 2, 20 groupes de voxels connexes sont détectés. Sur ces 20 groupes, seulement 2 groupes contiennent plus de 2 voxels et sont donc considérés comme des détections potentielles. Le premier groupe contient 7 voxels voisins et se situe dans la partie supérieure du cerveau, près d'un sillon. Cette modification est probablement due à une erreur de recalage. Notons qu'en pratique et au vu de la localisation et de la taille du changement, il est immédiatement mis à l'écart lors de l'interprétation des résultats par le neurologue. Le deuxième groupe contient 5 voxels voisins (*i.e.*, $\sim 60 \text{ mm}^3$) et se situe près du tronc cérébral dans la substance blanche. La taille de cette évolution reste toutefois petite par rapport aux modifications observées lors de l'application de cette méthode sur des patients atteints de NMO ($> 450 \text{ mm}^3$). Ainsi, cette expérience montre qu'utiliser des images ayant des ratios d'anisotropie différents n'introduit pas de détections parasites pouvant être confondues avec des changements pertinents causés par la pathologie étudiée.

11.2 Application à des données cliniques

11.2.1 Etude détaillée de deux patients

Un patient atteint de SEP

Deux images ont été acquises sur un scanner IRM General Electric de 3,0 Tesla ($t_2 - t_1 = 13$ mois). La séquence de diffusion comporte 31 directions de gradient pour une b-valeur de 1000 s/mm^2 . Les images pondérées en diffusion sont chacune de taille $256 \times 256 \times 34$ et ont une résolution spatiale anisotrope de $1 \times 1 \times 3 \text{ mm}^3$. L'évaluation est basée sur la comparaison entre les modifications détectées par le système automatique proposé et les modifications labélisées par le neurologue à partir d'images anatomiques. Le neurologue a pris sa décision en se basant sur deux modalités : une image pondérée en T_1 et une image FLAIR. Le système automatique réalise un test statistique sur les valeurs propres. Il met en œuvre un voisinage spatial $3 \times 3 \times 3$ pour les stratégies de rééchantillonnage S et BS et utilise $N = 200$ échantillons bootstrap pour les méthodes BL et BS.

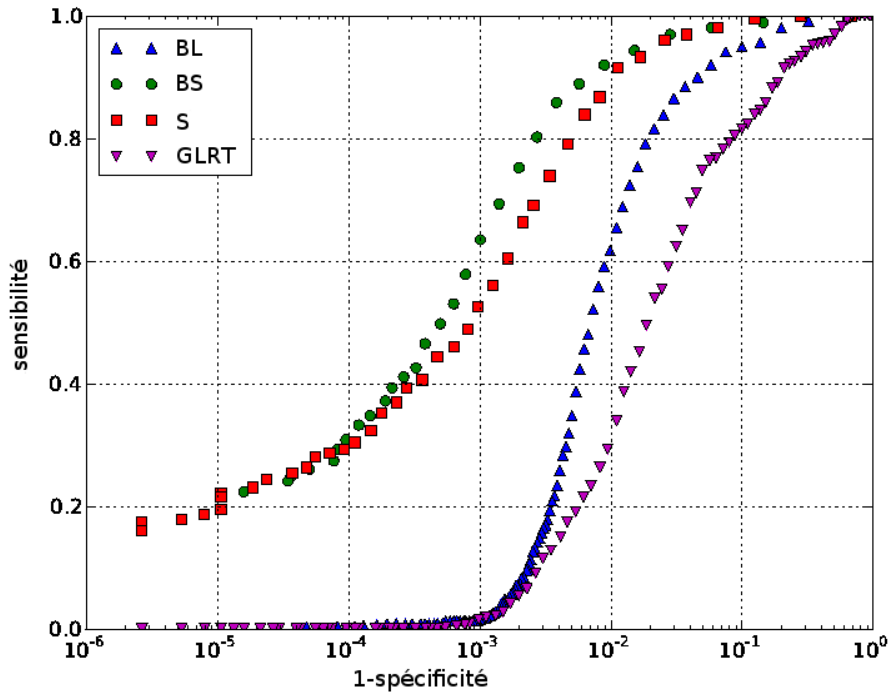


FIG. 11.4 – Courbes COR (tracé semilogx) : résultats du suivi longitudinal chez un patient atteint de SEP en utilisant les trois méthodes de rééchantillonnage proposées, à savoir S, BL, BS et le GLRT.

Les comparaisons entre les évolutions labélisées par le neurologue et les évolutions détectées par la chaîne de traitement proposée se font via des courbes COR. De plus, une comparaison est effectuée entre les trois méthodes de rééchantillonnage proposées et le GLRT (cf. figure 11.4). Les courbes COR doivent être analysées avec prudence car la labélisation des lésions par le neurologue ne peut pas être considérée comme une réelle vérité terrain dans la mesure où des changements peuvent survenir en imagerie de diffusion sans être visible en imagerie anatomique. D'ailleurs, c'est justement l'apport espéré de cette nouvelle modalité d'imagerie. La figure 11.4 confirme les résultats obtenus précédemment sur données simulées. Les deux méthodes basées sur une extension spatiale (S et BS) ressortent de manière significative, le BS restant la méthode la plus performante. Chose intéressante, le BL semble donner de meilleurs résultats que le GLRT, mais il est toutefois difficile de conclure que le résultat est significativement meilleur.

Une analyse visuelle des résultats par un neurologue permet d'affirmer que la plupart des évolutions détectées par les approches proposées sont aussi visibles en imagerie anatomique. Par ailleurs, d'autres changements sont détectés, menant à une interprétation différente des résultats. Dans la figure 11.5, deux changements correspondant à des lésions en imagerie anatomique sont détectés. Mais, d'après le neurologue (en interprétant les images anatomiques uniquement), seule la lésion située dans la ROI inférieure évolue. La lésion dans la ROI

supérieure est une lésion pré-existante, d'intérêt certain, puisque, conformément aux images anatomiques et à l'analyse du neurologue, elle ne semble pas évoluer entre les deux examens. La lésion dans la ROI inférieure est une nouvelle lésion qui apparaît dans le deuxième examen. Au vu des résultats, les méthodes basées sur un voisinage spatial donnent des cartes de détection plus lisses (*i.e.*, contenant moins de détections isolées). Néanmoins, elles tendent à élargir la lésion détectée (par rapport à l'évolution que l'on peut observer sur les images pondérées en T_2). Contrairement aux simulations, la méthode de BL permet une meilleure localisation des lésions. Elle reste toutefois sensible aux erreurs de recalage, particulièrement importantes dans les aires corticales. Les trois méthodes proposées permettent la détection d'une petite évolution (261 mm^3) dans la partie centrale de la lésion pré-existante. Cette observation montre que la lésion est encore active. Concernant la deuxième détection, la taille de cette nouvelle lésion est sur-estimée par toutes les méthodes. Cette sur-segmentation pourrait corroborer l'idée d'une détection de l'infiltration des tissus sains en périphérie de la lésion. La figure 11.6 met en évidence l'apparition d'une petite lésion. Cette modification n'est visible que sur une coupe. Il semble que les méthodes spatiales atteignent leurs limites sur cet exemple puisqu'elles ne permettent pas de détecter ce changement (sûrement dû au lissage induit par l'utilisation d'un voisinage spatial). La méthode locale permet de faire une détection de la lésion. Cette situation (évolution sur une seule coupe) n'est pas un fait isolé en imagerie de diffusion puisque l'épaisseur de coupe est relativement importante.

Un patient atteint de NMO

Les images ont été acquises par un scanner IRM Siemens de 1,5 Tesla lors de deux examens ($t_2 - t_1 = 41$ mois). La séquence de diffusion comporte 30 directions de gradient pour une b-valeur de 1000 s/mm^2 . Les images pondérées en diffusion sont chacune de taille $128 \times 128 \times 41$ et ont une résolution spatiale anisotrope de $1,8 \times 1,8 \times 3,5 \text{ mm}^3$. Il est important de noter que l'analyse proposée dans cette étude s'intéresse à des modifications induites par la NMO au niveau de la substance blanche apparemment normale et non à des lésions de NMO (connues pour être plus destructives que des lésions de SEP [12]). La figure 11.7 présente une comparaison des trois méthodes de rééchantillonnage sur un patient atteint de NMO. Comme pour la SEP, les méthodes de BS et S donnent des résultats similaires. Cependant, la méthode de BL détecte uniquement des modifications induites par des artefacts de bruit et des erreurs de recalage. Ces détections se situent préférentiellement dans les aires corticales, et souligne le fait que la méthode de BL ne semble pas adéquate pour le suivi longitudinal en NMO.

Il n'est pas surprenant que les approches locales et spatiales donnent des résultats différents puisqu'elles s'appuient sur des procédures de rééchantillonnage statistique reflétant différents types de variabilité. Ces différences d'estimation de variabilité dans les images de diffusion sont illustrées en figure 11.8. Il apparaît qu'avec la méthode de BL (cf. figure 11.8-c), la variabilité en orientation et forme des tenseurs induit par la variabilité du signal de diffusion est plus importante dans les régions de forte anisotropie (*i.e.*, dans la substance blanche). Ainsi, la sensibilité du test statistique est diminuée dans ces endroits, puisque le test statistique utilisé est inversement proportionnel à la variance. Mais ces régions de forte anisotropie sont également des régions d'intérêt susceptibles de contenir des modifications lors du suivi

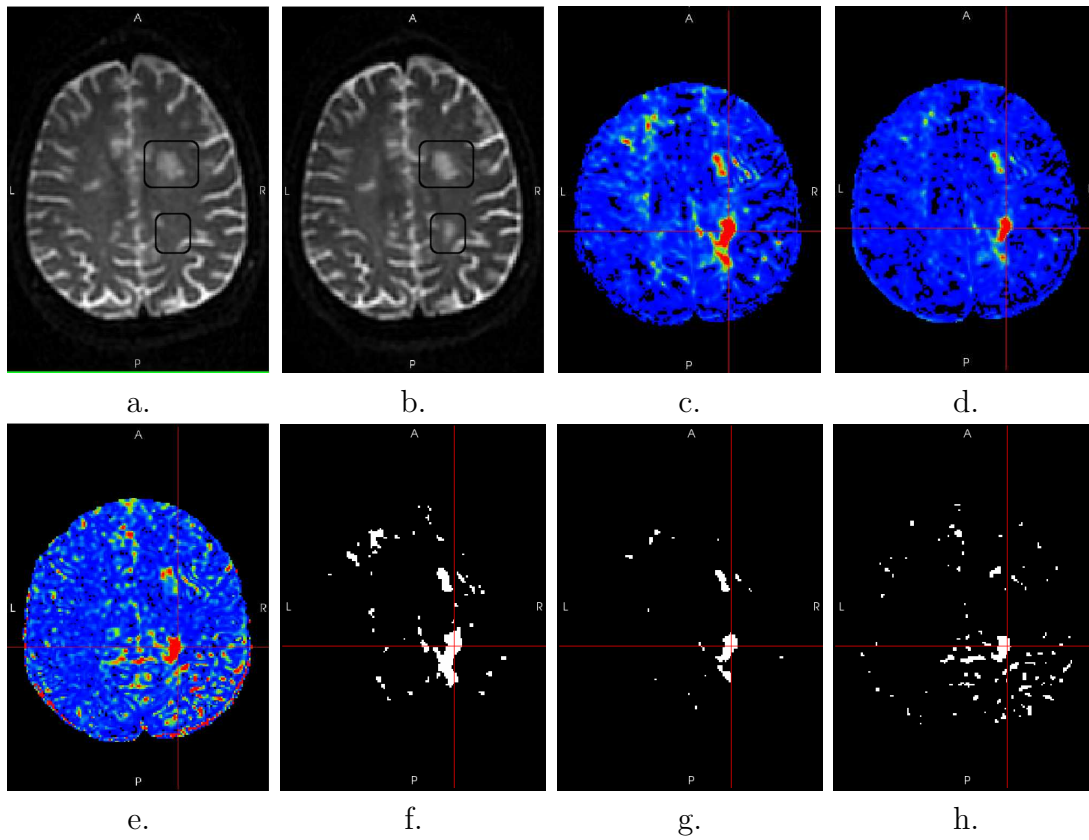


FIG. 11.5 – Cartes de détection obtenues lors du suivi longitudinal (avec le test sur les valeurs propres) d'un patient atteint de SEP et l'utilisation des trois méthodes de rééchantillonnage proposées. a. images pondérée en T_2 et acquise au temps t_1 , b. images pondérée en T_2 et acquise au temps t_2 , méthodes c. de BS, d. S et e. de BL. Les cartes seuillées ont été obtenues après une correction FDR ($p_{FDR}=0,001$) pour les méthodes de BS, S et de BL. Les résultats sont présentés dans les images f., g. et h., dans cet ordre. La plage des couleurs varie du bleu au rouge. Tous les voxels contenant une valeur au dessus du seuil maximal de visualisation sont représentés par une couleur rouge, alors que tous les voxels en dessous du seuil minimal de visualisation sont représentés par une couleur bleue.

longitudinal d'un patient atteint de NMO. Considérer un voisinage spatial donne en moyenne des valeurs de variance plus grandes et réparties plus uniformément sur l'image, avec des transitions entre tissus caractérisées par une augmentation des valeurs de variance (cf. figure 11.8-a,b). Bien que les populations de tenseurs générées par la méthode de BL et de BS diffèrent en forme et en orientation, les moyennes de tenseur estimées sont assez similaires (cf. figures 11.9 et 11.10).

Il est naturel de se demander pourquoi la méthode de BL fonctionne de manière satisfaisante en SEP (figure 11.5), alors que cette même méthode ne semble pas donner de résultats satisfaisants en NMO. Les populations de tenseurs et les moyennes associées ont été calculées à l'intersection des curseurs rouges de visualisation de la figure 11.5 et les résultats obtenus sont présentés en figure 11.10. Les différences entre les deux examens à

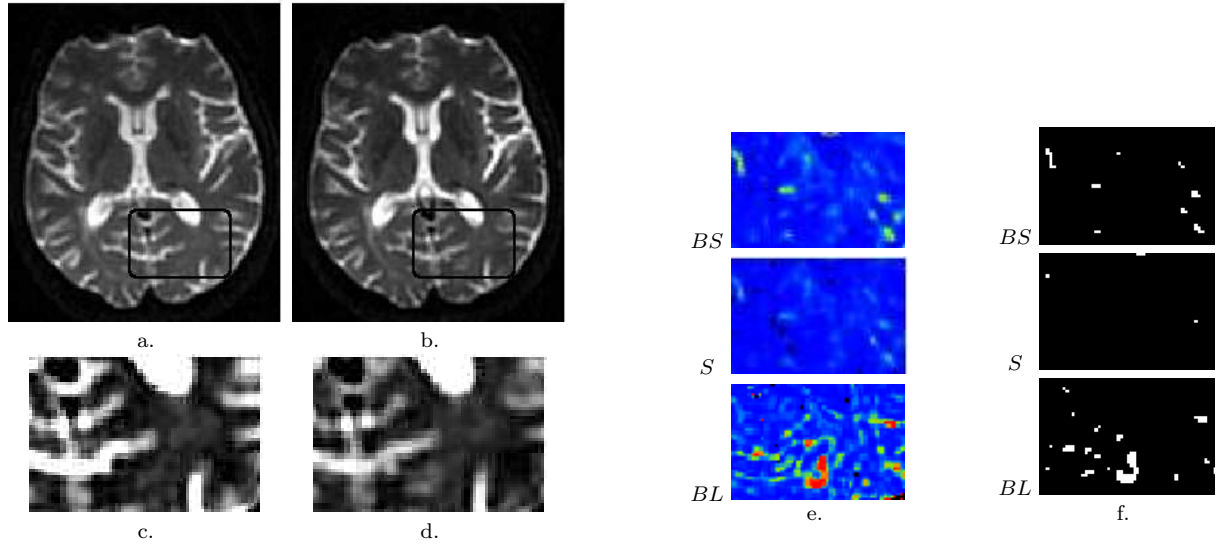


FIG. 11.6 – Exemple d’une lésion de SEP uniquement détectée avec la méthode de BL : a. images pondérée en T_2 et acquise au temps t_1 , b. images pondérée en T_2 et acquise au temps t_2 , c. zoom sur la ROI de la figure a., d. zoom sur la ROI de la figure b., e. test statistique sur les valeurs propres obtenu pour les trois méthodes de rééchantillonnage proposées et f. cartes de détection seuillées après correction FDR ($p_{FDR}=0,001$).

cette emplacement sont plus importantes (*i.e.*, une augmentation des trois valeurs propres de (44,9; 56,9; 155,0)%) alors qu’en NMO, nous avons une augmentation des trois valeurs propres de (10,6; 26,8; 43,1)%, figure 11.9). Ceci étant, on comprend mieux pourquoi la méthode de BL donne de bon résultats en SEP mais pas en NMO. A titre d’information, le manque de sensibilité de la méthode BL, observé sur ce patient, a aussi été observé sur toute la base de données NMO.

11.2.2 Base de données NMO

La base de données utilisée contient 11 patients et a été présentée en section 10.2. L’étude menée consiste à mettre en relation les modifications issues d’un suivi longitudinal effectué avec la chaîne de traitement proposée, l’évolution des lésions au niveau de la moelle épinière et l’évolution de l’état physique de chaque patient. Cette étude se base sur les travaux de [15], dans lesquels des altérations de diffusion dans la substance blanche apparemment normale chez des patients NMO ont été associées à une dégénérescence axonale due à des lésions dans la moelle épinière et dans les nerfs optiques. Au vu des résultats obtenus précédemment sur le patient NMO, la méthode spatiale S est retenue pour effectuer cette étude. Elle permet d’obtenir des résultats proches de la méthode BS pour un temps de calcul significativement réduit. L’état physique du patient est évalué avec l’échelle élargie du handicap ou, en anglais, *Expanded Disability Status Scale* (EDSS) [10]. Le score EDSS varie de 0 (examen neurologique normal) à 10 (mort). La tableau 11.3 résume les informations des 11 patients de la

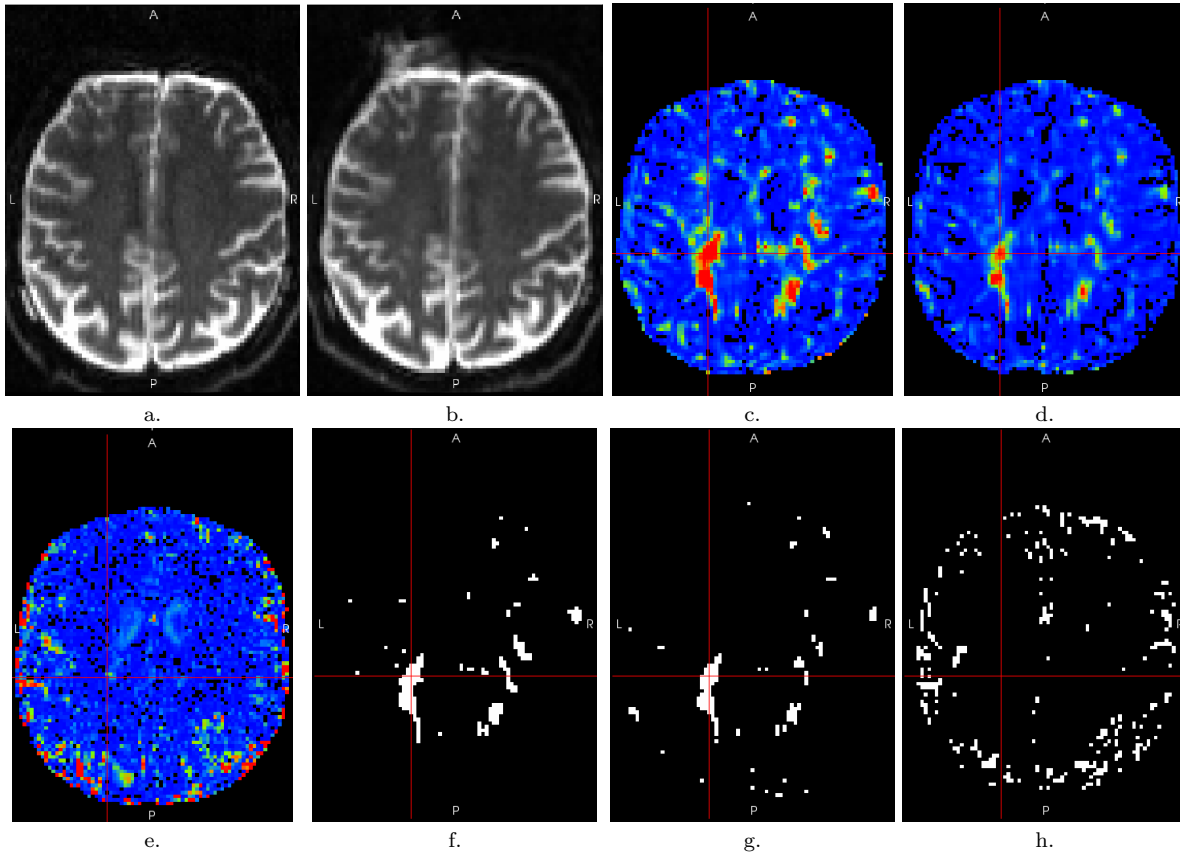


FIG. 11.7 – Cartes de détection obtenues lors du suivi longitudinal d'un patient atteint de NMO (le patient P_2), a. images pondérée en T_2 et acquise au temps t_1 , b. images pondérée en T_2 et acquise au temps t_2 , méthodes c. de BS, d. S et e. de BL. Les cartes seuillées ont été obtenues après une correction FDR ($p_{FDR}=0,001$) pour les méthodes de BS, S et de BL. Les résultats sont présentés dans les images f., g. et h., dans cet ordre.

base de données NMO et les résultats associés. Ce tableau présente l'évolution du score EDSS ($\Delta EDSS$), l'intervalle de temps entre les deux examens (Δt), l'évolution de l'étendue spatiale des lésions au niveau de la moelle épinière et les changements détectés en ITD. Une analyse visuelle des résultats par un neurologue révèle que les changements apparaissent de façon récurrente dans certaines régions chez les différents patients : le corps calleux, le sillon de Rolando et le cervelet. Pour cette raison, seules ces trois régions sont présentées dans la tableau synthétique 11.3.

Tout d'abord, une analyse détaillée des patients P_1 et P_2 est réalisée. Lors de l'étude du patient P_1 , deux régions sont détectées, à savoir, les parties antérieure et postérieure du corps calleux (cf. figure 11.11). Dans ces régions, des modifications des propriétés de diffusion ont été observées lors de la comparaison de patients NMO avec des contrôles sains [9]. De plus, les modifications des propriétés de diffusion dans le corps calleux ont été corrélées avec des troubles cognitifs (troubles de l'apprentissage et de la mémoire, diminution de la vitesse de traitement de l'information et troubles d'attention [9]). Ainsi, les modifications semblent en

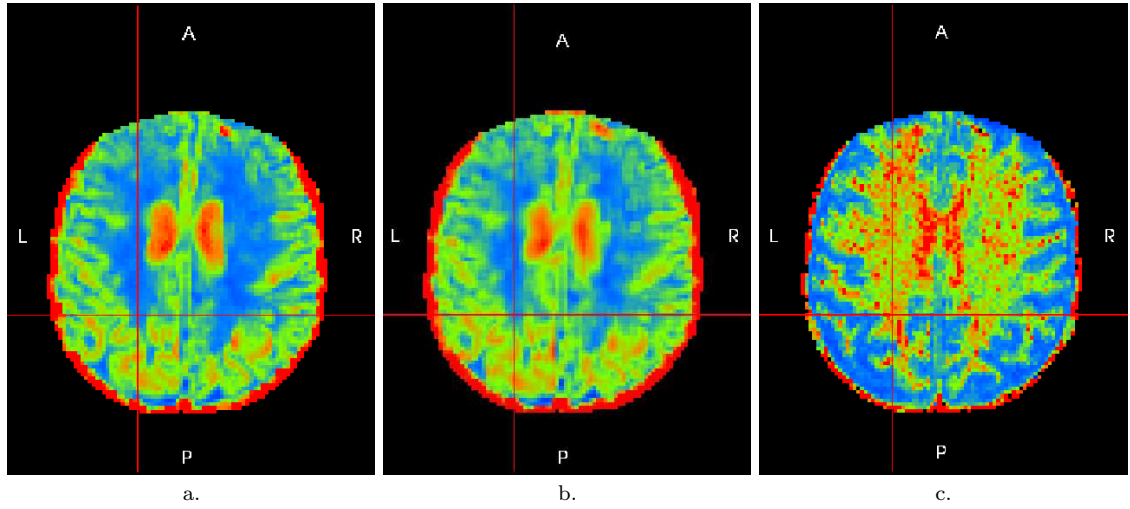


FIG. 11.8 – Cartes de variance obtenues lors du suivi longitudinal d'un patient atteint de NMO (patient P_2) avec les méthodes a. de BS (fenêtrage d'intensité : $0 - 0,3$), b. S (fenêtrage d'intensité : $0 - 0,6$), et c. de BL (fenêtrage d'intensité : $0 - 0,08$).

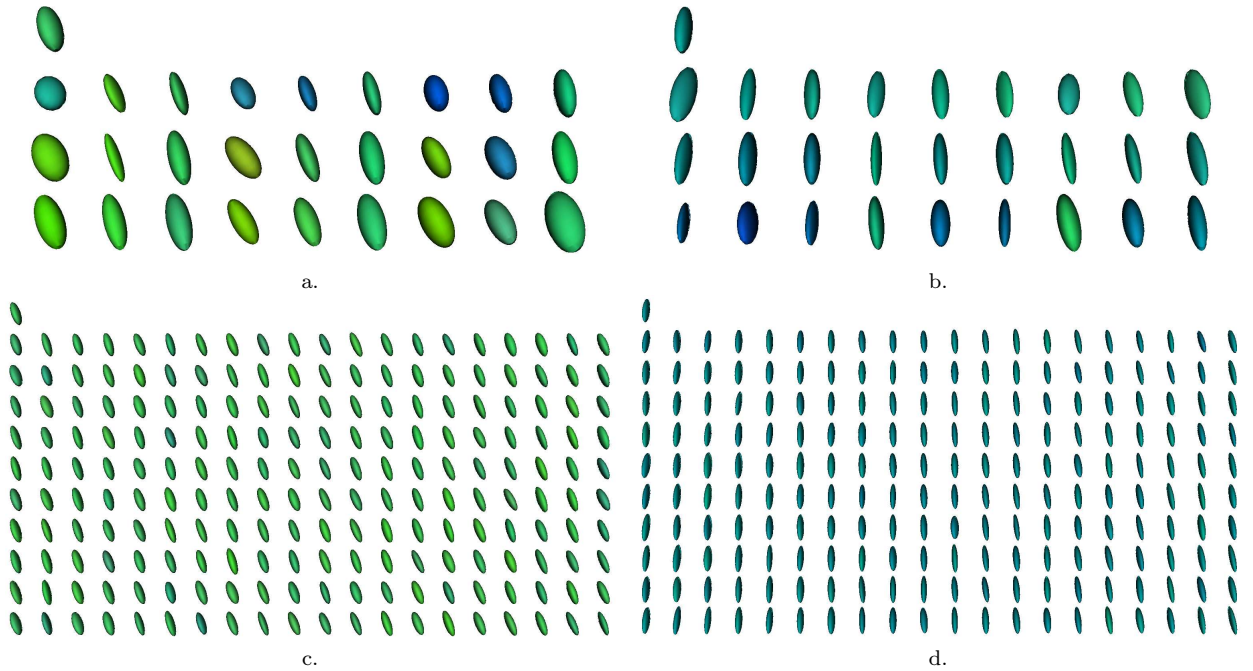


FIG. 11.9 – Population de tenseurs et tenseur moyen associé (tenseur en haut à gauche) obtenus lors du suivi longitudinal d'un patient atteint de NMO (patient P_2) avec les méthodes a. S sur le premier examen ($N = 27$), b. S sur le deuxième examen, c. de BL sur le premier examen ($N = 200$) et d. de BL sur le deuxième examen. Ces tenseurs proviennent du processus de rééchantillonnage calculé à l'intersection des curseurs rouges de visualisation de la figure 11.7.

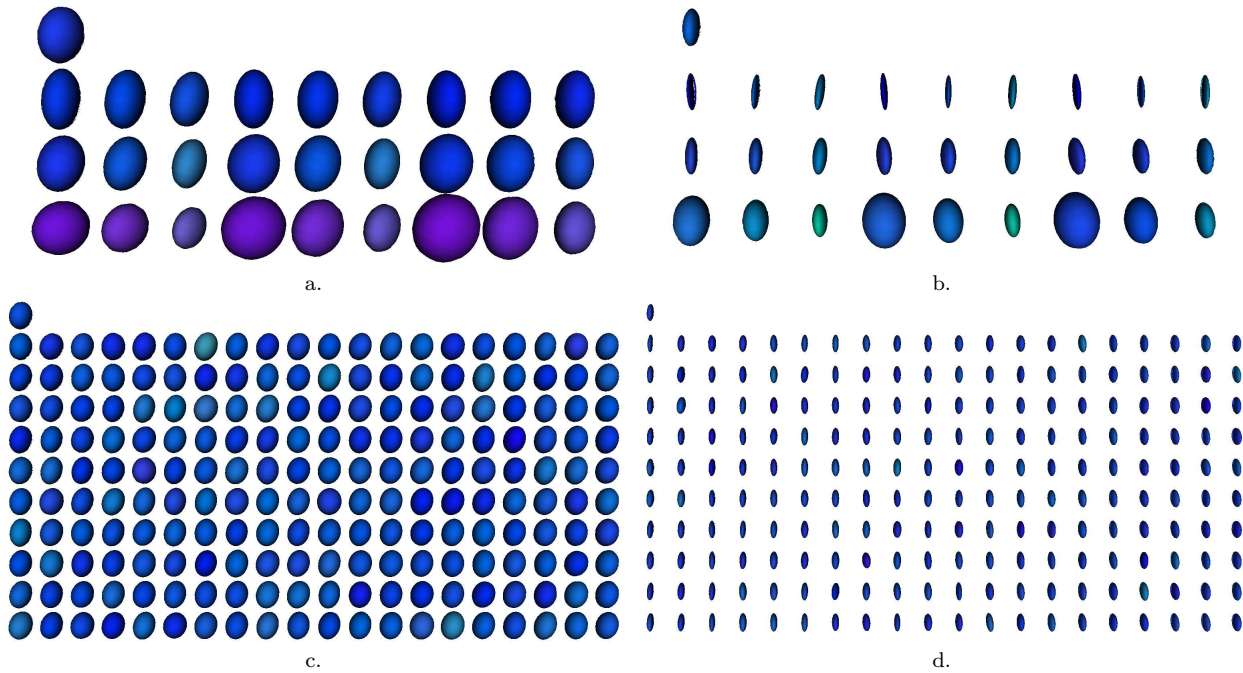


FIG. 11.10 – Population de tenseurs et tenseur moyen associé (tenseur en haut à gauche) obtenus lors du suivi longitudinal d'un patient atteint de SEP avec les méthodes a. S sur le premier examen ($N = 27$), b. S sur le deuxième examen, c. de BL sur le premier examen ($N = 200$) et d. de BL sur le deuxième examen. Ces tenseurs proviennent du processus de rééchantillonnage calculé à l'intersection des curseurs rouges de visualisation de la figure 11.5.

patient	1	2	3	4	5	6	7	8	9	10	11
anticorps aquaporin-4	-	-	-	-	-	-	+	-	+	+	+
$\Delta t(\text{mois})$	18	41	17	10	13	19	14	22	20	21	11
$\Delta EDSS$	0,5	2	3,5	0	-1,5	3	0,5	-1	0	0	1
IRM standard											
extension spatiale des lésions de la moelle épinière au temps t_1 (en nombres de segments vertébraux)	2	5	-	-	-	1	2	3	5	-	3
extension spatiale des lésions de la moelle épinière au temps t_2 (en nombres de segments vertébraux)	2	8	3	-	-	2	2	4	7	-	3
Changements détectés en ITD											
sillon de Rolando	G/D	G/D	G/D	-	-	G	G/D	-	G/D	-	G/D
corps calleux	A/P	-	A	-	-	-	-	-	-	-	-
cervelet	-	-	-	G	G/D	D	G/D	G/D	-	G/D	-

TAB. 11.3 – Résultats obtenus sur la base de données NMO. Pour chaque patient, le tableau renseigne le statut d'anticorps aquaporine 4, l'évolution du score EDSS ($\Delta EDSS$), l'intervalle de temps entre les deux examens (Δt), l'évolution de l'étendue spatiale des lésions au niveau de la moelle épinière et les changements détectés en ITD (G : gauche / D : droit, A : antérieur / P : postérieur).

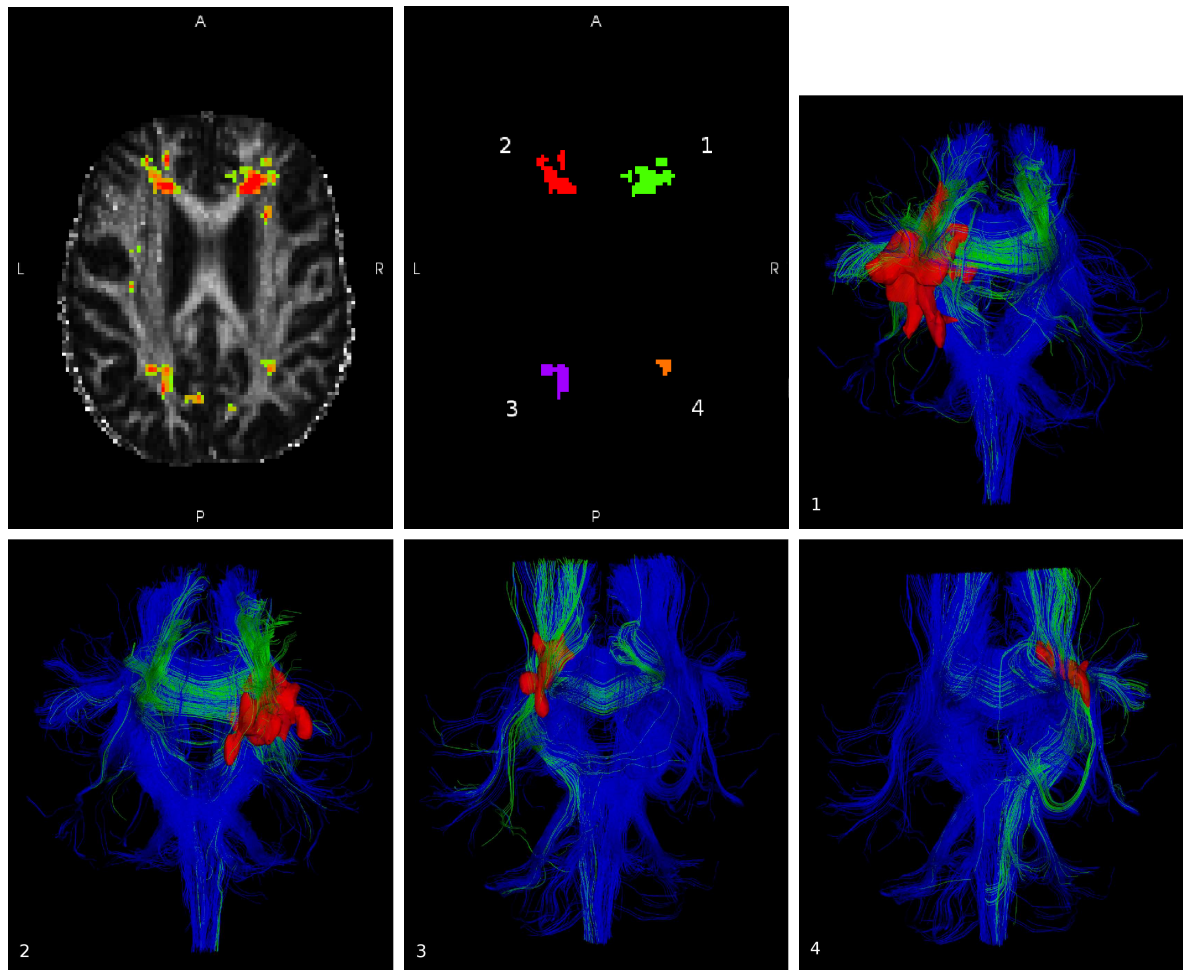


FIG. 11.11 – Cartes de détection obtenues lors du suivi longitudinal du patient NMO P_1 après correction FDR ($p_{FDR}=0,001$) et visualisation des fibres de la substance blanche intersectant avec les quatre plus grands changements (annotés de 1 à 4 sur la figure).

accord avec la dégradation des fonctions cognitives et comportementales du patient P_1 .

Concernant le patient P_2 , deux modifications bilatérales autour du sillon de Rolando sont détectées. L'outil de visualisation développé permet de colorer en vert les fibres qui passent par ces modifications (cf. figure 11.12). Ces fibres appartiennent au faisceau pyramidal. Les modifications détectées peuvent alors être attribuées aux altérations de diffusion dans la substance blanche apparemment normale liées à la dégénérescence axonale induite par des lésions dans la moelle épinière (augmentation de l'extension spatiale des lésions de 5 à 8 segments vertébraux) [15]. Cette conclusion concorde avec la perte de mobilité importante du patient entre les deux examens. En effet, le score EDSS du patient P_2 est passé de 6,5 au temps t_1 (capable de marcher avec une assistance bilatérale comme des cannes) à 8,5 au temps t_2 (contraint à rester au lit la plupart de la journée, pas capable de marcher).

Une analyse complète de la base de données NMO (cf. tableau 11.3) montre qu'il existe une

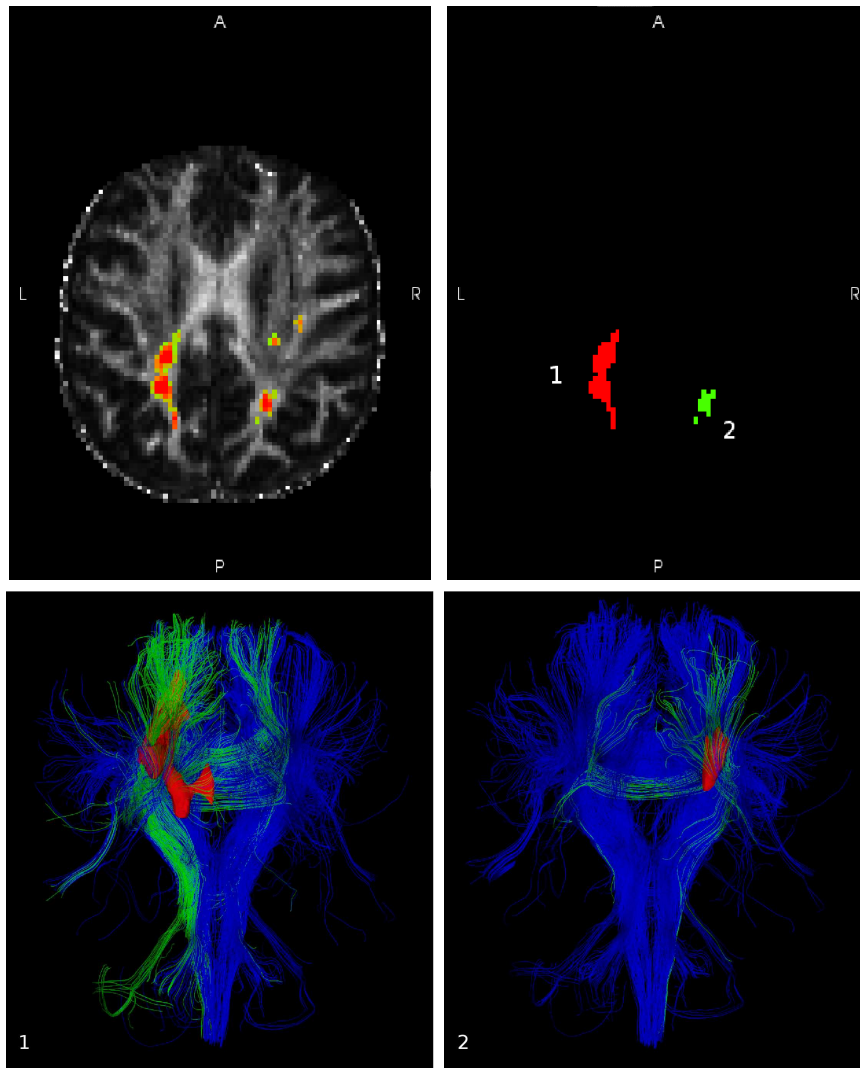


FIG. 11.12 – Cartes de détection obtenues lors du suivi longitudinal du patient NMO P_2 après correction FDR ($p_{FDR}=0,001$) et visualisation des fibres de la substance blanche intersectant avec les deux plus grands changements (annotés de 1 à 2 sur la figure).

certainne symétrie entre hémisphères dans les cartes de détection (changements bilatéraux). La méthode de visualisation proposée permet de montrer que le faisceau pyramidal intersecte les modifications détectées au niveau du sillon de Rolando et au niveau du corps calleux. Il se trouve que les sept patients présentant des évolutions au niveau du sillon de Rolando et/ou du corps calleux ont également des lésions au niveau de la moelle épinière. Trois d'entre eux présentent une baisse significative du score EDSS (≥ 2), associée à une augmentation de l'extension spatiale des lésions au niveau de la moelle épinière (patients P_2 , P_3 et P_6). Ces observations sont en accord avec les conclusions de [15], où il a été montré que des lésions de la moelle épinière induisent des altérations le long du faisceau pyramidal. Étonnamment, les patients dont les scores EDSS diminuent ne présentent aucune modification (ni au niveau du sillon de Rolando, ni au niveau du corps calleux : patients P_5 et P_8). Nous pensons que les

fonctions motrices ont été restaurées par un phénomène de compensation plutôt que par des phénomènes de remyélinisation qui auraient dû induire des modifications de la diffusion. Une exception a été trouvée pour le patient P_9 . Le score EDSS de ce patient ne change pas malgré l'évolution de lésions dans la moelle épinière et l'apparition de changements bilatéraux près du sillon de Rolando. C'est peut être un signe avant coureur de l'évolution clinique de la pathologie. Une corrélation avec le score EDSS lors du prochain examen permettra de confirmer ou d'infirmer cette hypothèse.

Finalement, il est intéressant de remarquer que différentes parties du cervelet ont été détectées lors de cette étude, alors qu'aucun patient ne souffre de syndrome cérébelleux. Ces changements pourraient apporter une information supplémentaire puisque le cervelet est connu pour être un lieu de haute expression de l'aquaporine 4. Il faut toutefois rester prudent dans l'interprétation de ces changements. En effet, ils pourraient tout aussi bien provenir d'erreurs de recalage, fréquents dans la région du cervelet, ce dernier étant constitué de nombreux sillons et circonvolutions rendant la mise en correspondance difficile.

Chapitre 12

Evaluation des méthodes de détection de changements de diffusion sur une modélisation des faisceaux de fibres

Sommaire

12.1 Résultats sur données simulées	184
12.2 Application à des données cliniques	190
12.2.1 Etude détaillée de deux patients	191
12.2.2 Base de données NMO	192

LA chaîne de traitement « détection de changement sur une modélisation des faisceaux de fibres » est mise en œuvre dans le cas particulier du suivi longitudinal de patients atteints de NMO (cf. section 9.2). Rappelons que cette stratégie permet de détecter des modifications des propriétés de diffusion en s'intéressant aux valeurs propres des tenseurs de diffusion le long des principaux faisceaux de fibres de la substance blanche. La nature multivariée de ces données impose la mise en place de tests statistiques multivariés adaptés : nous proposons d'utiliser un test local (T_{local}) et un test global (T_{global}) (cf. section 9.2). Le test local repose sur une méthode de voisinage adaptatif défini sur les faisceaux de fibres pour construire des échantillons de tenseurs à comparer reflétant la variabilité en chaque point de l'image. Le test global repose sur la représentation des fibres en faisceaux pour construire des échantillons de différences de tenseurs le long de ce modèle. Avant d'exploiter les données des patients, la méthode est évaluée sur données simulées. Dans un premier temps, des simulations sont menées pour comparer le test local, le test global et la méthode spatiale (S) validée dans le chapitre 11. Nous verrons que la distribution multinormale matricielle pour le logarithme des matrices symétriques définies positives rend bien compte des données de diffusion lorsque les populations de tenseurs sont contruites sur une modélisation des faisceaux de fibres de la substance blanche. Ensuite, une étude de deux patients NMO permet de mettre en avant les avantages des stratégies proposées dans ce chapitre, comparées à la méthode S. Finalement, l'ensemble de la base de patients NMO est traitée avec la méthode locale.

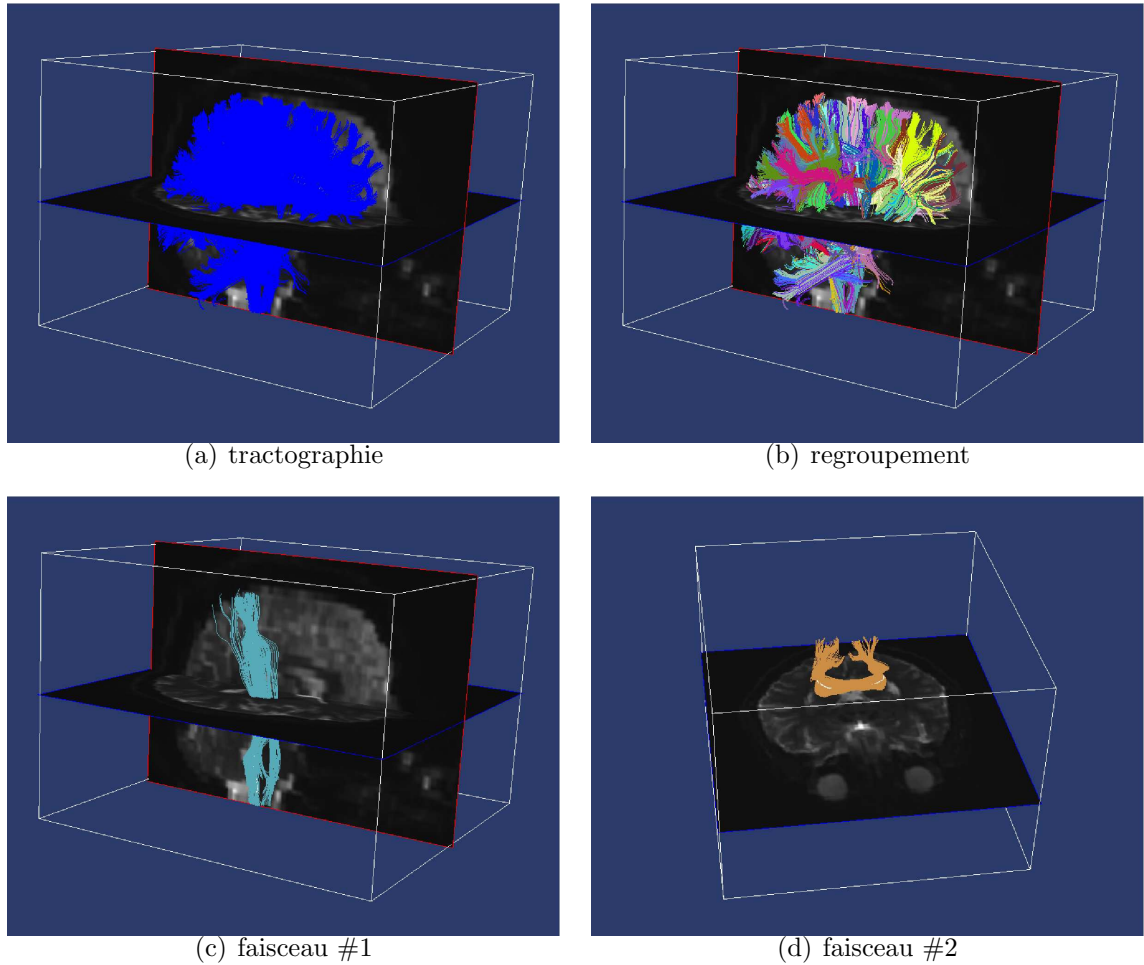


FIG. 12.1 – Représentations des fibres sélectionnées pour les simulations : (a) tractographie obtenue avec Slicer, (b) résultat de l'étape de regroupement des fibres en faisceaux (pour les paramètres, voir le texte), (c) faisceau #1 le long du faisceau pyramidal, et (d) faisceau #2 situé au niveau du corps calleux postérieur.

12.1 Résultats sur données simulées

Pour les raisons présentées dans la section 11.1.2, les méthodes sont dans un premier temps validées sur données simulées. Ces données de validation sont obtenues en augmentant la diffusion radiale le long d'un faisceau de fibres préalablement sélectionné (cf. figure 12.1). Une telle modification est caractéristique des changements qui peuvent avoir lieu en NMO. En effet, une augmentation ou une diminution de la diffusion radiale reflète un processus de démyélinisation ou de remyélinisation [8].

Soient deux acquisitions successives d'un sujet sain sur un scanner IRM Siemens de 3 Tesla. La séquence de diffusion comporte 30 directions de gradient pour une b-valeur de 1000 s/mm^2 . Les acquisitions ont été réalisées le même jour (le sujet est ressorti du scanner entre les deux acquisitions). Comme en routine clinique, aucune précaution particulière

Simulations et études de cas cliniques

Détection de changements de diffusion sur une modélisation des faisceaux de fibres

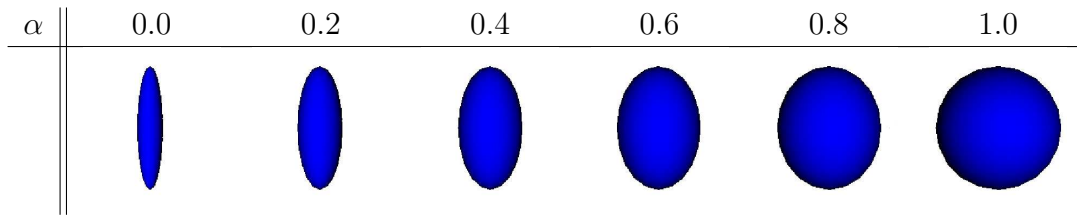


FIG. 12.2 – Modification radiale d’un tenseur avec comme valeurs propres $(\lambda_1, \lambda_2, \lambda_3) = (1, 0; 0, 2; 0, 2)$. Une telle modification est simulée en appliquant l’équation suivante : $\lambda'_i(\mathbf{s}) = \lambda_i(\mathbf{s}) + \alpha(\lambda_1(\mathbf{s}) - \lambda_i(\mathbf{s}))$, $i=(2,3)$ où $\alpha \in [0; 1]$ contrôle le degré de la modification.

n’a été prise concernant la position de la tête du patient dans l’IRM entre les deux acquisitions. Un recalage rigide rétrospectif des deux séquences a permis d’évaluer le mouvement du patient entre les deux acquisitions, soit une translation de $[-0,292; -4,378; 5,899]$ mm et une rotation autour des axes x , y , et z de $[2,596; 0,628; 0,226]$ degrés. Les images pondérées en diffusion sont chacune de taille $106 \times 106 \times 40$ et la résolution spatiale est de $1,8 \times 1,8 \times 3,5$ mm^3 .

La tractographie est réalisée avec Slicer¹ (module LabelSeeding : un seuil sur la FA de 0,2, sur la courbure de 0,8, une taille minimale de fibre de 50 mm et un pas de 0,5 mm), figure 12.1-a. Elle est constituée de 11 737 courbes 3-D (ou fibres) qui modélisent les voies de communication de la substance blanche. Le regroupement des fibres en faisceaux se fait avec la méthode d’agglomération paramétrée avec un seuil de fusion $d_f = 5$ mm (cf. section 8.3). En sortie de cet algorithme, 237 faisceaux de fibres ont été créés (cf. figure 12.1-b). Une région d’intérêt (ROI) est obtenue en discrétisant un faisceau d’intérêt. A l’intérieur de ce faisceau, la diffusion est modifiée en simulant une modification radiale dans un des deux examens avec la relation suivante :

$$\lambda'_i(\mathbf{s}) = \lambda_i(\mathbf{s}) + \alpha(\lambda_1(\mathbf{s}) - \lambda_i(\mathbf{s})), i=(2,3) \quad (12.1)$$

où $\lambda_i(\mathbf{s})$ et $\lambda'_i(\mathbf{s})$ représentent les i ème valeurs propres originales et modifiées à l’emplacement $\mathbf{s}$. $\alpha \in [0, 1]$ est un facteur multiplicatif qui permet de contrôler le degré de la modification. La figure 12.2 illustre l’influence de ce paramètre lors d’une modification radiale. Il est intéressant de noter qu’avec la relation proposée en (12.1), l’ordre de chaque valeur propre est préservé.

Les résultats obtenus avec le test local sont comparés avec les résultats de la méthode de détection de changements à l’échelle du voxel (S). Les seules différences entre ces deux approches sont la technique d’échantillonnage mise en œuvre et la géométrie du voisinage considéré. En effet, le test S est mené à l’échelle du voxel (sans calcul des fibres) sur un voisinage cubique fixé, alors que le test local est conduit en chaque point des fibres (qui ne sont pas nécessairement sur la grille de l’image) en utilisant un voisinage adaptatif (cf. figures 12.3 et 12.4). Le voisinage adaptatif n’est pas nécessairement symétrique. Tout d’abord, la validité de l’hypothèse de loi multinormale est explorée pour les méthodes S, locale et globale (cf. figure 12.5). A un emplacement choisi, les histogrammes d’un élément diagonal et

¹<http://www.slicer.org/>

Simulations et études de cas cliniques

Détection de changements de diffusion sur une modélisation des faisceaux de fibres

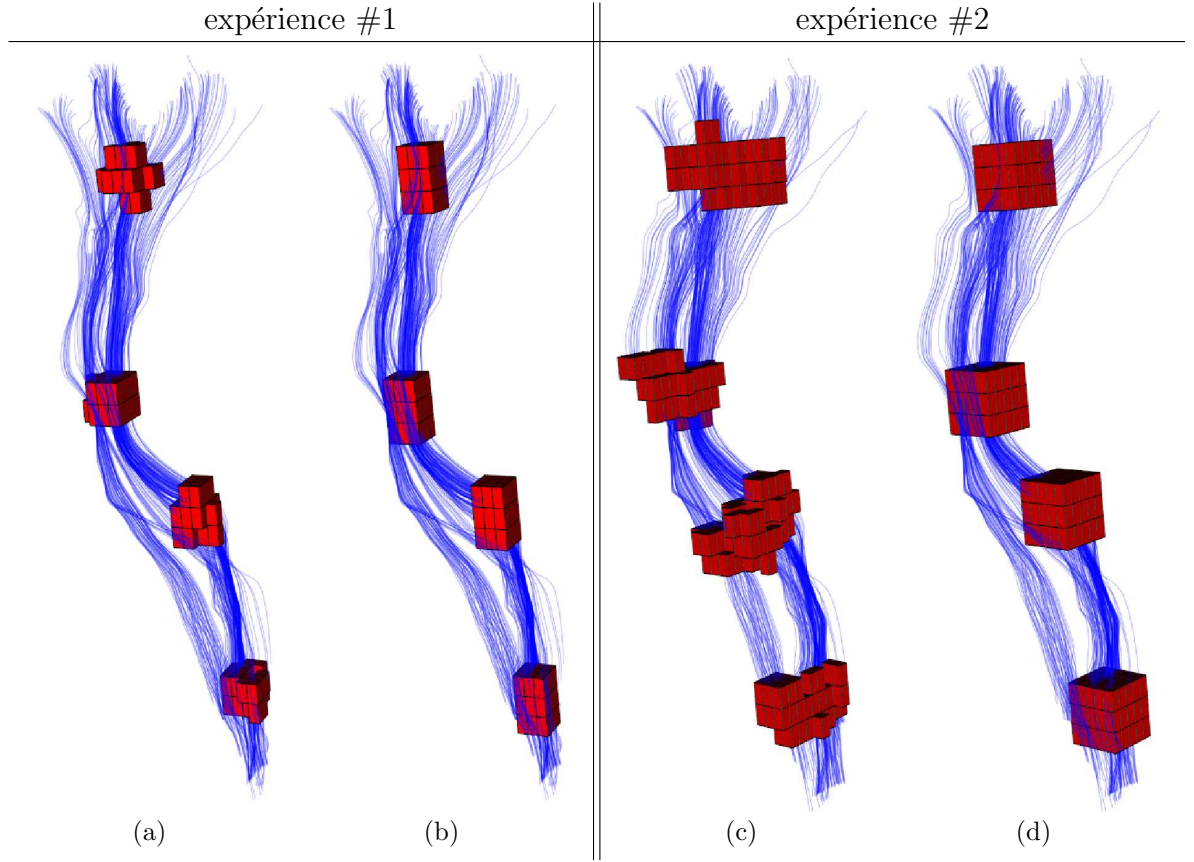
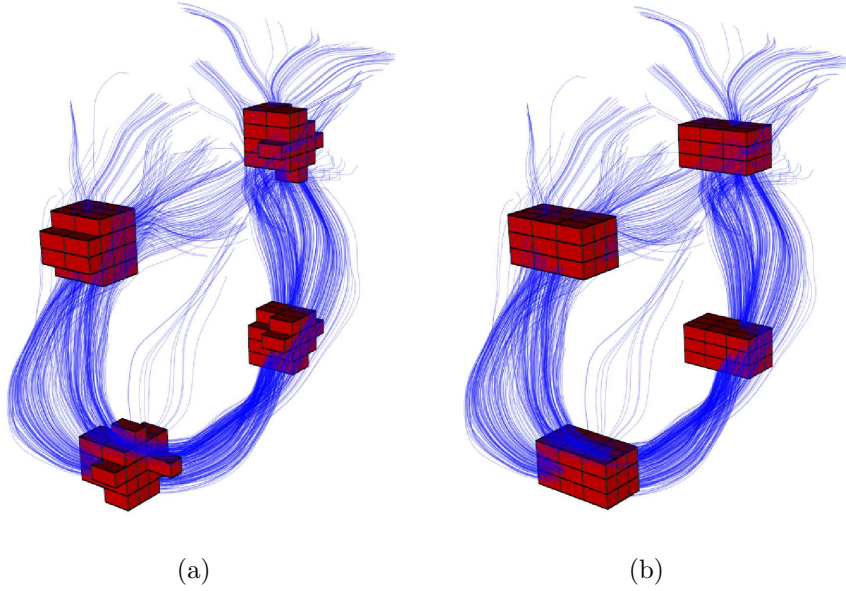


FIG. 12.3 – Géométrie du voisinage spatial obtenue en quatre échantillons du faisceau #1 (cf. figure 12.1-c) pour la méthode spatiale (S) et locale. Dans l'expérience #1 (b) un voisinage fixe $3 \times 3 \times 3$ est utilisé pour la stratégie S et (a) $N = 27$ points sont considérés pour générer le voisinage adaptatif de la méthode locale. Dans l'expérience #2 (d) un voisinage fixe $5 \times 5 \times 3$ est utilisé pour la stratégie S et (c) $N = 75$ points sont considérés pour générer le voisinage adaptatif de la méthode locale.

d'un élément anti-diagonal sont tracés. On remarque immédiatement que les méthodes locale et spatiale sont celles qui vérifient le mieux l'hypothèse gaussienne. Ensuite, pour être capable de comparer la méthode S avec la méthode locale, une projection P permet de passer du repère fibre (*i.e.*, domaine continu) vers le repère usuel (i, j, k) (*i.e.*, grille discrète de l'image). Cette étape est réalisée en utilisant une interpolation au plus proche voisin. Ainsi, lorsque plusieurs scores sont associés à un même voxel, le maximum est retenu. Le choix de projeter la carte de détection calculée sur chaque représentant des faisceaux vers le repère (i, j, k) peut pénaliser ces approches. Le critère utilisé pour comparer ces méthodes est l'aire sous la courbe COR.

Deux faisceaux ont été sélectionnés pour introduire une modification : un le long du faisceau pyramidal (faisceau #1, figure 12.1-b) et un dans la partie postérieure du corps calleux (faisceau #2, figure 12.1-c). Les tableaux 12.1-a,b et 12.2-a,b présentent les résultats pour le faisceau #1 et #2, dans cet ordre. Le test multivarié basé sur le modèle des fibres de la

expérience #1



expérience #2

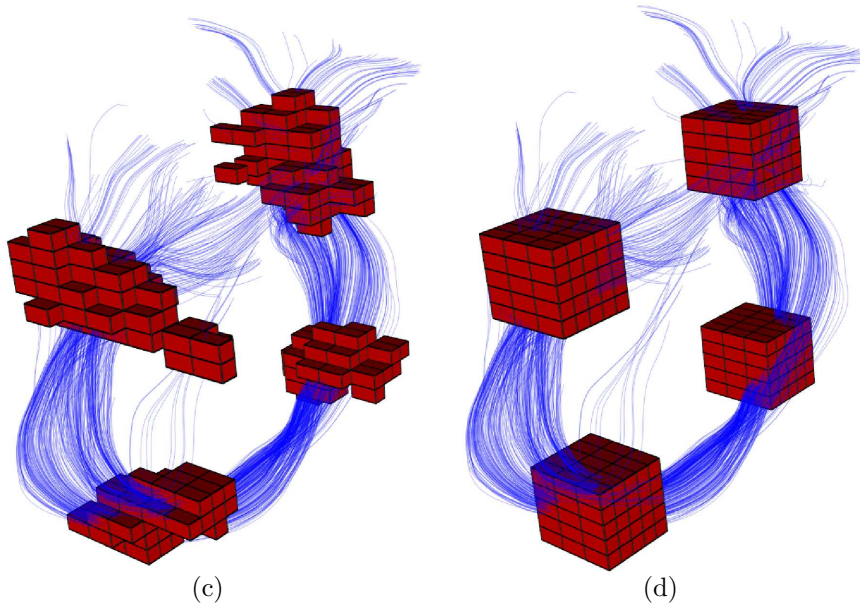


FIG. 12.4 – Géométrie du voisinage spatial obtenue en quatre échantillons du faisceau #2 (cf. figure 12.1-d) pour la méthode spatiale (S) et locale. Dans l'expérience #1 (b) un voisinage fixe $3 \times 3 \times 3$ est utilisé pour la stratégie S et (a) $N = 27$ points sont considérés pour générer le voisinage adaptatif de la méthode locale. Dans l'expérience #2 (d) un voisinage fixe $5 \times 5 \times 3$ est utilisé pour la stratégie S et (c) $N = 75$ points sont considérés pour générer le voisinage adaptatif de la méthode locale.

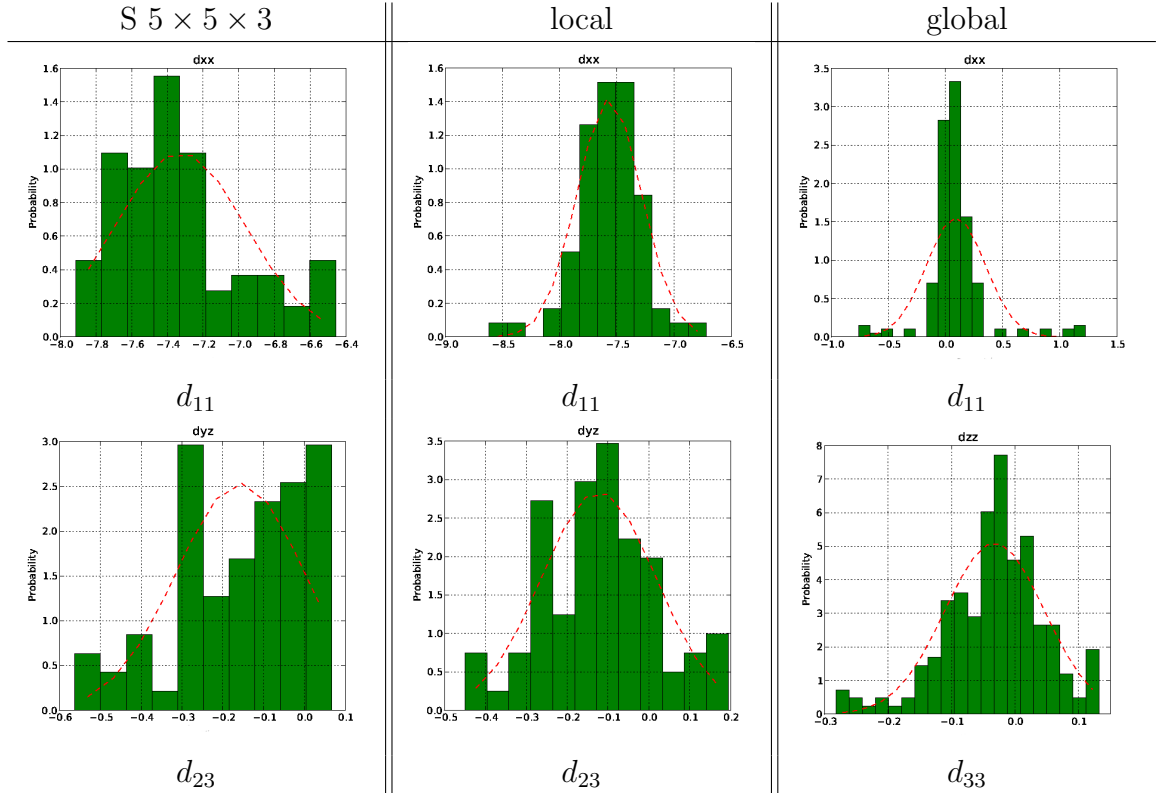


FIG. 12.5 – Distribution des éléments d_{11} et d_{23} des log-tenseurs pour le faisceau #1 avec la méthode S (un voisinage spatial $5 \times 5 \times 3$) et la méthodes locale ($N = 75$). Puisque les éléments anti-diagonaux des différences de log-tenseurs dans la méthode globale sont nuls, la distribution associée est remplacée par la distribution des éléments d_{33} . La courbe en pointillée représente une distribution gaussienne ajustée sur les données.

substance blanche est plus performant que la méthode S, mettant ainsi en évidence l'importance d'un voisinage respectant au mieux l'hypothèse de modèle constant par morceaux inhérent au test statistique. L'influence de la taille du voisinage a aussi été explorée en imposant que la taille des populations soit la même pour les deux méthodes. On remarque qu'un voisinage de petite taille donne de meilleurs résultats. Ceci peut s'expliquer par le fait que plus le voisinage augmente et moins l'hypothèse de modèle constant par morceau est vérifiée.

L'aire sous la courbe COR du test global T_{global} est également présentée pour une modification des faisceaux #1 et #2 (cf. tableaux 12.1-c et 12.2-c). Le critère utilisé pour évaluer cette méthode est l'aire sous les courbes COR, mais cette fois, directement calculée dans le système de coordonnées défini par les fibres. Dans ce cas particulier, chaque faisceau est associé à un score unique. Ainsi, un faisceau est détecté dans sa totalité ou non. Le test global donne de bien meilleurs résultats que le test local (cf. tableaux 12.1 et 12.2). La raison principale de cette performance accrue provient du type de modifications simulées. En effet, des changements globaux sur la totalité d'un faisceaux ont été simulés, ce qui avantage significativement la méthode globale. Une autre limitation due au schéma de modification proposé concerne l'étendue spatiale de la modification simulée. En effet, plusieurs faisceaux

Simulations et études de cas cliniques

Détection de changements de diffusion sur une modélisation des faisceaux de fibres

α	0,02	0,04	0,06	0,08	0,1	0,2	0,4	0,6
(a) expérience #1								
S	0,648	0,703	0,758	0,789	0,819	0,896	0,913	0,918
T_{local}	0,759	0,826	0,874	0,904	0,927	0,969	0,981	0,984
(b) expérience #2								
S	0,630	0,685	0,740	0,771	0,800	0,879	0,898	0,902
T_{local}	0,659	0,761	0,835	0,881	0,913	0,948	0,954	0,957
(c) expérience #3								
T_{global}	0,873	0,949	0,966	0,966	0,983	0,996	0,996	1,0

TAB. 12.1 – Faisceau #1 : aires sous les courbes COR pour les méthodes locale et globale, comparées avec la méthode spatiale (S) sur des modifications simulées. α est un facteur multiplicatif appliqué pour modifier la diffusion radiale. (a) dans l'expérience #1, $N = 27$ et $\tau = 3$ pour la méthode locale et un voisinage spatial $3 \times 3 \times 3$ est considéré pour la méthode S, (b) dans l'expérience #2, $N = 75$ et $\tau = 10$ pour la méthode locale et un voisinage spatial $5 \times 5 \times 3$ est considéré pour la méthode S et (c) dans l'expérience #3, les performances de la méthode globale sont présentées.

α	0,02	0,04	0,06	0,08	0,1	0,2	0,4	0,6
(a) expérience #1								
S	0,583	0,663	0,718	0,753	0,780	0,850	0,889	0,899
T_{local}	0,656	0,721	0,768	0,804	0,831	0,887	0,905	0,910
(b) expérience #2								
S	0,535	0,610	0,660	0,691	0,713	0,769	0,797	0,805
T_{local}	0,680	0,717	0,748	0,758	0,770	0,795	0,825	0,846
(c) expérience #3								
T_{global}	0,797	0,930	0,962	0,983	0,992	0,998	1,0	1,0

TAB. 12.2 – Faisceau #2 : aires sous les courbes COR pour les méthodes locale et globale, comparées avec la méthode spatiale (S) sur des modifications simulées. α est un facteur multiplicatif appliqué pour modifier la diffusion radiale. (a) dans l'expérience #1, $N = 27$ et $\tau = 3$ pour la méthode locale et un voisinage spatial $3 \times 3 \times 3$ est considéré pour la méthode S, (b) dans l'expérience #2, $N = 75$ et $\tau = 10$ pour la méthode locale et un voisinage spatial $5 \times 5 \times 3$ est considéré pour la méthode S et (c) dans l'expérience #3, les performances de la méthode globale sont présentées.

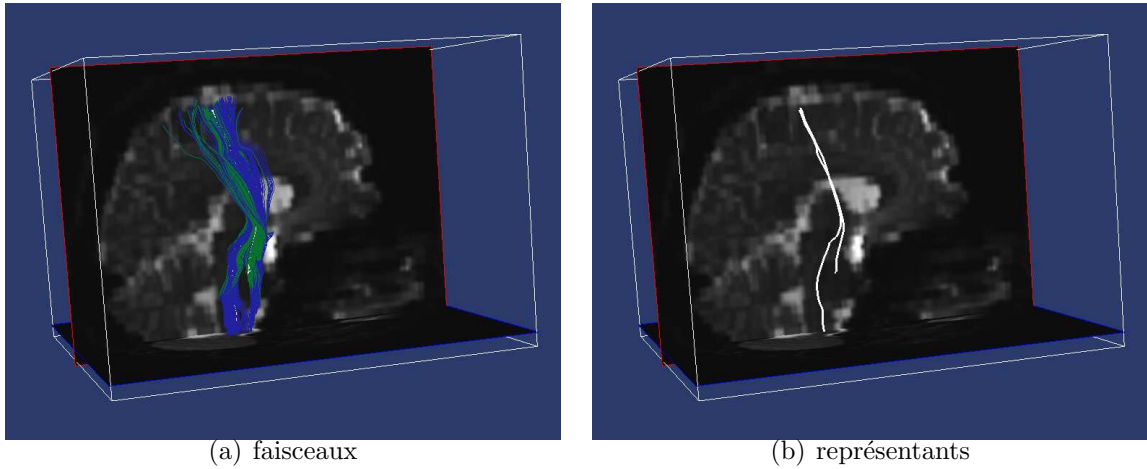


FIG. 12.6 – Chevauchement de faisceaux : (a) le faisceau #1 (bleu) et un faisceau proche (vert), et (b) les représentants associés.

peuvent être touchés par ce changement. Par exemple, l'expérience menée sur le faisceau #1 avec $\alpha = 0,4$ nous donne une aire sous la courbe COR de 0,996 (et non de 1 comme lors de l'expérience #2). Si on s'intéresse de près à cet exemple précis, on remarque qu'un faisceau a obtenu un score plus élevé d'où ce score de 0,996. Ce dernier et le faisceau #1 sont représentés en figure 12.6. On remarque que ces deux faisceaux ont un parcours dans une même région de l'espace image, avec pour seule différence que le faisceau #1 s'étend dans le tronc cérébral. Cette détection abaisse le score issu de l'analyse COR, mais est en fait une détection valide.

12.2 Application à des données cliniques

Dans cette étude, la corrélation entre les évolutions longitudinales détectées par la chaîne de traitement proposée, l'évolution des lésions dans la moelle épinière, l'évolution des fonctions cognitives et de l'état physique de chaque patient est étudiée. Cette étude se base sur les travaux de [15], dans lesquels les altérations de diffusion dans la substance blanche apparemment normale de patients NMO sont associées à une dégénérescence axonale due à l'apparition de lésions dans la moelle épinière et dans les nerfs optiques. Les troubles des fonctions cognitives sont évalués avec une batterie de tests cognitifs constituée de 14 sous-tests issus de la BRB-N (*Brief Repeatable Battery*) [5]. Les scores de chaque test sont calibrés sur une population saine. Les résultats fournis par les tests cognitifs sont le nombre de sous-tests qui ont donné des résultats inférieurs aux résultats obtenus par les 5% plus « mauvais » de la population normale. L'état physique du patient est évalué avec l'échelle EDSS [10]. Rappelons que le score EDSS varie de 0 (examen neurologique normal) à 10 (mort). Le tableau 12.3 résume les informations et les résultats pour 10 des 11 patients de la base de données NMO (le patient P_9 a été enlevé de l'étude car aucune tractographie satisfaisante n'a pu être obtenue). Ce tableau contient l'évolution du BRB-N ($\Delta BRB - N$), l'évolution du score EDSS ($\Delta EDSS$), l'intervalle de temps entre les deux examens (Δt), l'évolution de l'étendue

Simulations et études de cas cliniques

Détection de changements de diffusion sur une modélisation des faisceaux de fibres

patient	1	2	3	4	5	6	7	8	10	11
$\Delta t(\text{mois})$	18	41	17	10	13	19	14	22	21	11
$\Delta BRB - N$	1	2	0	-1	-5	0	-4	-3	-2	-1
$\Delta EDSS$	0,5	2	3,5	0	-1,5	3	0,5	-1	0	1
IRM standard										
extension spatiale des lésions de la moelle épinière au temps t_1 (en nombres de segments vertébraux)	2	5	0	0	0	1	2	3	0	3
extension spatiale des lésions de la moelle épinière au temps t_2 (en nombres de segments vertébraux)	2	8	3	0	0	2	2	4	0	3
Changements détectés en ITD										
tronc cérébral	O	O	O	-	O	O	-	O	-	-
sillon de Rolando	G	G/D	G/D	-	G	-	G	-	-	G/D
faisceaux pyramidaux	G/D	G/D	D	-	-	G/D	D	D	-	G/D
cervelet	-	O	-	O	-	-	O	O	-	O
corps calleux	A/M	A/M	-	-	-	-	A	-	A/M	-
faisceau longitudinal supérieur	D	G	D	-	G	G	G	-	-	G/D
régions préfrontales	O	-	-	-	-	O	-	-	O	-

TAB. 12.3 – Résultats obtenus sur la base de données NMO. Pour chaque patient, le tableau renseigne l'évolution du score EDSS ($\Delta EDSS$), l'intervalle de temps entre les deux examens (Δt), l'évolution de l'étendue spatiale des lésions au niveau de la moelle épinière et les changements détectés en ITD avec le test local (G : gauche / D : droit, A : antérieur / P : postérieur, M : médian, O : oui). Après une analyse en composantes connexes, les évolutions ont été codées en couleur, une couleur rouge signifiant une augmentation de FA et une couleur bleue une diminution de FA. Si deux évolutions contradictoires sont présentes dans une même région d'intérêt, aucune couleur n'est associée au score de cette région.

spatiale des lésions au niveau de la moelle épinière et les changements détectés en ITD. Des valeurs positives pour les scores $\Delta BRB - N$ et $\Delta EDSS$ correspondent à une aggravation des fonctions cognitives et motrices. Les cartes statistiques sont seuillées en utilisant une correction FDR pour tenir compte des comparaisons multiples. Des niveaux de signification de 0,001 et de 0,05 sont considérés pour le test local et le test global, respectivement. Ces seuils ont été choisis empiriquement pour détecter le plus de modifications en NMO tout en minimisant le nombre de fausses alarmes sur un sujet sain. Une analyse visuelle des données par un neurologue révèle que les modifications apparaissent dans des régions semblables à travers la base de données. Certaines de ces régions sont connues pour avoir un impact potentiel sur l'état physique des patients (évolution de l'EDSS) : le tronc cérébral, le sillon de Rolando, les faisceaux pyramidaux et le cervelet, et d'autres régions sont connues pour être liées aux fonctions cognitives (évolution du BRB-N) : le corps calleux, le faisceau longitudinal supérieur et les régions préfrontales. Ainsi, seules ces régions sont présentes dans la tableau synthétique 12.3.

12.2.1 Etude détaillée de deux patients

Tout d'abord, mettons en avant l'intérêt des stratégies locale et globale de l'approche fibre, comparé à la méthode spatiale (S). La figure 12.7 illustre l'efficacité de la méthode locale pour détecter des modifications dans le tronc cérébral. La méthode S ne détecte rien puisque

Simulations et études de cas cliniques

Détection de changements de diffusion sur une modélisation des faisceaux de fibres

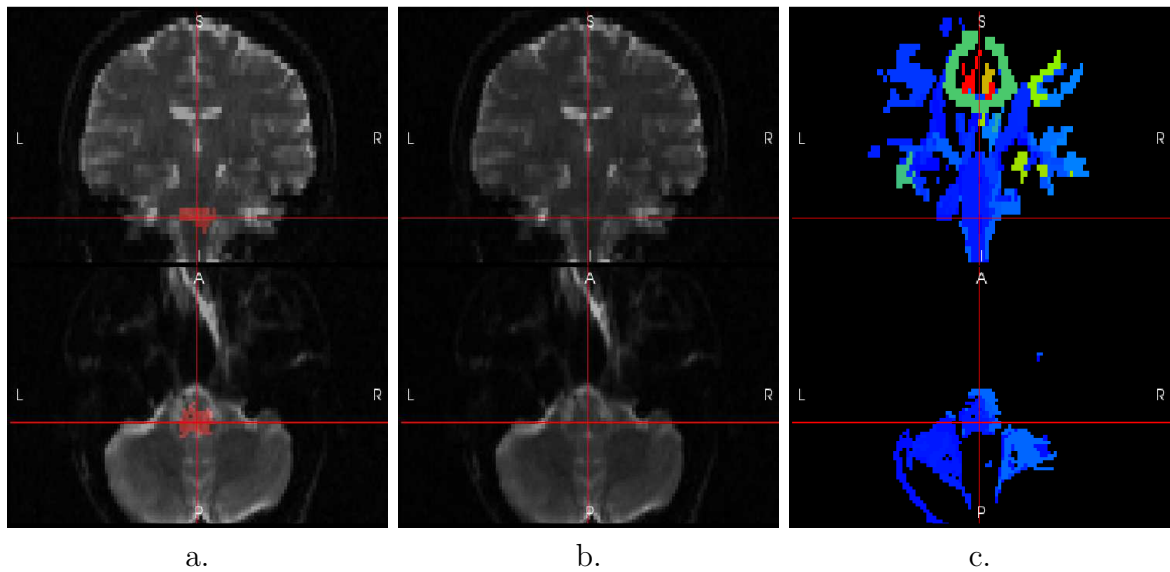


FIG. 12.7 – Cartes de détection obtenues lors du suivi longitudinal d'un patient NMO (le patient P_6) avec les méthodes a. locale, b. spatiale et c. globale. Le test local permet de détecter des changements dans le tronc cérébral, non visible avec le test S. Aucune détection significative à 0,05 n'est détectée avec le test global.

les modifications se font dans une région de petite taille, où un voisinage cubique, même de petite taille, contient un mélange de plusieurs tissus. Cette modification du tronc cérébral se situe au niveau du bulbe qui fait la jonction avec la moelle épinière. De telles évolutions (*i.e.*, des évolutions au niveau du tronc cérébral détectées uniquement avec la méthode locale) ont été observées chez 6 patients. Ces changements sont d'un grand intérêt puisqu'ils peuvent expliquer des variations de l'EDSS. Le test global ne détecte pas le/les faisceaux rattachés à cette modification car l'évolution est très localisée comparé à la taille des faisceaux dans cette région. Au regard de la carte de p-valeurs, à un seuil de signification de 0,05, aucune modification significative n'est détectée avec le test global. La figure 12.8 illustre l'efficacité du test local pour détecter des altérations de la diffusion le long des faisceaux pyramidaux. Cette méthode met en avant des modifications bilatérales au niveau des faisceaux pyramidaux, alors qu'aucune évolution spatialement cohérente n'est détectée avec le méthode S. Cette observation est d'un intérêt certain pour la NMO où une étude a montré que l'évolution de lésions au niveau de la moelle épinière induit des modifications de la diffusion le long des faisceaux pyramidaux [15]. A l'inverse du cas précédent, les modifications s'étendent sur une portion significative des faisceaux pyramidaux, qui sont alors détectés avec la méthode globale. Le test global montre également une évolution significative au niveau du corps calleux, trop subtile pour pouvoir être détectée par les deux méthodes locales. C'est en inférant sur tout le faisceau que le test global a pu détecter un tel changement.

12.2.2 Base de données NMO

Les résultats obtenus sur la base de données NMO sont présentés dans le tableau 12.3. On retrouve une certaine symétrie entre hémisphères dans les cartes de détection, en particulier

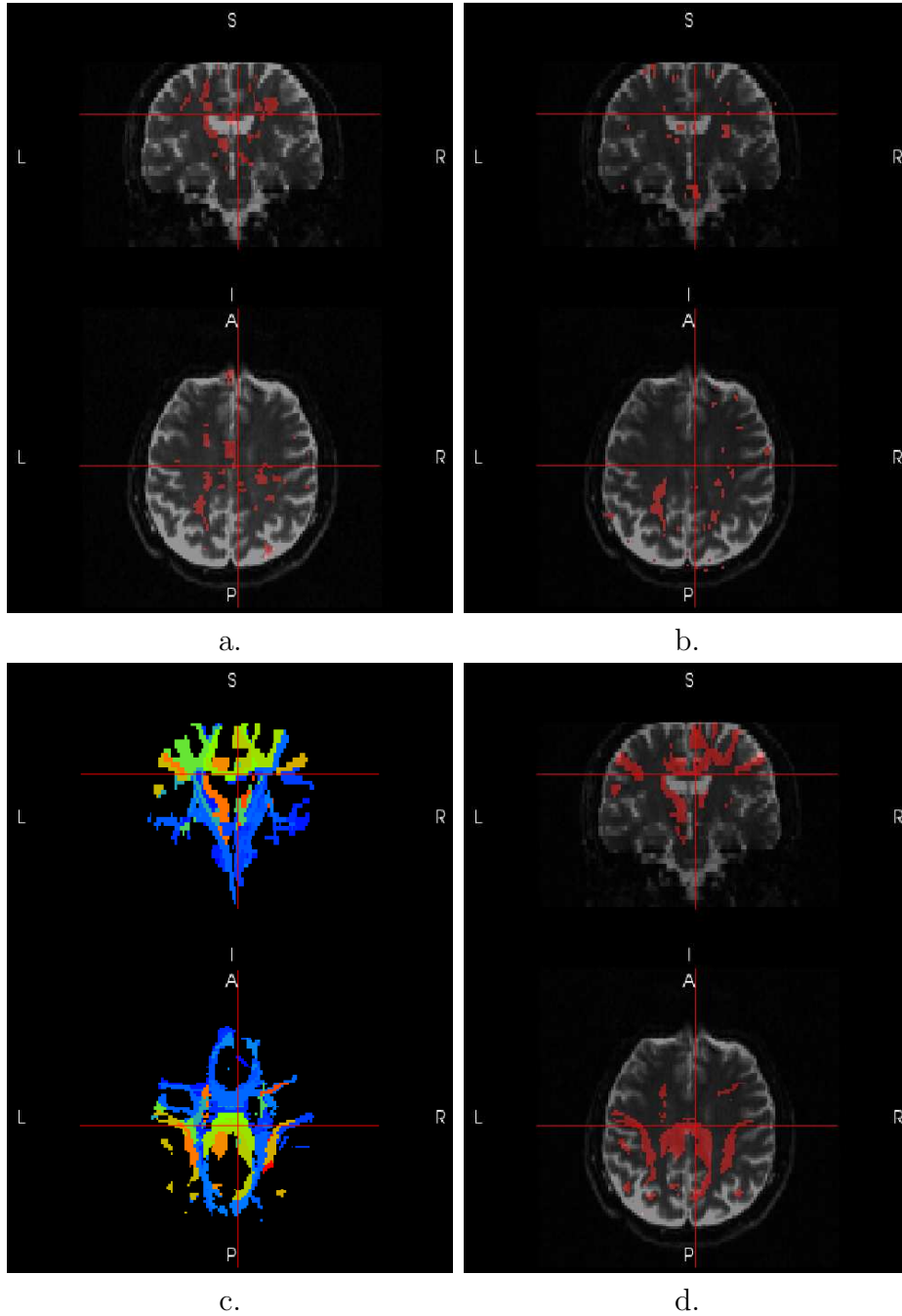


FIG. 12.8 – Cartes de détection obtenues lors du suivi longitudinal d'un patient NMO (le patient P_6) avec les méthodes a. locale, b. spatiale (S) et c. globale. La carte d. représente la carte de détection du test global à un niveau de signification 0,05. Le test local permet de mettre en évidence des altérations de la diffusion le long des faisceaux pyramidaux, alors qu'aucune détection spatialement cohérente n'est visible avec le test S.

Simulations et études de cas cliniques

Détection de changements de diffusion sur une modélisation des faisceaux de fibres

au niveau du sillon de Rolando et des faisceaux pyramidaux. Les variations de l'EDSS sont cette fois mise en regard avec des modifications au niveau du sillon de Rolando et/ou des faisceaux pyramidaux, mais également au niveau du tronc cérébral. Comme dans la section 11.2.2, sept patients présentent une telle évolution. À part le patient P_5 , tous ont des lésions dans la moelle épinière. Comme évoqué en section 11.2.2, les lésions au niveau de la moelle épinière impliquent des altérations de la diffusion le long des faisceaux pyramidaux (du tronc cérébral jusqu'au sillon de Rolando). Des lésions au niveau du corps calleux, du faisceau longitudinal supérieur et des régions préfrontales expliquent généralement une détérioration des fonctions cognitives. Les changements détectés dans ces régions sont en accord avec les changements des fonctions cognitives et comportementales des patients P_1 , P_2 , P_5 , P_7 , P_8 , P_{10} et P_{11} . En se focalisant sur le sillon de Rolando et les faisceaux pyramidaux, on remarque qu'une augmentation ou une diminution de l'EDSS est reliée à une diminution (démýélinisation) ou une augmentation (remýélinisation) de la FA. Cependant, aucune corrélation n'a été trouvée entre les scores EDSS et les évolutions détectées au niveau du tronc cérébral. Finalement, il est intéressant de remarquer que différentes parties du cervelet ont été détectées lors de cette étude, alors qu'aucun patient ne souffre de syndrome cérébelleux (ces détectations ont été discutées en section 11.2.2).

Par rapport aux résultats de la section 11.2.2 obtenus avec la méthode spatiale (S), la méthode locale fibre permet de détecter des changements au sein du faisceau pyramidal, du faisceau longitudinal supérieur et des régions préfrontales. Au niveau du sillon de Rolando 89% des modifications détectées avec la méthode locale fibre le sont également avec la méthode spatiale (S). Inversement, la méthode locale fibre permet de détecter 62% des modifications obtenues avec la méthode spatiale. Aucun recouvrement significatif des résultats n'est observé au niveau du cervelet. Cette observation semble venir confirmer le fait que les modifications observées dans cette région du cerveau pourraient provenir d'erreurs de recalage. Pour finir, au niveau du corps calleux, le test local fibre permet de détecter plus de modifications. Le corps calleux étant proche des ventricules, le voisinage adaptatif basé sur les fibres de la substance blanche semble résoudre les problèmes survenant aux interfaces entre tissus.

Conclusion partielle

Deux chaînes de traitement pour la détection de changements dans des séquences longitudinales d'IRM du tenseur de diffusion ont été proposées. Des tests statistiques multivariés, exploitant l'information contenue dans les matrices de valeurs propres des tenseurs, ont été validés sur simulations, puis menés sur des examens longitudinaux de patients atteints de SEP et de NMO. Les performances des méthodes proposées sont supérieures au GLRT. En SEP, les résultats ont montré qu'il se produisait des changements de diffusion correspondant à l'apparition de nouvelles lésions, à des modifications dans des lésions qui ne semblent pas évoluer en IRM conventionnelle et à d'autres évolutions qu'on ne peut corrélérer avec l'IRM anatomique. En NMO, les résultats ont montré qu'il se produisait des changements dans la substance blanche d'apparence normale en IRM conventionnelle. Une corrélation a été trouvée entre ces détections et l'évolution de l'état clinique des patients. Les deux chaînes de traitement proposées aboutissent aux mêmes conclusions générales, même si celle basée sur une modélisation des faisceaux de fibres de la substance blanche permet d'obtenir une détection plus riche (elle permet notamment de détecter des lésions au niveau du tronc cérébral expliquant les variations de l'EDSS). La partie suivante présente une extension des travaux sur les tenseurs d'ordre 2 et, pour cela, s'intéresse à un modèle d'ordre supérieur, le tenseur d'ordre 4.

Bibliographie

- [1] F. BLANC : *Neuromyéélite optique de DEVIC, troubles cognitifs et imagerie cérébrale par résonance magnétique*. Thèse de doctorat, Université de Strasbourg, 2010.
- [2] H. BOISGONTIER, V. NOBLET, F. HEITZ, L. RUMBACH et J. ARMSPACH : Generalized likelihood ratio tests for change detection in diffusion tensor images : Application to multiple sclerosis. *Medical Image Analysis*, 16(1):325 – 338, 2012.
- [3] La sclérose en plaques de l'enfant. <http://www.nie-enfant.com/>.
- [4] N. COLLONGUES, R. MARIGNIER, H. ZEPHIR et COLL. : Neuromyelitis optica in France : A multicenter study of 125 patients. *Neurology*, 74(9):736 – 742, 2010.
- [5] K. DUJARDIN, P. SOCKEEL, M. CABARET, J. DE SÈZE et P. VERMERSCH : BCcog-SEP : a French test battery evaluating cognitive functions in multiple sclerosis. *Revue Neurologique*, 160(1):51 – 62, 2004.
- [6] C. DUPEL-POTTIER : Critères diagnostiques de la sclérose en plaques en neuroimagerie. *Revue Neurologique*, 157(8-9):949 – 962, 2001.
- [7] T. FAWCETT : An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861 – 874, 2006.
- [8] L. HARSAN, P. POULET, B. GUIGNARD, J. STEIBEL, N. PARIZEL, P. de SOUSA, N. BOEHM, D. GRUCKER et M. GHANDOUR : Brain dysmyelination and recovery assessment by noninvasive in vivo diffusion tensor magnetic resonance imaging. *Journal of Neuroscience Research*, 83(3):392 – 402, 2006.
- [9] D. HE, Q. WU, X. CHEN, D. ZHAO, Q. GONG et H. ZHOU : Cognitive impairment and whole brain diffusion in patients with neuromyelitis optica after acute relapse. *Brain and Cognition*, 77(1):80 – 88, 2011.
- [10] J. KURTZKE : Rating neurologic impairment in multiple sclerosis : an expanded disability status scale. *Neurology*, 33(1):1444 – 1452, 1983.
- [11] V. LENNON, D. WINGERCHUK, T. KRYZER, S. PITTOCK, C. LUCCHINETTI, K. FUJIHARA, I. NAKASHIMA et B. WEINSHENKER : A serum autoantibody marker of neuromyelitis optica : distinction from multiple sclerosis. *Lancet*, 364(9451):2106 – 2112, 2004.
- [12] C. LUCCHINETTI, R. MANDLER, D. MCGAVERN, W. BRUCK, G. GLEICH, R. RANSOHOFF, C. TREBST, B. WEINSHENKER, D. WINGERCHUK, J. PARISI et H. LASSMANN :

- A role for humoral mechanisms in the pathogenesis of Devic's neuromyelitis optica. *Brain*, 125(Part 7):1450 – 1461, 2002.
- [13] Sclérose en plaques. <http://www.passeportsante.net>, 2011.
- [14] D. WINGERCHUK, V. LENNON, S. PITTOCK, C. LUCCHINETTI et B. WEINSHENKER : Revised diagnostic criteria for neuromyelitis optica. *Neurology*, 66(10):1485 – 1489, 2006.
- [15] C. YU, F. LIN, K. LI, T. JIANG, W. QIN, H. SUN et P. CHAN : Pathogenesis of normal-appearing white matter damage in neuromyelitis optica : Diffusion-tensor MR imaging. *Radiology*, 246(1):222 – 228, 2008.

PARTIE V

Vers un suivi longitudinal sur des modèles d'ordre supérieur

Introduction

LA cinquième et dernière partie décrit une extension des méthodes précédentes en s'intéressant aux modèles d'ordre supérieur, plus particulièrement le modèle d'ordre 4. Le modèle utilisé jusque là (le tenseur d'ordre 2) présente une limitation majeure lorsque la contribution de plusieurs faisceaux est observée. En effet, il ne permet pas de modéliser des profils de diffusion hétérogènes avec plusieurs orientations. Le modèle du tenseur d'ordre 4 offre une alternative en autorisant plusieurs orientations privilégiées. L'étude développée dans cette partie a pour but d'apporter de premiers éléments d'analyse sur l'apport d'une telle modélisation lors d'études longitudinales. L'exploration des modèles d'ordre 4 proposée ici est encore très préliminaire, il conviendra de poursuivre ces travaux à la fois sur le volet modélisation et expérimental.

Le premier chapitre présente les prétraitements spécifiques au tenseur d'ordre 4 permettant d'étendre les chaînes de traitement présentées dans le chapitre 9. Nous verrons que seule la stratégie de réorientation diffère significativement. Le deuxième chapitre propose une extension au tenseur d'ordre 4 des tests statistiques multivariés utilisés précédemment pour les tenseurs d'ordre 2 (cf. chapitre 9). Puisque les valeurs propres d'un tenseur d'ordre 4 sous sa forme matricielle 6×6 n'ont pas d'interprétations physique (cf. section 2.1.3), le test d'hypothèses non contraint sur la structure matricielle est utilisé (cf. section 6.2.2). Les différences ainsi mesurées peuvent provenir d'un changement des propriétés de diffusion, d'un changement d'organisation des structures sous-jacentes ou une combinaison de ces deux phénomènes. Le test non contraint sur la structure matricielle ne permet pas de dissocier l'origine des modifications observées. Avec le tenseur d'ordre 4 sous sa forme matricielle 6×6 , une décomposition spectrale ne suffit pas pour sélectionner l'information d'intérêt. Afin de quantifier des changements d'orientation, le troisième chapitre présente une mesure de similarité qui reflète l'organisation des structures anatomiques. Cette mesure est construite à partir des directions des faisceaux de fibres de la substance blanche en considérant des tenseurs d'ordre 2 ou d'ordre 4.

Ce chapitre s'appuie sur le modèle du tenseur d'ordre 4 présenté dans la section 2.1.3. Nous proposons ici un bref rappel. Un tenseur d'ordre 4 $\mathbf{T}$ contient 15 éléments distincts. Il peut être représenté par une matrice 6×6 symétrique. Rappelons qu'à partir de $\mathbf{T}$, le profil de

diffusion s'écrit :

$$d(\vec{g}) = \sum_{i,j,k \geq 0, i+j+k=4} t_{i,j,k} g_1^i g_2^j g_3^k \quad (12.2)$$

où $\vec{g} = [g_1, g_2, g_3]$ est un vecteur unité et les $t_{i,j,k}$ sont les éléments distincts de $\mathbf{T}$. A partir de l'équation de Stejskal-Tanner (1.26) (qui établit la relation entre le signal de diffusion S et le tenseur de diffusion), un tenseur d'ordre 4 est estimé en chaque voxel de l'image (cf. section 2.2). Le tenseur estimé modélise soit la diffusivité (dont les maxima ne sont pas nécessairement alignés avec l'orientation locale des fibres), soit la probabilité de déplacement des molécules d'eau représentée par la fonction de distribution des orientations (ODF) qui est représentative de l'orientation locale des fibres. Rappelons que le lien entre la base des polynômes homogènes et la base des harmoniques sphériques permet notamment de calculer rapidement une ODF angle solide (cf section 3.1.3). Les maxima d'un tenseur d'ordre 4 sont calculés en cherchant les solutions d'un système d'équations non-linéaires (stratégies de recherches analytique et différentielle) ou en cherchant les points d'un maillage de la sphère unité vérifiant le critère du maximum local (stratégie de recherche exhaustive) (cf. section 2.1.3). Finalement, différents scalaires peuvent être calculés pour caractériser un tenseur d'ordre 4, dont notamment, l'anisotropie généralisée (GA) qui permet de quantifier l'anisotropie d'un tenseur d'ordre 4 (cf. section 2.1.3).

Chapitre 13

Les prétraitements des images : la réorientation des tenseurs d'ordre 4

Sommaire

13.1 Etat de l'art des méthodes de réorientation	204
13.1.1 Réorientation de l'ODF discrétisée	204
13.1.2 Réorientation par décomposition sur une base de diracs	204
13.1.3 Réorientation rigide	205
13.1.4 Réorientation par décomposition spectrale ou par décomposition de Hilbert	205
13.2 Méthode proposée pour la réorientation des tenseurs d'ordre 4	205
13.2.1 Calcul des maxima de l'ODF	206
13.2.2 Décomposition des tenseurs d'ordre supérieur	208
13.2.3 Application à la réorientation	209
13.3 Résultats expérimentaux	210
13.3.1 Utilisation de l'approche proposée	210
13.3.2 Comparaison avec des méthodes de la littérature	212

LA comparaison inter-images basée sur une modélisation de la diffusion par des tenseurs d'ordre 4 nécessite des prétraitements des images spécifiques. Les prétraitements généraux ont été présentés dans le chapitre 7. La modélisation de la diffusion n'intervient que dans l'étape d'interpolation et de réorientation des images de tenseurs. Tous les autres prétraitements restant identiques, ils ne seront donc pas redétaillés dans ce chapitre. L'interpolation des tenseurs d'ordre 4 est identique à l'exception près qu'elle est ici directement conduite dans le domaine euclidien sur les 15 éléments du tenseur d'ordre 4 (aucune contrainte simple ne permettant d'imposer la positivité des tenseurs d'ordre 4). La méthode de réorientation, quant à elle, diffère significativement puisqu'elle doit permettre de réorienter chaque faisceau indépendamment.

Tout d'abord, ce chapitre présente différentes solutions de la littérature au problème de réorientation des tenseurs d'ordre 4. Dans la littérature, certaines stratégies permettent de

réorienter directement l'ODF discrétisée [8] ou décompose l'ODF sur une base de diracs [5]. D'autres stratégies tentent de réorienter directement un tenseur d'ordre 4, soit en appliquant une transformation rigide qui ne permet pas de réorienter chaque faisceau indépendamment [3], soit en décomposant un tenseur d'ordre 4 en plusieurs tenseurs d'ordre 2 à réorienter [11]. Dans cette dernière approche, la décomposition ne garantit pas que les tenseurs d'ordre 2 soient alignés dans la direction des fibres. Cette contrainte a été introduite dans une extension de la méthode [11] que nous avons proposée dans [6]. Elle constitue une contribution originale de cette thèse sur ce point particulier. Cette stratégie de réorientation des tenseurs d'ordre 4 ainsi que les résultats expérimentaux associés sont également présentés dans ce chapitre.

Ce chapitre étend la réorientation des tenseurs d'ordre 2 présentée en section 7.3. Les mêmes notations sont utilisées. $\mathbf{F}$ représente la matrice Jacobienne du champ de déformation H , définie en chaque voxel de l'image. Deux stratégies de réorientation ont été présentées pour les tenseurs d'ordre 2 : une réorientation rigide (méthode dite du *Finite Strain*, FS) qui ne tient compte que de la déformation contenue dans $\mathbf{F}$, et une réorientation qui tient compte également des tenseurs observés en préservant leur direction principale (méthode dite de *Preservation of the Principal Direction*, PPD).

13.1 Etat de l'art des méthodes de réorientation

13.1.1 Réorientation de l'ODF discrétisée

Dans [8], l'ODF est échantillonnée sur la sphère unité et la transformation $\mathbf{F}$ est appliquée sur chacune de ces directions. Les valeurs de l'ODF dans les directions d'échantillonnage sont normalisées. Finalement, l'ODF transformée est représentée par cet ensemble de données et doit être estimée. Dans [8], l'auteur propose d'utiliser une méthode de déconvolution sphérique.

13.1.2 Réorientation par décomposition sur une base de diracs

Dans [5], une extension du PPD est proposée. Les coefficients $\vec{c}_l$ de l'ODF à l'ordre l exprimés dans une base d'harmoniques sphériques sont décomposée sur une base de diracs $\mathbf{\Pi}$:

$$\vec{c} \approx \mathbf{\Pi} \vec{w} \quad (13.1)$$

où $\vec{w}$ est un vecteur de taille n contenant les coefficients de la décomposition et $\mathbf{\Pi} = [\delta_l(\theta_1, \phi_1), \dots, \delta_l(\theta_n, \phi_n)]$. Pendant la réorientation, les coefficients $\vec{w}$ estimés restent constants et les n directions (θ_i, ϕ_i) sont réorientées. L'ODF réorienté est obtenue en considérant la base de diracs réorientée, pondérée avec les coefficients $\vec{w}$.

13.1.3 Réorientation rigide

Dans [3], la méthode du FS pour des tenseurs d'ordre 2 est étendue pour des tenseurs d'ordre 4. Soit $\mathbf{R}$ la matrice de rotation estimée localement à partir du champ de déformation en chaque voxel de l'image. Il existe une formule analytique permettant d'appliquer une rotation à un tenseur d'ordre 4 $\mathbf{D}$ exprimé dans une base de polynômes homogènes (il est également possible d'effectuer cette réorientation sur la base des harmoniques sphériques). Cette formule est la suivante [3] :

$$\mathbf{D}_{reo}(\vec{g}) = \sum_{i,j,k,l=1}^3 d_{ijkl}(\mathbf{R}\vec{g})_i(\mathbf{R}\vec{g})_j(\mathbf{R}\vec{g})_k(\mathbf{R}\vec{g})_l \quad (13.2)$$

avec $(\mathbf{R}\vec{g})_i$ la i -ème composante du vecteur $\mathbf{R}\vec{g}$. Cette méthode est efficace pour appliquer des transformations rigides, mais est mise en échec dès lors que des transformations non-linéaires sont considérées. En effet, avec une simple rotation, il n'est pas possible de réorienter les directions des faisceaux séparément.

13.1.4 Réorientation par décomposition spectrale ou par décomposition de Hilbert

Dans [11], un tenseur ODF d'ordre 4, exprimé sur une base de polynômes homogènes, est décomposé comme une fonction de plusieurs tenseurs d'ordre 2. Cette décomposition se fait en utilisant soit la décomposition spectrale, soit la décomposition de Hilbert. Ensuite, une version modifiée du PPD, permettant de prendre en compte des tenseurs d'ordre 2 non nécessairement définis positifs, est appliquée sur chaque tenseur d'ordre 2. Ces décompositions ne reflètent pas nécessairement le substrat anatomique sous-jacent (*i.e.*, chaque tenseur d'ordre 2 n'est pas forcément aligné sur un paquet de fibres). Ainsi, la réorientation peut échouer.

13.2 Méthode proposée pour la réorientation des tenseurs d'ordre 4

Dans [6], nous avons proposé d'améliorer la stratégie de réorientation par décomposition spectrale ou de Hilbert, en introduisant une nouvelle décomposition basée sur les maxima de l'ODF. L'idée de cette décomposition est d'exprimer un tenseur ODF d'ordre 4 (noté $\mathbf{T}$), exprimé sur une base de polynômes homogènes, comme une somme de N tenseurs d'ordre 2 symétriques définis positifs et alignés sur les faisceaux principaux (notés $\mathbf{D}_i$). Ainsi, le PPD classique peut être utilisé pour réorienter chaque $\mathbf{D}_i$. Le tenseur d'ordre 4 réorienté est ainsi obtenu par recombinaison des différents tenseurs d'ordre 2 réorientés.

13.2.1 Calcul des maxima de l'ODF

La recherche des maxima d'une ODF est un problème abordé par plusieurs articles (une revue de ces articles est proposée en section 2.1.3). Nous souhaitons évaluer les performances de ces différentes méthodes. Les recherches analytique et différentielle nécessitent de résoudre des systèmes d'équations non linéaires qui n'admettent pas de solution simple et qui ont parfois plusieurs solutions. Afin d'avoir des performances optimales pour ces méthodes, nous avons proposé d'utiliser des méthodes d'optimisation globale basée sur l'analyse par intervalle (*i.e.*, une amélioration des implantations proposées dans les articles). Quant à la recherche exhaustive, elle nécessite une discrétisation régulière plus ou moins fine de la sphère unité. Nous proposons une comparaison des ces trois méthodes sur données synthétiques en évaluant leur précision et sensibilité au bruit.

Amélioration des méthodes analytique et différentielle

Lors de la recherche des points critiques avec la méthode analytique ou la méthode différentielle, et afin de trouver toutes les solutions, une optimisation par intervalle est utilisée [7]. Notons $\mathcal{D}$ le domaine de recherche. Dans la cas de la méthode analytique, il faut définir un domaine de recherche $2D$ pour trouver les solution (u, v) (cf. section 2.1.3). De la même manière, en utilisant les coordonnées sphériques, on peut définir un domaine de recherche $2D$ (θ, ϕ) pour la méthode différentielle. Dans ce cas précis, les limites du domaine de recherche sont connues : $\theta \in [0, 2\pi]$ et $\phi \in [0, \pi]$, ce qui simplifie la recherche. Ensuite, l'idée générale de l'algorithme d'optimisation est de subdiviser successivement le domaine de recherche en boîtes de taille de plus en plus petites (dans notre cas, chaque boîte est subdivisée en quatre boîtes de même taille) et de ne conserver que les boîtes susceptibles de contenir un ou des points critiques. Ce choix est fait à l'aide d'une fonction d'inclusion qui permet de définir un encadrement des valeurs que peut prendre la fonction optimisée sur une boîte donnée. Les fonctions d'inclusions des méthodes analytique et différentielle sont données en section 2.1.3. Le principe de la procédure de recherche est le suivant. Soient deux listes L_1 et L_2 . L_1 est la liste courante contenant les boîtes à traiter. L_2 est une liste résultat de petites boîtes de taille $SIZE_MIN$, chacune d'entre elles étant susceptible de contenir un zéro du problème. $SIZE_MIN$ est un paramètre permettant de contrôler la taille minimale en dessous de laquelle une boîte n'est plus subdivisée, ayant ainsi une conséquence directe sur la précision de la solution estimée (cf. l'algorithme en annexe-D).

Recherche exhaustive

Pour la recherche exhaustive, il faut tout d'abord générer un échantillonnage uniforme sur la sphère unité et ensuite extraire les voisins d'un point du maillage (cf figure 13.1). Naturellement, on pense à discrétiser uniformément en (θ, ϕ) . Néanmoins, le maillage résultant de cette discrétisation n'est pas uniforme car la densité des échantillons est plus importante au pôle qu'à l'équateur. Pour générer un échantillonnage triangulaire uniforme d'une sphère unité, un icosaèdre à 20 faces est choisi comme point de départ. Une procédure récursive de subdivision de chaque face de l'icosaèdre en quatre sous-faces de forme triangulaire est utilisée. Pour ce faire, on projette sur la sphère unité les points médians de chaque arête. Selon

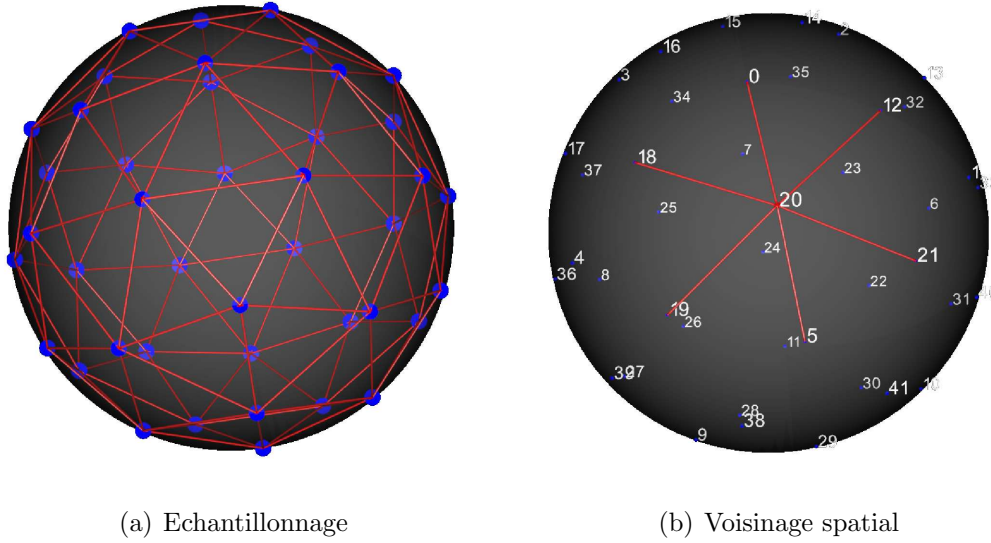
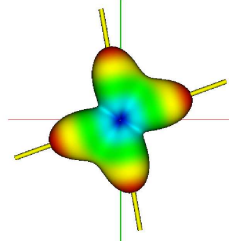


FIG. 13.1 – Recherche exhaustive : (a) échantillonnage triangulaire uniforme d'une sphère unité en utilisant une procédure de subdivision et (b) représentation des six voisins (0, 5, 12, 16, 19, 21) du point 20.

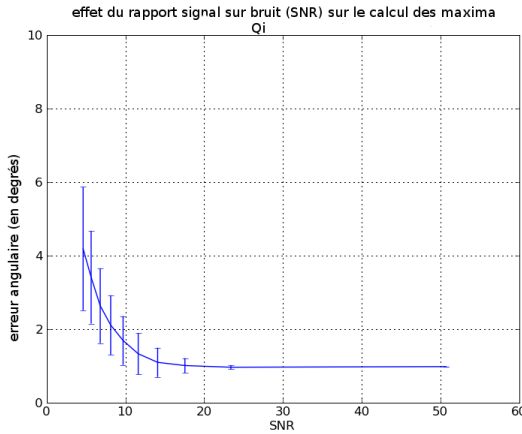
le niveau de détail souhaité, il est maintenant très facile de générer le maillage d'une sphère unité en subdivisant autant de fois que nécessaire. Avec cette procédure d'échantillonnage, chaque point du maillage a six voisins. Il suffit ensuite d'appliquer, en chaque point du maillage, le critère du maximum relatif (cf. section 2.1.3).

Comparaison des trois approches

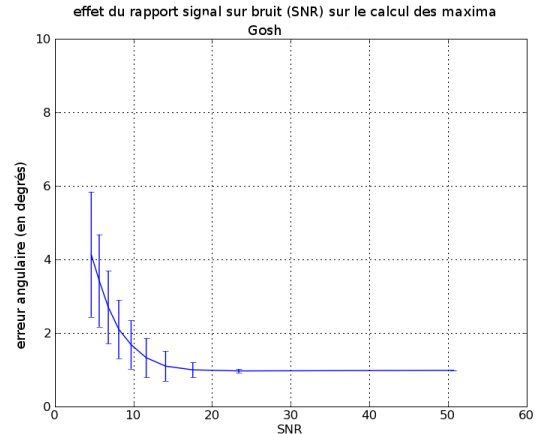
Les méthodes sont comparées sur un cas simulé (cf. figure 13.2). Pour ce faire, les maxima d'un croisement de fibre synthétique de 80° sont extraits mille fois à un certain niveau de bruit pour différentes réalisations de ce bruit. On mesure l'erreur de reconstruction des maxima estimés par rapport à la vérité terrain. Les méthodes analytique et différentielle donnent les meilleurs résultats. La recherche exhaustive donne des résultats très satisfaisants, la qualité du résultat pouvant être améliorée en augmentant le nombre d'échantillons utilisés lors de la recherche sur la sphère unité. Nous utiliserons cette méthode pour extraire les maxima d'une ODF notées $\vec{d}_i$ (avec un échantillonnage de 642 points) car elle nécessite un temps de calcul plus court et donne de bons résultats pour un rapport signal sur bruit $\geq 20dB$ (hypothèse vérifiée lors d'acquisitions EPI avec une antenne de tête multi-canaux). Notons que les maxima d'une image constituée de 8000 tenseurs sont calculés en 21,32 secondes avec la recherche exhaustive (un CPU de 2,67GHz) alors que les deux autres stratégies nécessitent un temps de calcul $\geq 117,62$ secondes (quatre CPU de 2,67GHz).



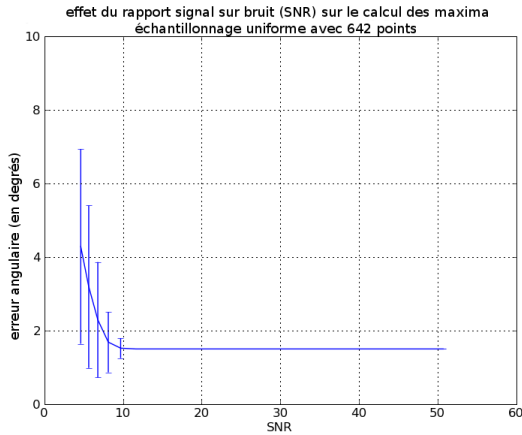
(a) fibre synthétique



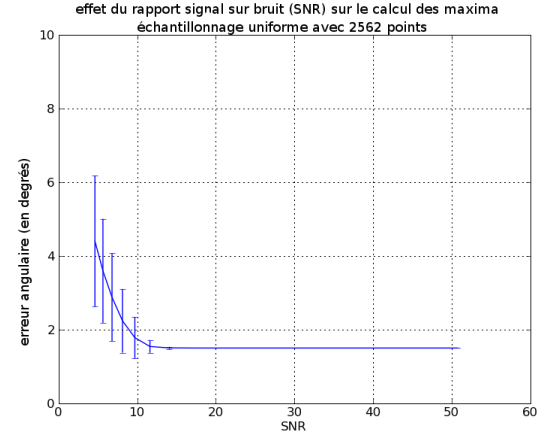
(b) recherche analytique



(c) recherche différentielle



(d) recherche exhaustive : 642 points



(e) recherche exhaustive : 2562 points

FIG. 13.2 – Comparaison des méthodes d'extraction des maxima : les maxima d'un croisement de fibre synthétique de 80° sont extraits mille fois à différents niveaux de bruit pour différentes réalisations de bruit. L'erreur de reconstruction des maxima estimés par rapport à la vérité terrain est tracée en fonction du rapport signal sur bruit.

13.2.2 Décomposition des tenseurs d'ordre supérieur

La décomposition que nous proposons commence par la détermination des maxima de l'ODF (cf. section 13.2.1). Ces maxima sont alignés sur les faisceaux principaux. Le nombre N de termes dans la décomposition est donc égal au nombre de maxima. Les $\mathbf{D}_i$ représentent

des faisceaux, ils sont définis positifs avec comme directions principales les directions des maxima de diffusion : $\vec{d}_i = (\theta_i, \phi_i)$. Pour ce faire, la décomposition spectrale d'une matrice symétrique définie positive est utilisée : $\mathbf{D}_i = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{U}_i^T$, où $\mathbf{U}_i = (\vec{e}_{i1}, \vec{e}_{i2}, \vec{e}_{i3})$ est une matrice orthonormale et $\mathbf{\Sigma}_i$ est une matrice diagonale avec des coefficients strictement positifs : $\lambda_{i1} \geq \lambda_{i2} = \lambda_{i3} > 0$. On fait l'hypothèse que la diffusion se fait dans une direction privilégiée et qu'un modèle linéaire de tenseur est suffisant pour cette modélisation (*i.e.*, une hypothèse de symétrie cylindrique). Finalement, les trois vecteurs propres ont pour expression : $\vec{e}_{i1} = (\theta_i, \phi_i)$, $\vec{e}_{i2} = (\theta_i + \pi/2, \phi_i)$ et $\vec{e}_{i3} = \vec{e}_{i1} \wedge \vec{e}_{i2}$. En écrivant les $\mathbf{D}_i$ sous leurs formes canoniques $\vec{D}_i = \text{vec}(\mathbf{D}_i) = [d_{11}^i, d_{22}^i, d_{33}^i, \sqrt{2}d_{12}^i, \sqrt{2}d_{13}^i, \sqrt{2}d_{23}^i]$, la décomposition proposée est donnée par :

$$\mathbf{T} = \sum_{i=1}^N \vec{D}_i^T \vec{D}_i + \epsilon = \sum_{i=1}^N \text{vec}^T(\mathbf{U}_i \mathbf{\Sigma}_i \mathbf{U}_i^T) \text{vec}(\mathbf{U}_i \mathbf{\Sigma}_i \mathbf{U}_i^T) + \epsilon \quad (13.3)$$

où ϵ est le résidu. Pour décomposer $\mathbf{T}$, une distance entre des tenseurs d'ordre 4 est nécessaire. La distance $\text{dist}(\mathbf{T}_1, \mathbf{T}_2)$ donnée par la relation (2.33) est utilisée. La méthode de décomposition proposée revient à résoudre le problème d'optimisation non-linéaire suivant :

$$\min_{\lambda'_{i1}, \lambda'_{i2}} \text{dist} \left(\mathbf{T}, \sum_{i=1}^N (\vec{X}_i^T \vec{X}_i) \right)^2 \quad (13.4)$$

où $\lambda'_{i1}, \lambda'_{i2}$ sont les variables à optimiser et $\vec{X}_i = \text{vec}(\mathbf{U}_i \text{diag}(\lambda'_{i1}, \lambda'_{i2}, \lambda'_{i2}) \mathbf{U}_i^T)$. Il est intéressant de noter que l'optimisation se fait sur $\lambda'_{ij} = \lambda_{ij}$ pour s'assurer que chaque λ_{ij} est positif. Les valeurs propres dans les directions principales λ'_{i1} sont initialisées avec la racine carrée des valeurs prises par l'ODF dans ces directions. La fonction de coût est minimisée en utilisant la méthode du simplex [9]. Notons qu'une telle décomposition peut se généraliser pour des tenseurs de diffusion d'ordre quelconque.

13.2.3 Application à la réorientation

Deux stratégies de réorientation des tenseurs d'ordre 4 sont proposées :

- Réorientation des seules directions de diffusion (RD) : Soient les directions principales de diffusion $\vec{d}_i$ calculées en section 13.2.1. La direction réorientée $\vec{d}_i'$ est alors donnée par $\vec{d}_i' = F\vec{d}_i / \|F\vec{d}_i\|$ et les coefficients de diffusion associés aux directions $\vec{d}_i'$ sont conservés. Cette méthode ne conserve pas le profil de diffusion de l'ODF. Toutefois, une telle stratégie peut être intéressante lors du calcul de tractographies.
- Réorientation basée sur une décomposition en tenseurs d'ordre 2 (RT) : Chaque tenseur d'ordre 2, issu de la décomposition proposée dans la section 13.2.2, est un tenseur SDP orienté sur un faisceau de fibres. Le PPD peut alors être utilisé pour réorienter chaque tenseur d'ordre 2. La combinaison de tous les tenseurs d'ordre 2 à partir de (13.3) permet de retrouver le tenseur d'ordre supérieur réorienté.

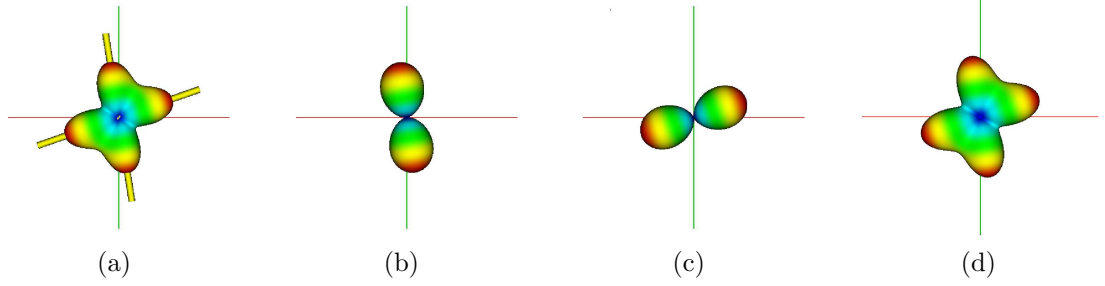


FIG. 13.3 – Réorientation proposée : (a) un croisement de deux faisceaux synthétiques, (b) premier et (c) deuxième tenseurs d'ordre 2 symétriques définis positifs issus de la décomposition et (d) le tenseur ODF d'ordre 4 reconstruit.

Δ_{angle}	Δs_1	Δs_2
90	0,018	0,31
80	0,91	0,10

(a)

Δ_{angle}	err
90	0.0017
80	0.0019

(b)

TAB. 13.1 – Erreurs de reconstruction : (a) erreur angulaire Δs_i (en degrés) entre l'orientation des faisceaux 1 et 2 extraite par l'algorithme de recherche des maxima et la vérité terrain en considérant deux configurations pour l'angle Δ_{angle} (en degrés) entre les faisceaux. (b) précision de la reconstruction en considérant les mêmes configurations pour l'angle Δ_{angle} entre les faisceaux.

13.3 Résultats expérimentaux

13.3.1 Utilisation de l'approche proposée

La précision de cette décomposition ainsi que la précision de la recombinaison (13.3) associée est étudiée dans la figure 13.3 et dans le tableau 13.1. Le tableau 13.1-a résume avec quelle précision les directions des maxima de l'ODF sont extraits avec une recherche exhaustive. L'erreur commise est faible. Soit $\mathbf{T}_{ref}$ le tenseur ODF d'ordre 4 de référence qui est décomposé avec la méthode proposée et $\mathbf{T}_{rec}$ le tenseur ODF d'ordre 4 reconstruit à partir du mélange de tenseurs d'ordre 2. Le tableau 13.1-b présente l'erreur de reconstruction qui est calculée à partir de l'équation suivante :

$$err = \frac{dist(\mathbf{T}_{ref}, \mathbf{T}_{rec})}{dist(\mathbf{T}_{ref}, 0)} \quad (13.5)$$

où le numérateur mesure la différence entre les tenseurs ODF et le dénominateur est un terme de normalisation. D'après le tableau 13.1-b, l'erreur de reconstruction commise est du même ordre de grandeur pour les deux configurations angulaires simulées.

Afin de tester la méthode dans sa globalité, un croisement de fibres synthétique échantillonné sur une grille régulière $L \times L$ (dans la cas présent $L = 13$) est considéré (cf. figure 13.4-a).

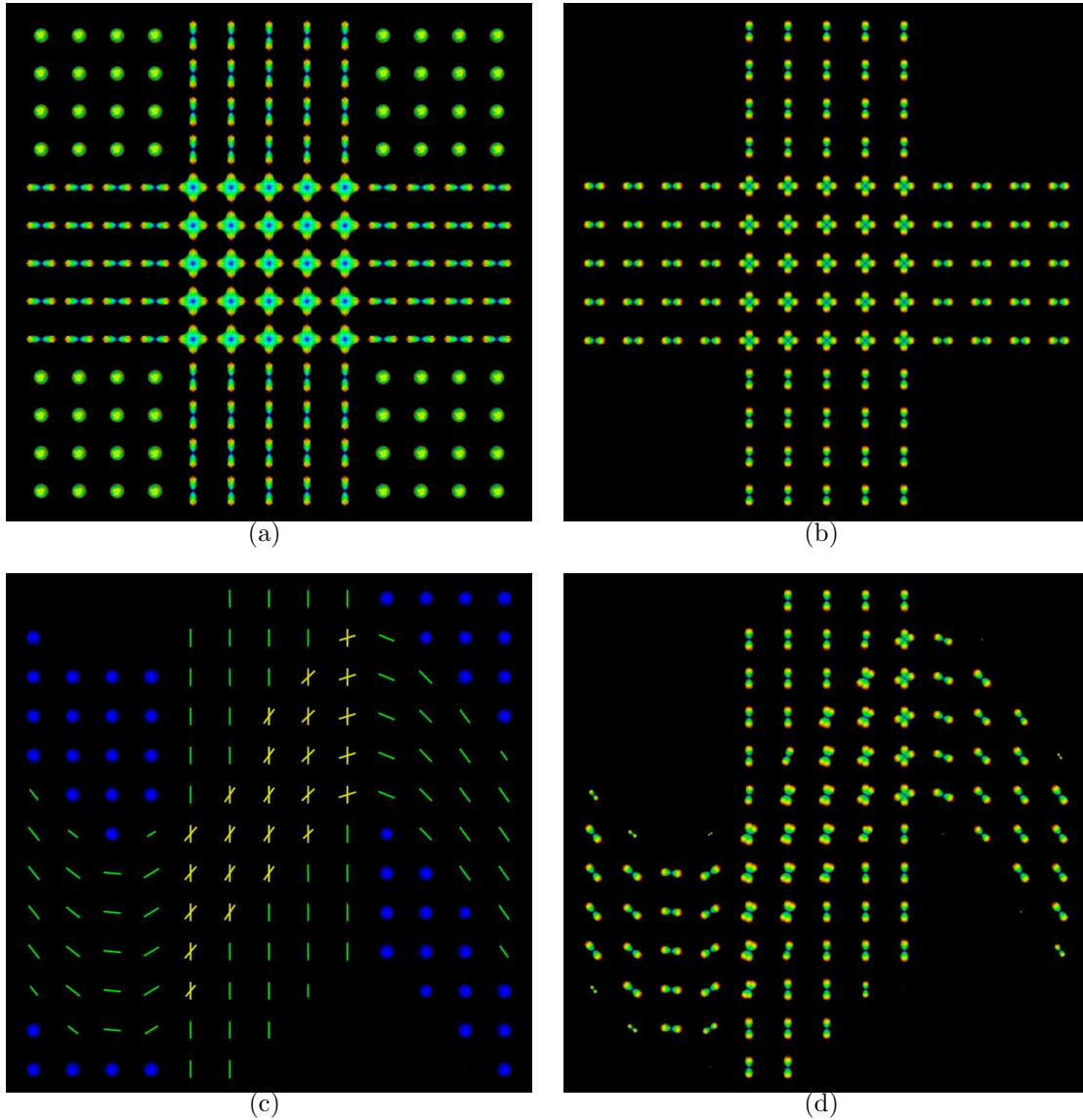


FIG. 13.4 – Réorientation proposée : (a) croisement de fibres simulé et (b) décomposition associée aux tenseurs ODF d'ordre 4. Application à l'image (b) d'un champ de déformation sinusoïdal avec les méthodes (c) RD et (d) RT.

Le signal de diffusion a été simulé en utilisant la méthode proposée dans [2]. Cette méthode calcule une approximation précise et continue du signal de diffusion en utilisant les noyaux adaptatifs¹. Un seuil sur l'anisotropie généralisée (GA) de 0,1 permet d'écarter les formes isotropes. La décomposition proposée est illustrée en figure 13.4-b. Ensuite, un champ de déformations sinusoïdal est appliqué. La figure 13.4-c illustre la réorientation RD, alors que la figure 13.4-d illustre la réorientation RT. On remarque que les résultats obtenus sont très satisfaisants. En effet, seul le faisceau horizontal est réorienté, alors que le faisceau vertical

¹<http://www.cise.ufl.edu/~abampou/lab/DWMRI.simulator.php>

Trans.	Meth.	$\Delta\theta_1$	$\Delta\theta_2$	Δd_1	Δd_2
T_1	R_{rigide}	0,92	1,06	0,0	0,0
T_1	$R_{Hilbert}$	16,96	63,04	9,38	90,26
T_1	RD	0,91	1,06	0,04	0,17
T_1	RT	0,57	1,40	1,07	1,24
T_2	R_{rigide}	7,66	9,76	3,24	5,26
T_2	$R_{Hilbert}$	4,47	0,95	35,66	3,54
T_2	RD	1,02	0,96	0,04	0,17
T_2	RT	1,14	0,84	0,96	1,27

TAB. 13.2 – Comparaisons avec les méthodes existantes : les angles sont en degrés et les erreurs en pourcentages (cf. texte).

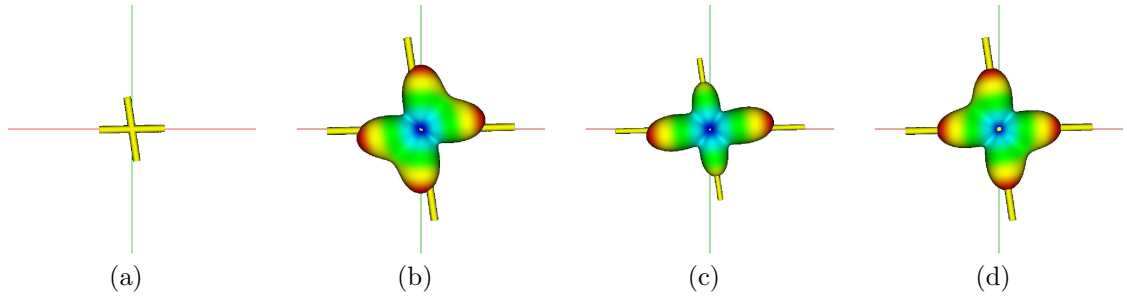


FIG. 13.5 – Transformation T_2 appliquée au tenseur de la figure 13.3 (a), pour les stratégies de réorientation (a) RD, (b) R_{rigide} , (c) $R_{Hilbert}$ et (d) RT.

conserve la même orientation.

13.3.2 Comparaison avec des méthodes de la littérature

Pour finir cette étude, la réorientation rigide (R_{rigide} , cf. section 13.1.3), la réorientation de Hilbert ($R_{Hilbert}$, cf. section 13.1.4) et la réorientation RT, proposée dans cette section, sont comparées en faisant deux simulations permettant de mettre en évidence les avantages et les inconvénients de chaque stratégie. Ainsi, deux transformations sont appliquées à un croisement de deux faisceaux séparés d'un angle de 80 degrés : une rotation d'un angle de 45° autour de l'axe z (notée T_1) et un cisaillement vertical :

$$T_2 : \begin{pmatrix} 1 & 0 & 0 \\ -\tan(18^\circ) & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (13.6)$$

Les performances de chaque méthode sont évaluées en terme d'erreur angulaire $\Delta\theta_i$ (entre les maxima du tenseur ODF réorienté et les directions théoriques obtenues après réorientation) et en terme de modifications du profil de l'ODF dans ces directions Δd_i . Les résultats sont comparés dans le tableau 13.2. La méthode proposée semble être préférable pour des transformations non-linéaires. La méthode RT donne des résultats légèrement moins bon que la

méthode RV puisque la reconstruction du tenseur ODF d'ordre 4 introduit une délocalisation des directions de chaque faisceau. Finalement, en figure 13.5, les tenseurs réorientés après application de la transformation T_2 sont présentés. Sur cet exemple, seule la méthode proposée RT semble permettre de reconstruire correctement le profil de l'ODF après réorientation.

Chapitre 14

Un test statistique multivarié pour des tenseurs d'ordre 4

Sommaire

14.1 Chaîne de traitement	215
14.2 Résultats	217
14.2.1 Simulations	217
14.2.2 Application à des données cliniques	218

Les deux chaînes de traitements présentées dans le chapitre 9 sont basées sur une modélisation de la diffusion des molécules d'eau par un tenseur d'ordre 2. En imagerie de diffusion, il est fréquent de mesurer, en un voxel, la contribution de différents faisceaux de fibres de la substance blanche. Le modèle du tenseur d'ordre 2 est alors inadapté car il permet de modéliser au plus un faisceau. Ce chapitre propose une extension de ces méthodes aux tenseurs $\mathbf{T}$ d'ordre 4 dans leurs représentations matricielles 6×6 . Puisque cette matrice est symétrique, il est possible d'utiliser directement les tests statistiques multivariés présentés dans le chapitre 6. On fait l'hypothèse que $\mathbf{T} \in \text{Sym}(6)$ suit une distribution multinormale. Aucune précaution n'est toutefois prise pour garantir la positivité des tenseurs d'ordre 4, qui ne se résume pas à une simple contrainte de positivité des valeurs propres, qui, rappelons le ici, sont sans interprétation physique immédiate (notamment en terme de maximum du profil de diffusion). Le test sélectionné pour cette étude est alors le test sur le tenseur de diffusion complet (cf. section 6.2.2). Des tests sur données simulées et l'application à un cas clinique permettent de comparer les résultats obtenus avec un modèle d'ordre 2 et un modèle d'ordre 4.

14.1 Chaîne de traitement

La détection de changement suit le schéma général présenté dans la section 9.1 et reproduit en figure 14.1. L'estimation de l'ITD d'ordre 2 se fait en utilisant la méthode classique des moindres carrés, alors que l'estimation de l'ITD d'ordre 4 se fait en utilisant la méthode prenant en considération les angles solides (*i.e.*, estimation de l'ODF, cf. section 3.1.3).

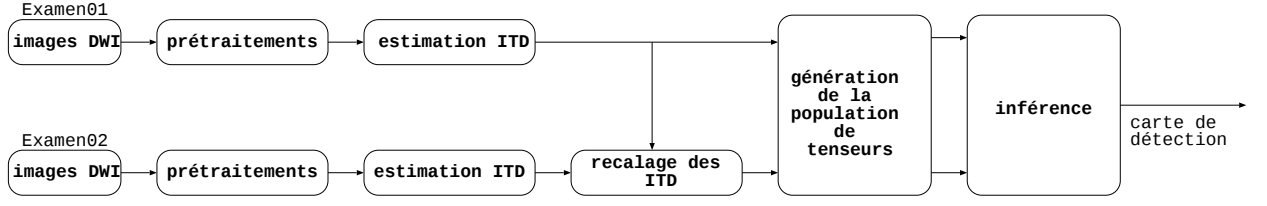


FIG. 14.1 – Etapes à suivre pour effectuer un suivi longitudinal lors d’une détection à l’échelle du voxel.

Rappelons que l’unique maximum d’un modèle d’ordre 2 est nécessairement aligné avec l’orientation principale des fibres sous-jacentes. Puisque aucune stratégie simple ne permet de garantir la positivité des tenseurs d’ordre 4, aucune précaution particulière n’est prise pour les tenseurs d’ordre 4 et d’ordre 2 (afin de ne pas introduire de biais dans la comparaison des deux méthodes). Lors de l’étape de recalage non-rigide, les tenseurs d’ordre 2 ou d’ordre 4 sont réorientés avec une méthode de réorientation rigide (méthode dite du *Finite Strain*) (cf. sections 7.3 et 13.1.3). A l’ordre 4, cette méthode a l’avantage d’être très rapide et donne, dans le cas du suivi d’un même patient, des résultats très satisfaisants. De plus, ce choix permet d’utiliser la même méthode sur les tenseurs d’ordre 2 et d’ordre 4 afin de ne pas introduire de biais dans la comparaison des deux méthodes. Pour le choix de la méthode de génération des populations de tenseurs à comparer, nous utiliserons la stratégie de prise en compte du voisinage à l’échelle du voxel (*i.e.*, un voisinage spatial cubique, centré sur le voxel d’étude). Il aurait été envisageable d’utiliser les stratégies de bootstrap local ou spatial ou la stratégie basée sur un *a priori* anatomique construit à partir des fibres de la substance blanche. Finalement, pour les raisons expliquées dans l’introduction de ce chapitre, un test statistique sur les tenseur de diffusion T_{e-tens} est calculé pour les tenseurs d’ordre 2 et d’ordre 4 (cf. section 6.2.2) :

$$T_{e-tens} = \frac{(n-2)n_1n_2}{qn^2\hat{\sigma}^2} tr \left[(\bar{\mathbf{T}}_1 - \bar{\mathbf{T}}_2)^2 \right] \sim F(q, q(n-2)) \quad (14.1)$$

où n_1, n_2 sont le nombre de tenseurs dans la première et deuxième population, $n = n_1 + n_2$, et $\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2$ sont les moyennes des deux populations de tenseurs à l’ordre 2 ou à l’ordre 4. q vaut 6 ou 21 à l’ordre 2 ou 4 respectivement. L’estimation des moyennes $\bar{\mathbf{T}}_1$ et $\bar{\mathbf{T}}_2$ de chaque population de tenseurs s’écrit (cf. section 6.2.2) :

$$\bar{\mathbf{T}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{T}_{i,j} \quad (14.2)$$

et l’estimation de la variance s’écrit (cf. section 6.2.2) :

$$\hat{\sigma}^2 = \frac{1}{qn} \left[\sum_{j=1}^{n_1} tr \left[(\mathbf{T}_{1,j} - \bar{\mathbf{T}}_1)^2 \right] + \sum_{j=1}^{n_2} tr \left[(\mathbf{T}_{2,j} - \bar{\mathbf{T}}_2)^2 \right] \right] \quad (14.3)$$

La carte de détection contient les p-valeurs corrigées avec la stratégie proposée en section 6.2.3.

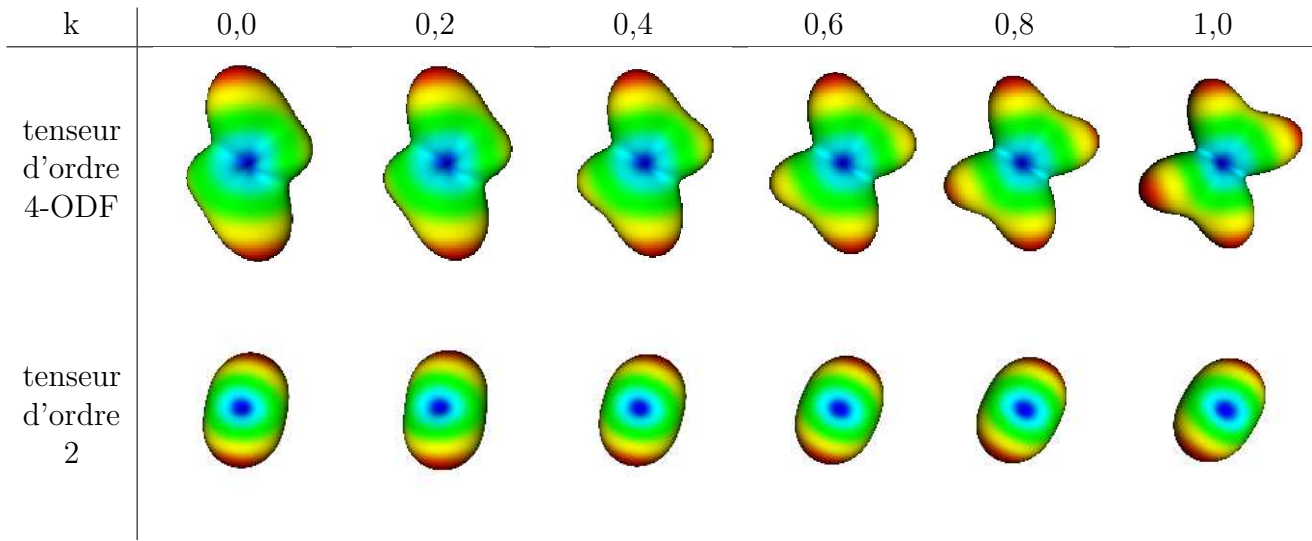


FIG. 14.2 – Modification du signal de diffusion dans un cône de 35° autour du faisceau principal. Cette modification correspond à une diminution appliquée en multipliant les valeurs de diffusion avec un facteur $k \in [0, 1]$. Nous représentons ici les ODF exprimées avec des tenseurs d'ordre 4 ou d'ordre 2. Lorsque $k = 1$, on applique aucune modification et on obtient le modèle de simulation sans bruit.

14.2 Résultats

14.2.1 Simulations

Dans cette section, le test statistique T_{e-tens} (14.1) est calculé sur des données synthétiques à partir de tenseurs d'ordre 2 et d'ordre 4. L'objectif est de comparer ces deux modélisations et de voir si l'extension au modèle d'ordre supérieur présente un intérêt. Pour générer les données synthétiques, le signal d'un croisement de deux faisceaux séparés de 80° est utilisé. Deux images de taille $50 \times 50 \times 50$ sont générées en ajoutant un bruit gaussien sur le signal en chaque voxel de l'image. Finalement, dans une des deux images, une modification est simulée dans un voisinage de taille $16 \times 16 \times 16$ (soit une modification sur environ 3% des voxels). Le changement est introduit en modifiant le signal dans un voisinage de la direction du faisceau associé au maximum de diffusion (*i.e.*, un cône de 35° autour de cette direction est considéré). Les valeurs des signaux de diffusion, dont les gradients appartiennent au cône de modification, sont diminuées (*i.e.*, la diffusion est augmentée). Cette diminution est réalisée en multipliant la valeur du signal de diffusion par un facteur $k \in [0, 1]$. Ce schéma de modification est illustré en figure 14.2.

Pour comparer les deux modélisations du signal de diffusion, un critère basé sur l'aire sous les courbes COR est utilisé. Ces courbes COR ont été calculées sur l'intégralité de l'image en utilisant les statistiques de test non corrigées pour les comparaisons multiples. Le tableau 14.1 présente les résultats obtenus pour le modèle du tenseur-ODF d'ordre 4 et le modèle du tenseur d'ordre 2. Pour les traitements réalisés, un voisinage spatial $3 \times 3 \times 3$ est considéré pour engendrer les populations de tenseurs à comparer et deux rapports signal à bruit sont

SNR=26dB									
k	1,00	0,975	0,950	0,925	0,900	0,875	0,850	0,825	0,800
tenseur d'ordre 4-ODF	0,560	0,621	0,743	0,871	0,937	0,976	0,988	0,995	0,997
tenseur d'ordre 2	0,559	0,726	0,934	0,989	0,996	0,999	0,999	0,999	0,999
SNR=20dB									
k	1,00	0,975	0,950	0,925	0,900	0,875	0,850	0,825	0,800
tenseur d'ordre 4-ODF	0,556	0,578	0,600	0,666	0,733	0,821	0,863	0,910	0,944
tenseur d'ordre 2	0,545	0,596	0,732	0,834	0,930	0,969	0,987	0,993	0,996

TAB. 14.1 – Aires sous les courbes COR obtenues pour le modèle du tenseur-ODF d'ordre 4 et le modèle du tenseur d'ordre 2. Les valeurs de diffusion sont diminuées dans le cône de modification en appliquant un facteur multiplicatif $k \in [0, 1]$. Lorsque $k = 1$, aucune modification n'est simulée. Une aire d'environ 0,5 est alors tout à fait normale.

considérés (26dB et 20dB).

Étonnamment, le test basé sur le modèle du tenseur d'ordre 2 donne les meilleurs résultats. La modification simulée sur un des deux faisceaux pourrait expliquer cette tendance. L'estimation du tenseur d'ordre 2 serait alors affectée de manière plus significative (cf. figure 14.2). Cette explication nécessiterait toutefois d'autres évaluations. Malgré ces simulations infructueuses, nous allons comparer les deux stratégies sur données réelles afin de s'affranchir de la stratégie de simulation mise en place.

14.2.2 Application à des données cliniques

Dans cette section, les tests statistiques T_{e-tens} (14.1) calculés avec un modèle d'ordre 2 ou d'ordre 4 sont comparés sur un patient atteint de SEP. Les données de diffusion de ce patient ont été décrites en section 11.2.1. Les traitements sont réalisés avec un voisinage spatial $3 \times 3 \times 3$ pour engendrer les populations de tenseurs à comparer. Une analyse visuelle des résultats permet d'affirmer que la plupart des évolutions détectées le sont par les deux approches. Les figures 14.3 et 14.4 présentent les résultats issus d'une modélisation à l'ordre 2 et à l'ordre 4. La détection en figure 14.3 correspond à une nouvelle lésion qui apparaît dans le deuxième examen. On remarque que cette évolution est diffuse avec le modèle d'ordre 2 et très localisée avec le modèle d'ordre 4. Ceci semble en accord avec les simulations où la stratégie basée sur un tenseur d'ordre 2 a permis de détecter de faibles changements. La détection en figure 14.4 peut être attribuée à une remyélinisation. Cette fois, la modélisation à l'ordre 4 permet de mieux détecter l'évolution et fournit une carte statistique présentant moins d'artefacts de détection (notamment au niveau des ventricules). Les résultats des deux approches sont similaires, avec toutefois une réduction de certains artefacts avec le modèle d'ordre 4.

Au vu de ces résultats préliminaires, le modèle d'ordre 4 ne semble pas améliorer significativement les résultats. Ces résultats peuvent être expliqués par l'ajout de dimensions supplémentaires lorsque l'on travaille avec un tenseur d'ordre 4, tout en conservant un nombre

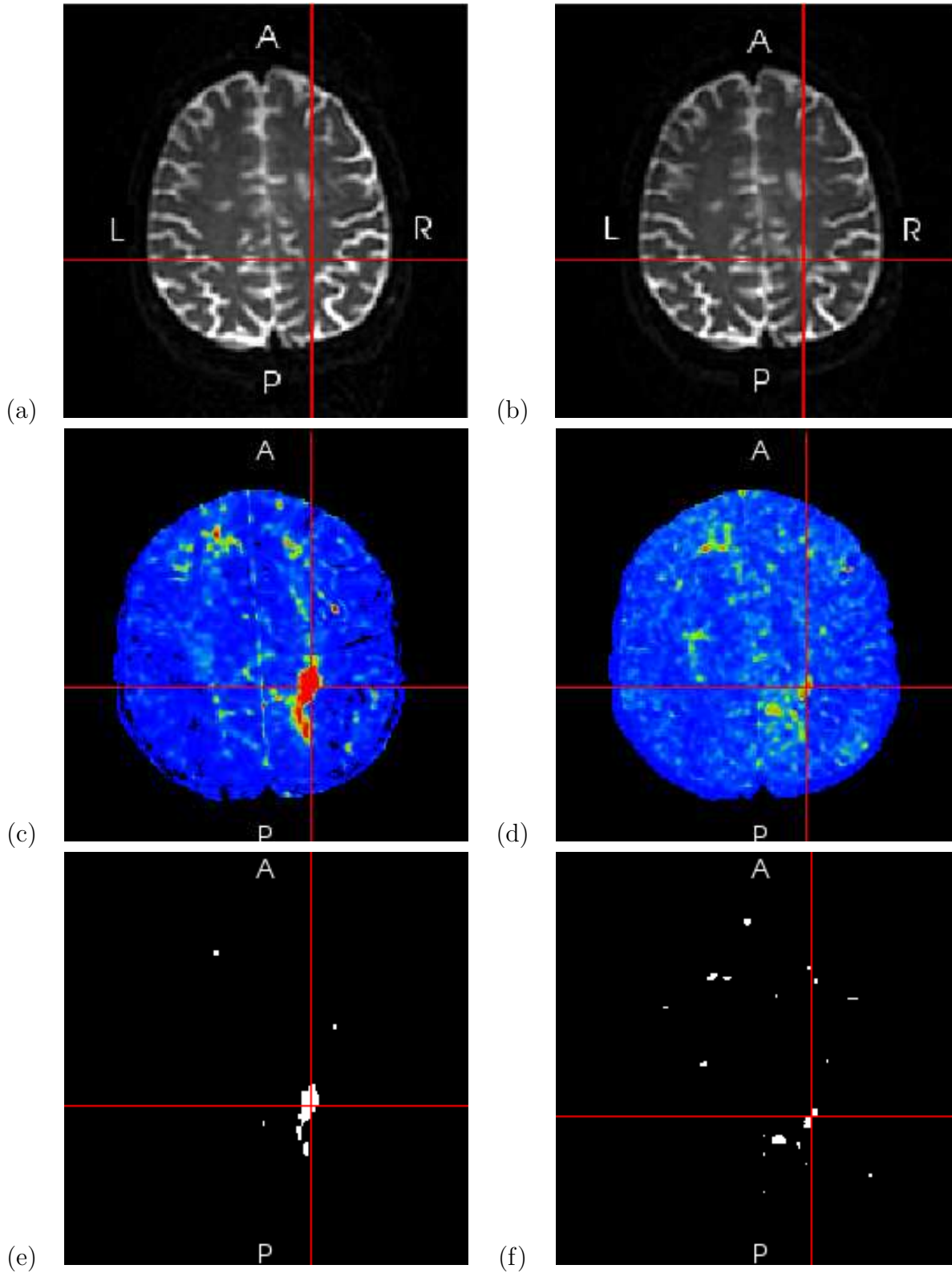


FIG. 14.3 – Cartes de détection obtenues lors du suivi longitudinal d'un patient atteint de SEP, avec un modèle du tenseur (c) d'ordre 2 et (d) d'ordre 4. Les cartes seuillées correspondantes (e) et (f) sont obtenues après correction FDR ($p_{FDR} = 0,001$). Les images pondérées en T_2 , acquises aux temps t_1 et t_2 , sont respectivement présentées en (a) et (b).

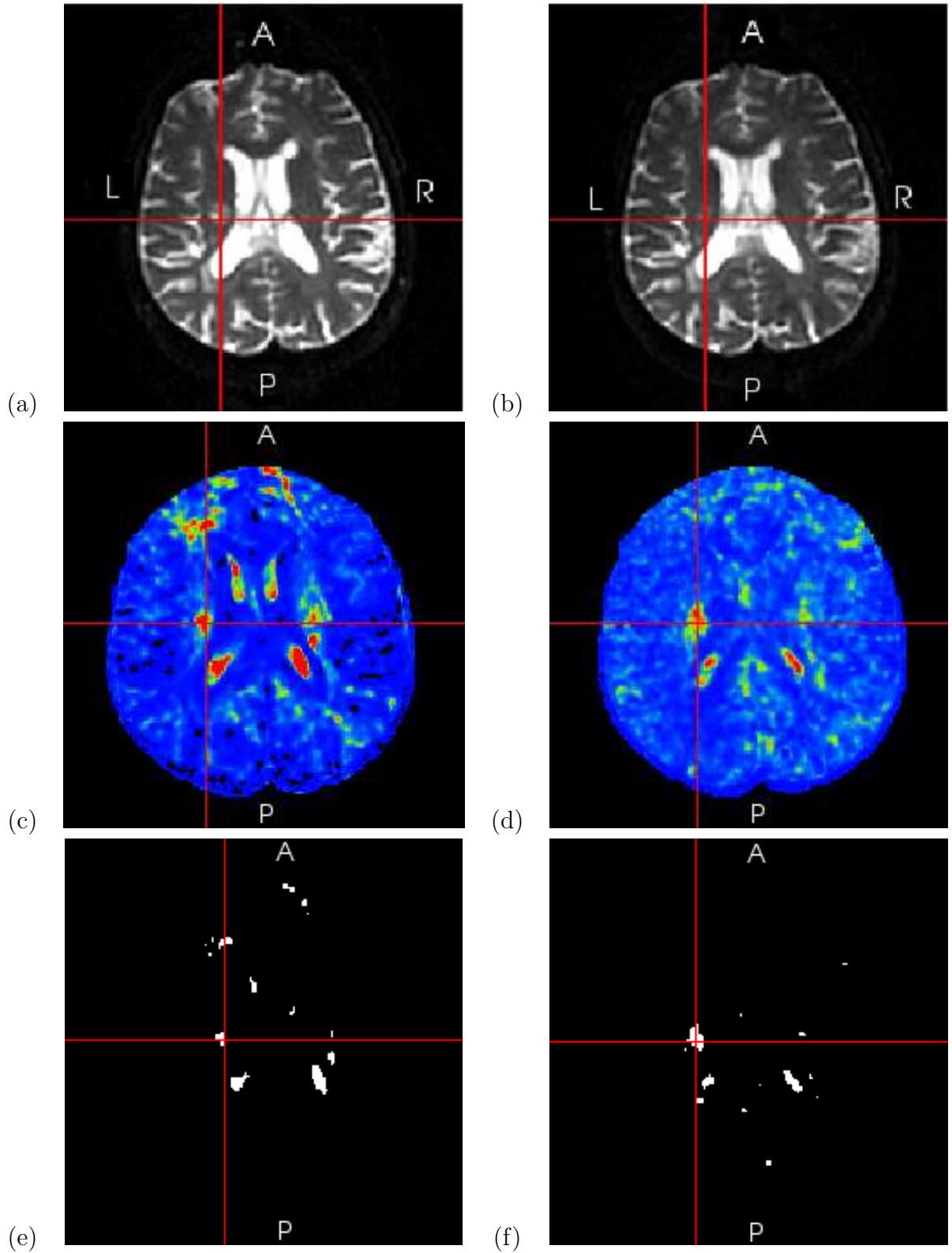


FIG. 14.4 – Cartes de détection obtenues lors du suivi longitudinal d'un patient atteint de SEP, avec un modèle du tenseur (c) d'ordre 2 et (d) d'ordre 4. Les cartes seuillées correspondantes (e) et (f) sont obtenues après correction FDR ($p_{FDR} = 0,001$). Les images pondérées en T_2 , acquises aux temps t_1 et t_2 , sont respectivement présentées en (a) et (b).

d'échantillon fixe dans chaque population. Une piste pour résoudre ce problème pourrait consister à utiliser une stratégie de bootstrap pour augmenter le nombre d'échantillons. Le chapitre suivant propose une autre alternative en s'intéressant uniquement à une information interprétable des tenseurs d'ordre 4, les directions de diffusion. Une mesure de similarité est construite à partir des directions de diffusion et devrait davantage permettre de mettre en évidence les contributions possibles d'un modèle d'ordre supérieur.

Chapitre 15

Détection de changements d'orientations des fibres par une mesure de similarité

Sommaire

15.1 Mesure de similarité pour des tenseurs d'ordre 2	224
15.1.1 La matrice de dispersion	224
15.1.2 La similarité entre deux populations	224
15.2 Mesure de similarité pour des tenseurs d'ordre 4	224
15.2.1 Une association intra-population	225
15.2.2 Une association inter-population	225
15.3 Chaîne de traitement	225
15.4 Résultats	226
15.4.1 Simulations	226
15.4.2 Application à des données cliniques	227

POUR des tenseurs d'ordre 2, nous avons vu en section 6.2.2 qu'il était possible de quantifier l'évolution de l'organisation de la matière blanche au cours du temps en utilisant le test T_{2VEC} . Ce test fait l'hypothèse que les matrices des valeurs propres sont égales. En pratique, cette hypothèse est bien souvent non vérifiée. De plus, appliquer ce test sur des tenseurs d'ordre 4 dans leur représentation matricielle 6×6 n'a que peu de sens. En effet, la matrice des vecteurs propres d'un tenseur d'ordre 4 n'a pas d'interprétation simple. Une possibilité pour détecter des changements d'orientations des fibres en considérant des tenseurs d'ordre 2 consiste à utiliser une mesure de similarité construite directement sur les vecteurs propres principaux dans une région d'intérêt [12]. Pour ce faire, la matrice de dispersion des vecteurs propres principaux est utilisée pour décrire la distribution de ces vecteurs dans la région anatomique d'intérêt. Après avoir présenté cette stratégie, une extension aux tenseurs d'ordre 4 permettant de tenir compte de plusieurs directions de diffusion privilégiées est introduite dans ce chapitre. Une chaîne de traitement est alors élaborée pour mesurer localement la cohérence des orientations des faisceaux de fibres de la substance blanche. Cette méthode explore les modèles d'ordre supérieur, permettant une modélisation plus réaliste de

la diffusion des molécules d'eau dans le cerveau humain. Finalement, la méthode est évaluée sur données simulées, avant d'effectuer une étude sur données réelles.

15.1 Mesure de similarité pour des tenseurs d'ordre 2

Les vecteurs propres principaux des tenseurs de la population sont utilisés pour calculer une matrice de dispersion permettant ensuite de calculer un critère de similarité. Ces vecteurs propres correspondent aux directions de diffusion privilégiées et sont calculés, pour un tenseur d'ordre 2, en utilisant la décomposition spectrale.

15.1.1 La matrice de dispersion

Soient $\vec{e}_{1,1}, \dots, \vec{e}_{1,n_1}$ et $\vec{e}_{2,1}, \dots, \vec{e}_{2,n_2}$ deux échantillons de vecteurs unitaires qu'il est possible de représenter sur la sphère unité (dans la cadre du suivi longitudinal, ces vecteurs unitaires correspondent aux vecteurs propres associés aux plus grandes valeurs propres d'un échantillon de tenseurs de diffusion). Sous ces hypothèses, la dispersion des vecteurs unitaires de chaque échantillon peut être représentée par une matrice de dispersion $\mathbf{T}_l$, $l \in (1, 2)$. L'expression de cette matrice de dispersion est fonction des vecteurs unitaires de chaque échantillon et est donnée par la relation suivante :

$$\mathbf{T}_l = \frac{1}{n_l} \sum_{i=1}^{n_l} \vec{e}_{l,i} \vec{e}_{l,i}^T \quad (15.1)$$

T_l est une matrice 3×3 symétrique qui quantifie la distribution des vecteurs unitaires dans une région d'intérêt. Les valeurs propres t_1, t_2, t_3 de cette matrice caractérisent la distribution des points en coordonnées sphériques. Les vecteurs propres associés $\vec{t}_1, \vec{t}_2, \vec{t}_3$ caractérisent l'orientation de cette distribution.

15.1.2 La similarité entre deux populations

Une mesure de cohérence entre deux échantillons représentés par leurs matrices de dispersion $\mathbf{T}_1$ et $\mathbf{T}_2$ peut être calculée en considérant la valeur absolue du produit scalaire entre les vecteurs propres principaux $\vec{t}_{1,1}$ et $\vec{t}_{2,1}$ de chacune de ces matrices :

$$S_{T_2} = |\vec{t}_{11}^T \cdot \vec{t}_{21}| \quad (15.2)$$

Une forte cohérence se traduit par une valeur S_{T_2} proche de un, alors que S_{T_2} se rapproche de zéro lorsque les matrices de dispersion sont presque orthogonales.

15.2 Mesure de similarité pour des tenseurs d'ordre 4

Les extrema d'un tenseur d'ordre 4 sont calculés en utilisant une recherche exhaustive sur la sphère unité échantillonnée avec 642 points (cf. section 13.2.1). Ensuite, à l'ordre 4, la stratégie à mettre en œuvre pour calculer la mesure de similarité est plus complexe. Puisqu'un

tenseur ODF d'ordre 4 peut modéliser plusieurs faisceaux, un processus d'association en deux étapes est utilisé : une association intra-population suivie par une association inter-population.

15.2.1 Une association intra-population

Les maxima calculés pour chaque population de tenseurs sont regroupés en groupes d'orientations homogènes en utilisant un algorithme de classification par k-moyennes sphérique [4] (la métrique utilisée est le produit scalaire). L'algorithme de k-moyennes sphérique est initialisé avec les directions des maxima du tenseur ODF de la population i présentant le plus grand nombre maxima m_i (on aura donc m_i classes dont les centroïdes sont initialisés avec les directions des maxima). Après cette première étape, m_1 et m_2 groupes sont obtenus à partir de chacune des populations. Une matrice de dispersion est alors calculée pour chaque groupe.

15.2.2 Une association inter-population

Les vecteurs propres principaux des matrices de dispersion de chaque groupe sont placés dans deux listes l_1 et l_2 . Il faut noter que les deux listes n'ont pas nécessairement le même nombre d'éléments. Ainsi, les listes l_1 et l_2 sont traitées afin d'obtenir les $m = \min(m_1, m_2)$ paires de directions les plus similaires selon le produit scalaire. La mesure pondérée de similarité proposée pour un tenseur ODF d'ordre 4 est donnée par :

$$S_{T4} = \sum_{i=1}^m w_i |\vec{l}_1(i)^T \cdot \vec{l}_2(i)| \quad (15.3)$$

avec :

$$w_i = \frac{\text{Card}(l_1(i)) + \text{Card}(l_2(i))}{\sum_{i=1}^{m_1} \text{Card}(l_1(i)) + \sum_{i=1}^{m_2} \text{Card}(l_2(i))} \quad (15.4)$$

où $\text{Card}(l_j(i))$ représente le nombre de maxima dans le i -ème groupe de la population j .

15.3 Chaîne de traitement

Pour le suivi longitudinal de l'organisation de la substance blanche, une chaîne de traitement dédiée a été imaginée (cf. figure 15.1). On s'intéresse ici à l'orientation des tenseurs. En imagerie de diffusion, on observe un phénomène microscopique d'un point de vue macroscopique. Il est donc possible (et même très fréquent) de mesurer les contributions de différents faisceaux en un même voxel. Pour cette raison, un modèle d'ordre 2 ou un modèle d'ordre 4 est considéré. L'estimation de l'ITD d'ordre 2 se fait en utilisant la méthode classique des moindres carrés, alors que l'estimation de l'ITD d'ordre 4 se fait en utilisant la méthode prenant en considération les angles solides (*i.e.*, estimation de l'ODF, cf. section 3.1.3). Lors du recalage des modèles d'ordre 2 ou 4, une méthode de réorientation rigide ou *Finite Strain* est utilisée (cf. sections 7.3 et 13.1.3). Ce choix permet d'effectuer un traitement comparable

Vers un suivi longitudinal avec des modèles d'ordre supérieur

Détection de changements d'orientations des fibres par une mesure de similarité

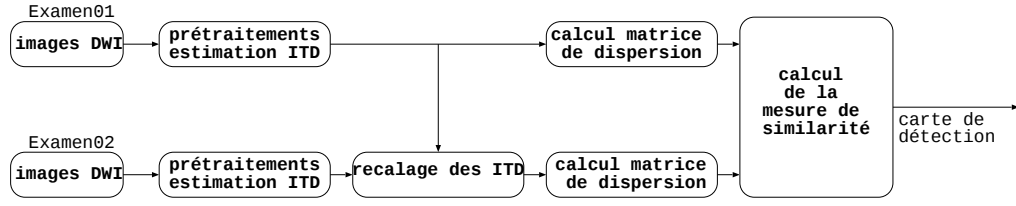


FIG. 15.1 – Etapes à suivre pour effectuer un suivi longitudinal des changements d'orientations des fibres par mesure de similarité.

sur les tenseurs d'ordre 2 et d'ordre 4 lors de la réorientation et facilite ainsi la comparaison des deux approches. Une mesure de similarité est calculée à partir des matrices de dispersion estimées en un voxel dans les deux examens recalés. Pour calculer ces matrices de dispersion, un voisinage spatial cubique centré sur le voxel d'étude et de taille variable fixée par l'utilisateur est utilisé. Il aurait été envisageable d'utiliser les stratégies de bootstrap local ou spatial ou la stratégie basée sur un *a priori* anatomique pour former les populations de tenseurs à comparer.

15.4 Résultats

15.4.1 Simulations

Les données synthétiques sont simulées en utilisant un mélange de gaussiennes. Soit $\mathbf{D}_1$ un tenseur diagonale de termes diagonaux $1,7 \times 10^{-3}$, $2,0 \times 10^{-4}$ et $2,0 \times 10^{-4}$. Le tenseur $\mathbf{D}_2(\phi)$ est calculée en tournant $\mathbf{D}_1$ d'un angle ϕ autour du troisième vecteur propre. Soit $\vec{g}$ un vecteur de la sphère unité. Pour deux faisceaux de fibres, le signal simulé est donné par :

$$s(\vec{g}, b) = \sum_{i=1}^2 f_i \exp(-b \vec{g}^T \mathbf{D}_i \vec{g}) \quad (15.5)$$

où f_i et $\mathbf{D}_i$ sont, respectivement, la proportion et le tenseur de diffusion associés au faisceau i . Dans notre expérience, les f_i valent $1/2$ et b vaut 3000.

A partir d'un ensemble de signaux notés $s(\vec{g}, b)$ calculés pour différentes configurations de croisements de fibres $\phi \in [0, \pi/4, 5\pi/6, 3\pi/8, 7\pi/16, \pi/2]$, les tenseurs d'ordre 2 (cf. figure 15.2-a) et les tenseurs ODF d'ordre 4 (cf. figure 15.2-b) associés sont calculés. Quand ϕ varie de 0 à $\pi/2$, l'organisation des fibres passe d'un faisceau homogène organisé autour d'une direction privilégiée à un croisement orthogonal de deux faisceaux de fibres. On remarque immédiatement que seul le modèle d'ordre 4 permet de restituer l'information d'in-homogénéité des structures au niveau voxel, dès lors que l'angle entre les faisceaux est plus grand que $\pi/4$.

Pour tester la méthode de détection de changements et pour mettre en avant l'intérêt d'un modèle d'ordre 4 par rapport un modèle d'ordre 2, de multiples paires d'images de taille $10 \times 10 \times 10$ sont générées pour différentes configurations angulaires ϕ . Un bruit gaussien

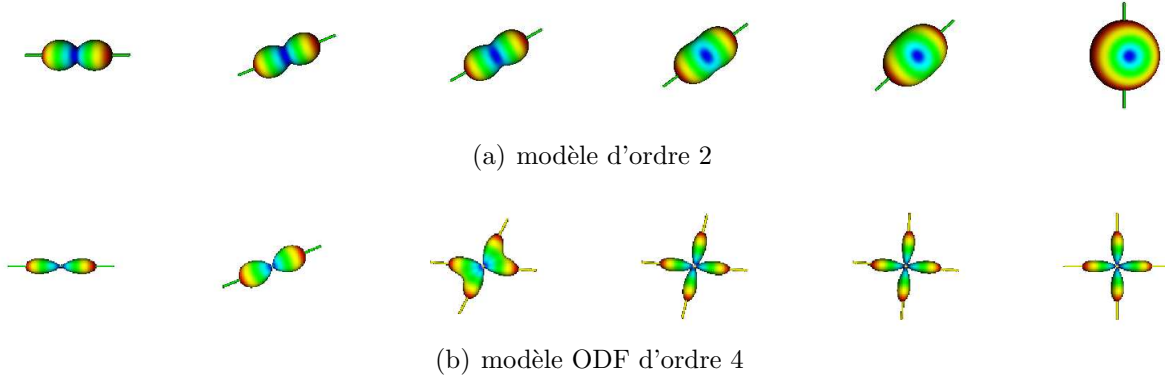


FIG. 15.2 – Estimation des modèles et des directions de fibre associées pour un signal idéal (pas de bruit) quand $\phi \in [0, \pi/4, 5\pi/6, 3\pi/8, 7\pi/16, \pi/2]$.

est ajouté au signal $s(\vec{g}, b)$ ($\sigma = 0,2$). Alors que ϕ varie de 0 à $\pi/2$, toutes les combinaisons angulaires possibles pour les images à comparer sont simulées. Les résultats des mesures de cohérence S_{T_2} à l'ordre 2 et S_{T_4} à l'ordre 4 sont disponibles dans le tableau 15.1. Les deux populations inter-images à comparer en chaque voxel sont extraites en utilisant un simple voisinage spatial de taille $3 \times 3 \times 3$. Les valeurs dans le tableau 15.1-b,c représentent la valeur moyenne de cohérence de chaque couple d'images. Les valeurs théoriques présentées dans le tableau 15.1-a sont calculées en considérant la moyenne des deux produits scalaires associés à chacun des faisceaux simulés.

Les meilleurs résultats sont obtenus avec le modèle qui permet de modéliser au mieux les données (*i.e.*, le tenseur ODF d'ordre 4). En effet, pour des configurations angulaires $\phi > \pi/4$, seul le modèle d'ordre 4 offre une mesure fidèle de la similarité inter-images. Pour des configurations angulaires $\phi \leq \pi/4$, le modèle d'ordre 4 est mis à défaut et ne modélise plus correctement le croisement de fibres. Ainsi, le maximum théorique est de 0,5 pour la mesure de cohérence S_{T_4} puisqu'un faisceau est manquant (*i.e.*, comparaison d'un faisceau *vs* deux faisceaux).

15.4.2 Application à des données cliniques

Nous avons traité le patient atteint de SEP présenté en section 14.2.2 avec cette chaîne de traitement, mais aucune modification de l'organisation des faisceaux de fibres de la substance n'a été détectée. L'étude que nous proposons ici porte sur l'analyse longitudinale de la maturation cérébrale d'un fœtus. Les images d'un cerveau de fœtus ont été acquises *in utero* avec un scanner IRM Siemens de 1,5 Tesla. Le protocole incluait une série d'images de diffusion avec 60 directions de gradients et une b-valeur de 700 s/mm^2 . Les deux images ont été acquises à 21 jours d'intervalle (correspondant à la 27-ème et à la 30-ème semaine de gestation). L'acquisition d'images fœtale est sujette à de nombreux artefacts liés au mouvement du fœtus. Pour compenser au mieux ces mouvements lors de l'acquisition, les images pondérées en diffusion sont reconstruites en utilisant l'algorithme proposé dans [10]. Après ce premier prétraitement, la taille des images du premier examen est de $42 \times 42 \times 31$ avec

Vers un suivi longitudinal avec des modèles d'ordre supérieur

Détection de changements d'orientations des fibres par une mesure de similarité

(a) théorique

$\phi \backslash \phi$	0,0	$\pi/4$	$5\pi/16$	$3\pi/8$	$7\pi/16$	$\pi/2$
0,0	1,000	0,500	0,500	0,500	0,500	0,500
$\pi/4$	0,500	1,000	0,986	0,944	0,903	0,854
$5\pi/16$	0,500	0,986	1,000	0,986	0,962	0,927
$3\pi/8$	0,500	0,944	0,986	1,000	0,994	0,977
$7\pi/16$	0,500	0,903	0,962	0,994	1,000	0,994
$\pi/2$	0,500	0,854	0,927	0,977	0,994	1,000

(b) S_{T_2}

$\phi \backslash \phi$	0,0	$\pi/4$	$5\pi/16$	$3\pi/8$	$7\pi/16$	$\pi/2$
0,0	0,999 0,001	0,922 0,422	0,873 0,373	0,801 0,301	0,716 0,216	0,444 0,056
$\pi/4$	0,924 0,424	0,999 0,001	0,992 0,006	0,970 0,026	0,942 0,039	0,794 0,060
$5\pi/16$	0,872 0,372	0,992 0,006	0,999 0,001	0,992 0,006	0,967 0,005	0,228 0,699
$3\pi/8$	0,784 0,284	0,960 0,016	0,988 0,002	0,999 0,001	0,997 0,003	0,650 0,327
$7\pi/16$	0,730 0,230	0,933 0,030	0,968 0,006	0,991 0,003	0,995 0,005	0,548 0,446
$\pi/2$	0,541 0,041	0,415 0,439	0,394 0,533	0,402 0,575	0,422 0,572	0,631 0,369

(c) S_{T_4}

$\phi \backslash \phi$	0,0	$\pi/4$	$5\pi/16$	$3\pi/8$	$7\pi/16$	$\pi/2$
0,0	0,999 0,001	0,834 0,334	0,536 0,036	0,497 0,003	0,500 0,000	0,498 0,002
$\pi/4$	0,873 0,373	0,923 0,077	0,529 0,457	0,468 0,476	0,479 0,424	0,477 0,377
$5\pi/16$	0,543 0,043	0,614 0,372	0,967 0,033	0,914 0,072	0,897 0,065	0,873 0,054
$3\pi/8$	0,499 0,001	0,465 0,479	0,923 0,063	0,991 0,009	0,986 0,008	0,978 0,001
$7\pi/16$	0,500 0,000	0,493 0,410	0,892 0,067	0,986 0,008	0,991 0,009	0,992 0,002
$\pi/2$	0,501 0,001	0,460 0,394	0,843 0,084	0,974 0,003	0,989 0,005	0,994 0,006

TAB. 15.1 – Mesures de similarité simulées, $\phi \in [0, \pi/4, 5\pi/16, 3\pi/8, 7\pi/16, \pi/2]$: (a) les valeurs théoriques, (b) à l'ordre 2 : S_{T_2} et (c) à l'ordre 4 : S_{T_4} . En gras la valeur absolue de la différence entre la valeur de similarité calculée et la valeur théorique.

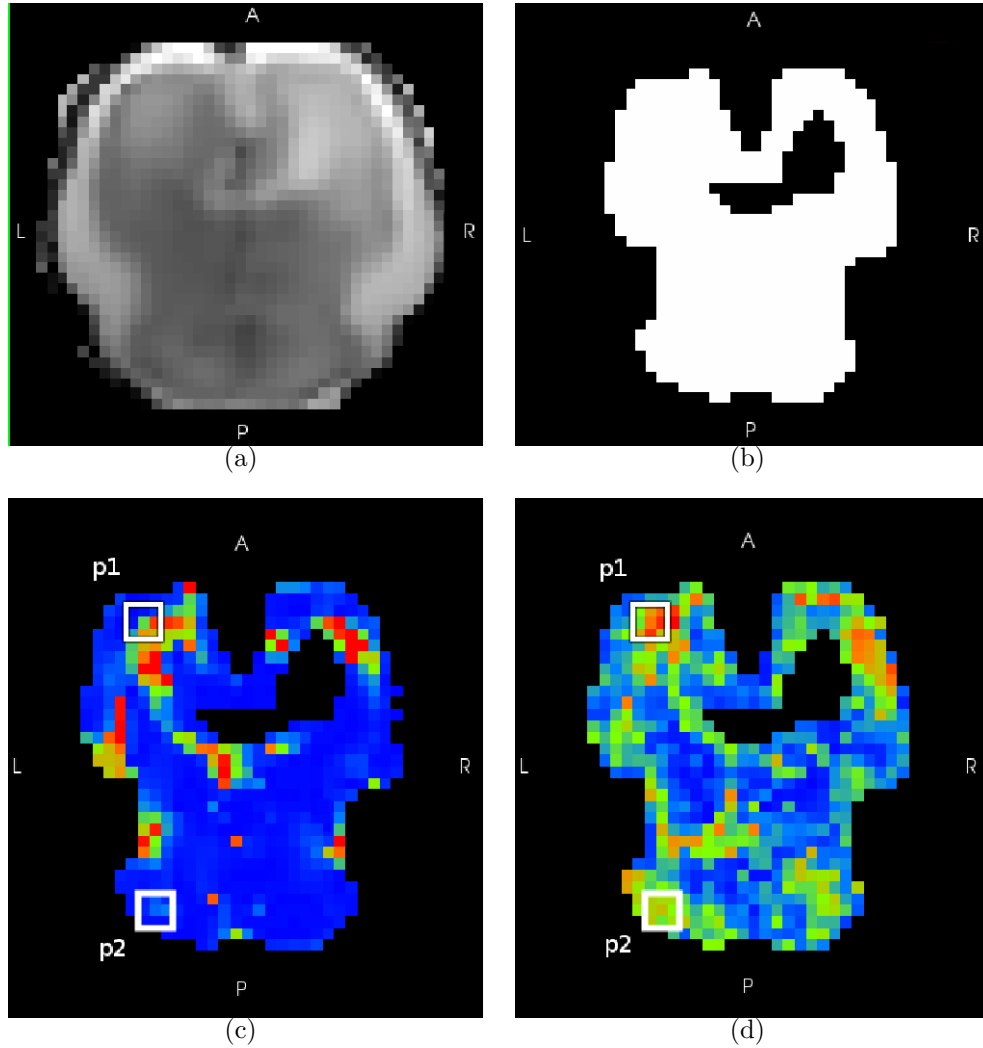


FIG. 15.3 – Résultats de la mesure de similarité sur des images de fœtus : (a) image pondérée en T_2 , (b) masque de la substance blanche où les maxima des tenseurs de diffusion peuvent être calculés, (c) S_{T2} , et (d) S_{T4} . Le code des couleurs va du bleu (un changement) au rouge (aucun changement).

une résolution spatiale de $2,3 \times 2,3 \times 2,3 \text{ mm}^3$ et la taille des images du second examen est de $48 \times 53 \times 36$ avec une résolution spatiale de $1,9 \times 1,9 \times 1,9 \text{ mm}^3$. La méthode de détection de changements requiert que les deux examens soient dans le même référentiel. Pour ce faire, une transformation affine entre deux images pondérées en T_2 est estimée en utilisant l'information mutuelle.

Cette étude porte sur les changements d'organisation de la substance blanche. Une première étape consiste à segmenter cette zone d'intérêt (*i.e.*, la substance blanche). En se basant sur les travaux proposés dans [1], une analyse de variance ANOVA est utilisée en chaque voxel pour déterminer si un modèle de diffusion d'ordre zéro est suffisant pour modéliser convenablement le signal de diffusion (*i.e.*, correspondant à une diffusion isotropique associée

Vers un suivi longitudinal avec des modèles d'ordre supérieur

Détection de changements d'orientations des fibres par une mesure de similarité

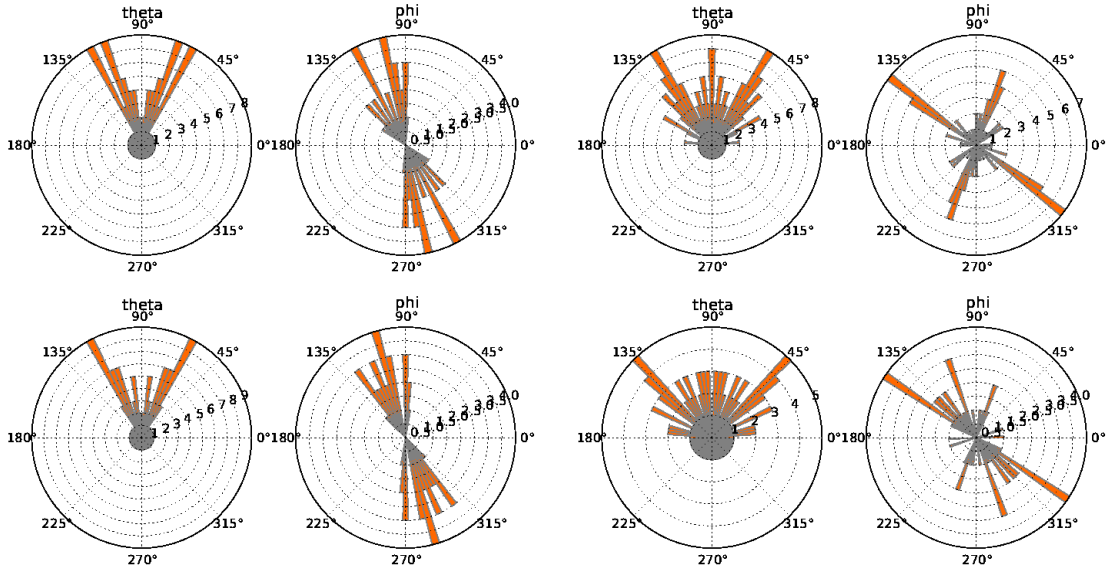


FIG. 15.4 – Distributions angulaires à l'emplacement p_1 . Ligne du haut : tenseur ODF d'ordre 4 et ligne du bas : tenseur d'ordre 2. Les mesures de cohérences associées sont respectivement 0,39 et 0,63.

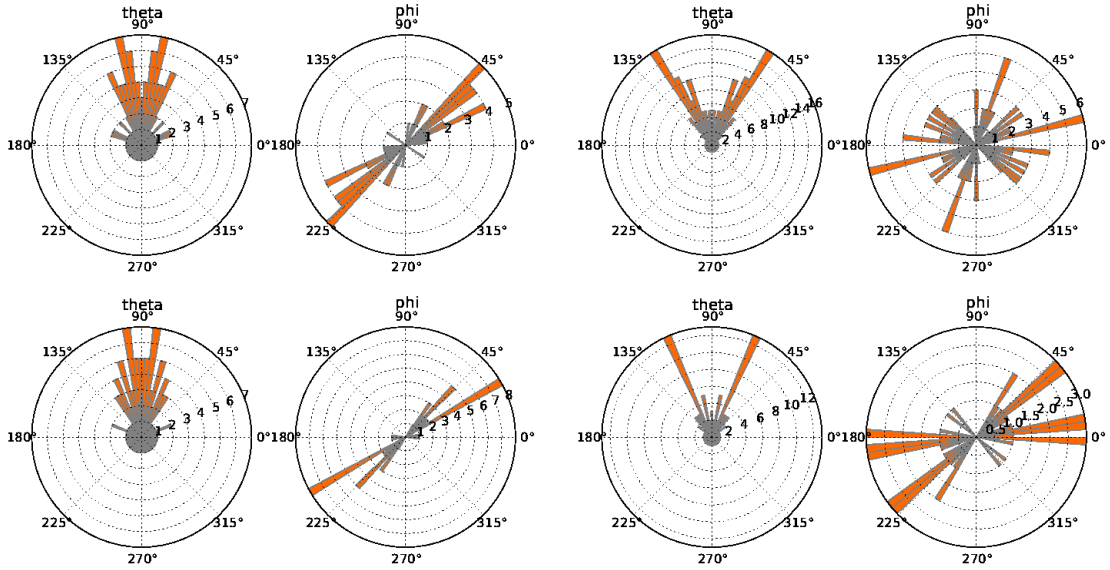


FIG. 15.5 – Distributions angulaires à l'emplacement p_2 . Ligne du haut : tenseur ODF d'ordre 4 et ligne du bas : tenseur d'ordre 2. Les mesures de cohérences associées sont respectivement 0,46 et 0,94.

aux régions présentant du liquide céphalo-rachidien) ou si un modèle d'ordre supérieur est préférable. Après cette première étape, les mesures de cohérence S_{T_2} et S_{T_4} pour des tenseurs d'ordre 2 ou des tenseurs ODF d'ordre 4 sont calculées (cf. figure 15.3-c,d). Les deux méthodes donnent des résultats similaires avec toutefois quelques variations intéressantes. Tout d'abord, la mesure S_{T_4} permet de quantifier plus finement les évolutions (transitions

rouge-vert). Ensuite, une nouvelle détection apparaît en bas à gauche de la figure 15.3-d (en p_2). En utilisant un diagramme de rose (*i.e.*, un histogramme sphérique), les distributions angulaires aux emplacements p_1 et p_2 sont présentées en figure 15.4 et 15.5. A l'emplacement p_1 , un modèle permettant de modéliser un seul faisceau est suffisant dans le premier examen, alors qu'un deuxième mode apparaît dans le deuxième examen. Ainsi, le modèle d'ordre 4 donne une mesure de cohérence plus pertinente. A l'emplacement p_2 , deux et trois maxima sont détectés dans le premier et deuxième examen respectivement. Le modèle d'ordre 2 n'est alors plus du tout adapté et donne un score peu significatif.

Conclusion partielle

Deux chaînes de traitement pour la détection de changements dans des séquences longitudinales d'IRM du tenseur de diffusion ont été proposées pour comparer l'apport d'un modèle d'ordre 4 par rapport à un modèle d'ordre 2. L'utilisation du test statistique multivarié proposé dans le chapitre 6.2.2 avec des tenseurs d'ordre 2 ou 4 donne des résultats similaires. Aucune amélioration majeure n'a été observée et permettrait de motiver ici le choix d'un modèle d'ordre supérieur. Une mesure de similarité plus simple, directement construite sur les directions de maxima de l'ODF, a alors été proposée. Elle permet de montrer que seuls les tenseurs d'ordre 4 donnent des résultats satisfaisant quand les contributions de plusieurs faisceaux sont mesurées en un même voxel. Dans cette partie, nous avons également proposé une méthode de réorientation des tenseurs d'ordre 4 permettant de réorienter chaque faisceau indépendamment. Les résultats présentés dans cette partie sont toutefois très préliminaires et demandent d'autres évaluations.

Bibliographie

- [1] D. ALEXANDER, G. BARKER et S. ARRIDGE : Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331 – 340, 2002.
- [2] A. BARMPOUTIS, B. JIAN et B. C. VEMURI : Adaptive kernels for multi-fiber reconstruction. *In International Conference on Information Processing in Medical Imaging, Williamsburg, IPMI '09*, p. 338 – 349, 2009.
- [3] A. BARMPOUTIS et B. C. VEMURI : Groupwise registration and atlas construction of 4th-order tensor fields using the R+ riemannian metric. *In Medical Image Computing and Computer Assisted Intervention, London*, vol. 5761 de *LNCS*, p. 640 – 647, 2009.
- [4] I. DHILLON et D. MODHA : Concept decompositions for large sparse text data using clustering. *Machine Learning*, 42(1-2):143 – 175, 2001.
- [5] T. DHOLLANDER, W. VAN HECKE, F. MAES, S. SUNAERT et P. SUETENS : Spatial transformations of high angular resolution diffusion imaging data in q-space. *In Medical Image Computing and Computer Assisted Intervention, Workshop, Pékin, MICCAI '10*, p. 73 – 83, 2010.
- [6] A. GRIGIS, F. RENARD, V. NOBLET, C. HEINRICH, F. HEITZ et J. ARMSPACH : A new high order tensor decomposition : application to reorientation. *In IEEE International Symposium on Biomedical Imaging, Chicago, ISBI '11*, p. 258 – 261, 2011.
- [7] T. HICKEY, Q. JU et M. van EMDEN : Interval arithmetic : from principles to implementation. *Journal of the ACM*, 48:1038 – 1068, 1999.
- [8] X. HONG, L. ARLINGHAUS et A. ANDERSON : Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging data. *Magnetic Resonance in Medicine*, 61:1520 – 1527, 2009.
- [9] J. A. NELDER et R. MEAD : A simplex method for function minimization. *The Computer Journal*, 7(4):308 – 313, 1965.
- [10] E. OUBEL, M. KOOB, C. STUDHOLME, J. DIETEMANN et F. ROUSSEAU : Reconstruction of scattered data in fetal diffusion MRI. *Medical Image Analysis*, 16(1), 2012.
- [11] F. RENARD, V. NOBLET, C. HEINRICH et S. KREMER : Reorientation strategies for high order tensors. *In IEEE International Symposium on Biomedical Imaging, Rotterdam, ISBI '10*, p. 1185 – 1188, 2010.

- [12] WU, YC, FIELD, AS, CHUNG, MK, BADIE, B et ALEXANDER, AL : Quantitative analysis of diffusion tensor orientation : Theoretical framework. *Magnetic Resonance in Medicine*, 52(5):1146–1155, 2004.

Conclusion générale et perspectives

Synthèse des travaux

LORS de ce travail de thèse, nous avons développé de nouvelles approches statistiques pour la détection de changement dans des séquences temporelles d'images de diffusion, avec pour objectif de découvrir de nouveaux marqueurs pour le suivi longitudinal de maladies neurodégénératives. Cette modalité étant relativement récente, de nombreux problèmes ont été résolus en amont, pour permettre la mise en place des outils statistiques de détection de changements adaptés à l'imagerie du tenseur de diffusion. Une première contribution apportée par ce travail de thèse concerne la meilleure prise en compte de la nature tensorielle des modèles d'ordre 2 lors d'un suivi longitudinal. Les tests statistiques multivariés utilisés tiennent compte de la nature des espaces associés (matrices définies positives). Une seconde contribution concerne l'extension aux tenseurs d'ordre 4 de ces tests statistiques et d'une mesure de similarité, initialement proposée pour des tenseurs d'ordre 2, qui permet de quantifier l'évolution de l'organisation de la substance blanche. Une troisième contribution concerne la gestion des voisinages sur lesquels les tests sont menés, avec en particulier la conception de tests statistiques sur des faisceaux de fibres ou la mise en place de procédures de bootstrap pour tenir compte de la variabilité du signal de diffusion.

D'autres contributions méthodologiques ont également été proposées, notamment au niveau du modèle tensoriel d'ordre 4. Ce modèle permet de tenir compte de la contribution de plusieurs faisceaux. La complexité de ce modèle a mené au développement d'une nouvelle stratégie de réorientation, nécessaire lors de l'application d'une transformation géométrique à un champ de tenseur, permettant de réorienter chaque faisceau indépendamment. Une autre contribution concerne l'évaluation des performances des méthodes de recherche des maxima d'une ODF. Nous avons proposé une amélioration des méthodes analytique et différentielle en ayant recours pour la résolution des systèmes d'équations à des méthodes issues de l'analyse par intervalle.

Enfin, un point important de ce travail a été la mise en place d'un cadre de validation des différentes chaînes de traitement proposées. En l'absence de vérité terrain, le recours à diverses simulations a permis de valider ces stratégies. Ces dernières ont ensuite été appliquées

sur des images d'une base de patients atteints de SEP et d'une base de patients atteints de NMO. L'étude de la base de données constituée d'examens longitudinaux de 12 patients atteints de sclérose en plaques a montré la capacité de la méthode à détecter des modifications au sein de lésions n'évoluant pas en IRM conventionnelle. L'étude de la base de données constituée d'examens longitudinaux de 11 patients atteints de neuromyéélite optique a montré qu'il se produisait des changements au sein de la substance blanche d'apparence normale qui pouvaient être corrélés avec l'évolution des capacités cognitives ou motrices des patients. Ces méthodes de détection de changements en IRMd pourraient ainsi offrir de nouveaux biomarqueurs pour le suivi des pathologies neurodégénératives.

Perspectives

LE travail entrepris lors de cette thèse et présenté dans ce mémoire laisse place à de nombreuses perspectives afin d'améliorer les prétraitements, d'améliorer et de généraliser les méthodes proposées pour la détection de changements, ou encore de mener des études plus poussées des données cliniques et des résultats associés.

Recalage des images : Les méthodes de détection de changements proposées dépendent du recalage mis en jeu lors de l'alignement spatial des images à comparer. Par exemple, l'utilisation inappropriée d'un recalage déformable risque de compenser tous les changements d'intérêt. Le choix de la métrique a aussi un effet sur les résultats. Lors de cette thèse, les tentatives de recalage basé sur des images anatomiques n'ont pas toujours permis de mettre en correspondance de façon satisfaisante les structures de la substance blanche. Un recalage basé sur la carte de FA a donné de meilleurs résultats et a été retenu lors de cette thèse. Toutefois, les chaînes de traitement proposées pourraient être améliorées en considérant un recalage des tenseurs d'ordre 2 [6] ou 4 [1], ou un recalage conjoint entre des images anatomiques et des images d'indice représentant les propriétés de la substance blanche. L'utilisation d'un recalage symétrique pourrait également améliorer les résultats de nos chaînes de traitements [5]. Un tel recalage permet notamment d'être indépendant au choix arbitraire d'une image de référence.

Association des faisceaux : Lors de l'étude de patients atteints de SEP et de NMO, la structure des faisceaux de fibres a été supposée inchangée entre les deux examens. Or, pour certaines pathologies, cette hypothèse n'est plus vérifiée (par exemple lors du suivi de l'évolution de tumeurs). Dès lors, il s'agirait de développer une méthode permettant la mise en correspondance entre les différents faisceaux de chaque examen [7]. De plus, la méthode de génération de population de tenseurs intégrant un *a priori* anatomique sur les fibres de la substance blanche dépend de l'algorithme de tractographie. Une tractographie déterministe exploitant un modèle d'ordre supérieur ou une tractographie probabiliste pourraient améliorer la qualité de l'*a priori* anatomique [4, 2].

Le modèle d'ordre 4 : De nombreux problèmes méthodologiques inhérents à la complexité du modèle de tenseur d'ordre 4 sont encore ouverts. Nous avons vu qu'il était possible d'esti-

mer des tenseurs d'ordre 4 définis positifs, mais aucune contrainte de positivité n'est imposée lors des tests statistiques. De plus, il serait intéressant de mener des tests sur des quantités interprétables des tenseurs d'ordre 4 (comme les trois valeurs propres d'un tenseur d'ordre 2 qui sont reliées à la diffusion dans les trois directions de l'espace). La mesure de similarité proposée pourrait également être améliorée en tenant compte de la distribution des orientations des faisceaux. Un test statistique, plus robuste, permettrait alors de comparer deux examens. Finalement, la validation des méthodes reposant sur le modèle du tenseur d'ordre 4 est une chose difficile, car une pathologie ne va pas nécessairement s'exprimer dans une région de croisement de fibres. Continuer de collaborer avec des cliniciens permettrait de cibler les pathologies d'intérêt pour les modèles d'ordre supérieur.

Analyse de groupe : Dans ce mémoire de thèse, un expert humain a établi une corrélation entre les détections obtenues par les chaînes de traitement et les signes cliniques pour chaque patient. Cette étude a permis de dégager de grandes tendances. Pour confirmer ces premiers indices et afin d'évaluer le caractère prédictif de l'imagerie de diffusion, une analyse d'un troisième examen et des scores cliniques associés pourraient être entrepris. Des études statistiques de groupe ou de groupe longitudinales pourraient également permettre d'expliquer qualitativement ces tendances :

- Une première approche consiste à exploiter les cartes de statistiques obtenues à partir du suivi de chaque patient et d'étudier la liaison entre ces mesures et un ensemble de variables explicatives (dans notre cas, un score clinique, mais de manière générale tout ce qui est pertinent pour l'étude : âge, sexe, ...). Dans une étude préliminaire nous avons utilisé le modèle linéaire général pour chercher des corrélations avec une variable explicative (l'EDSS ou le BRB-N). Néanmoins, cette stratégie a des limites. En effet, peut-on vraiment faire l'hypothèse gaussienne sur les cartes de statistique ? De plus, les observations sont-elles réellement indépendantes ? En effet, dans le cadre du suivi longitudinal d'une pathologie, on s'attend à obtenir des données observées corrélées entre elles.
- Une deuxième approche permet de modéliser des observations non-indépendantes. Pour ce faire, un modèle linéaire généralisé mixte doit être utilisé. Ce modèle est particulièrement adapté lorsque des mesures sont répétées (c'est exactement le cas lors d'une étude de groupe longitudinale). Comme pour le modèle linéaire généralisé, ce modèle tient compte de facteurs fixes (pour toute la population) mais également de facteurs aléatoires (dépendant de l'individu). Dans la littérature une étude longitudinale de groupe sur des patients atteints de sclérose en plaques, reposant sur ce modèle a été menée [3]. Des mesures fonctionnelles (les variations de MD le long du corps calleux ou de la diffusion parallèle le long du faisceau pyramidal droit) sont mesurées dans le temps et ajoutées au modèle comme des facteurs fixes. Le vecteur d'observations contient un score clinique mesurant la dégradation des fonctions cognitives des patients (le PASAT ou, en anglais, *functional predictors and the paced auditory serial additional test*). Cet indicateur du stade de la maladie est également mesuré dans le temps. Puisque ce modèle offre la possibilité de spécifier une distribution et une struc-

ture de covariance appropriées pour les termes d'erreur, il serait intéressant de mener une étude pour sélectionner la meilleure modélisation possible.

Bibliographie

- [1] A. BARMPOUTIS et B. C. VEMURI : Groupwise registration and atlas construction of 4th-order tensor fields using the R+ riemannian metric. *In Medical Image Computing and Computer Assisted Intervention, London*, vol. 5761 de *LNCS*, p. 640 – 647, 2009.
- [2] M. DESCOTEAUX, R. DERICHE, T. R. KNOSCHE et A. ANWANDER : Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transaction on Medical Imaging*, 28(2):269 – 286, 2009.
- [3] J. GOLDSMITH, C. CRAINICEANU, B. CAFFO et D. REICH : Longitudinal penalized functional regression for cognitive outcomes on neuronal tract measurements. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 61(3):453 – 469, 2012.
- [4] J. PONTABRY et F. ROUSSEAU : Probabilistic tractography using q-ball modeling and particle filtering. *In Medical Image Computing and Computer Assisted Intervention, Toronto*, vol. 6892 de *LNCS*, p. 209 – 216, 2011.
- [5] W. K. THOMPSON et D. HOLLAND : Bias in tensor based morphometry stat-ROI measures may result in unrealistic power estimates. *NeuroImage*, 57(1):1 – 4, 2011.
- [6] B. YEO, T. VERCAUTEREN, P. FILLARD, X. PENNEC, P. GOLLAND, N. AYACHE et O. CLATZ : DTI registration with exact finite-strain differential. *In IEEE International Symposium on Biomedical Imaging, Paris*, ISBI '08, p. 700 – 703, 2008.
- [7] O. ZVITIA, A. MAYER, R. SHADMI, S. MIRON et H. K. GREENSPAN : Co-registration of White Matter Tractographies by Adaptive-Mean-Shift and Gaussian Mixture Modeling. *IEEE Transactions on Medical Imaging*, 29(1):132 – 145, 2010.

Annexes

Annexe A : expression de la base des harmoniques sphériques modifiée à l'ordre 4

Les 15 premiers termes de la base des harmoniques sphériques modifiée sont donnés par :

$$\begin{aligned}
Y_1(\theta, \phi) &= \frac{1}{2\sqrt{\pi}} \\
Y_2(\theta, \phi) &= \frac{\sqrt{15}}{4\sqrt{\pi}}(\cos^2(\phi) - \sin^2(\phi))\sin^2(\theta) \\
Y_3(\theta, \phi) &= \frac{\sqrt{15}}{2\sqrt{\pi}}\cos(\phi)\cos(\theta)\sin(\theta) \\
Y_4(\theta, \phi) &= \frac{\sqrt{15}}{4\sqrt{\pi}}(3\cos^2(\theta) - 1) \\
Y_5(\theta, \phi) &= \frac{\sqrt{15}}{2\sqrt{\pi}}\sin(\phi)\cos(\theta)\sin(\theta) \\
Y_6(\theta, \phi) &= \frac{\sqrt{15}}{2\sqrt{\pi}}\cos(\phi)\sin(\phi)\sin^2(\theta) \\
Y_7(\theta, \phi) &= \frac{3}{16}\sqrt{\frac{35}{\pi}}(\sin^4(\phi) - 6\cos^2(\phi)\sin^2(\phi) + \cos^4(\phi))\sin^4(\theta) \\
Y_8(\theta, \phi) &= \frac{3}{4}\sqrt{\frac{35}{2\pi}}(\cos^2(\phi) - 3\sin^2(\phi))\cos(\phi)\cos(\theta)\sin^3(\theta) \\
Y_9(\theta, \phi) &= \frac{3}{8}\sqrt{\frac{5}{\pi}}(7\cos^2(\theta) - 1)(\cos^2(\phi) - \sin^2(\phi))\sin^2(\theta) \\
Y_{10}(\theta, \phi) &= \frac{3}{4}\sqrt{\frac{5}{2\pi}}(7\cos^2(\theta) - 3)\cos(\phi)\cos(\theta)\sin(\theta) \\
Y_{11}(\theta, \phi) &= \frac{3}{16}\sqrt{\frac{1}{\pi}}(3 - 30\cos^2(\theta) + 35\cos^4(\theta)) \\
Y_{12}(\theta, \phi) &= \frac{3}{4}\sqrt{\frac{5}{2\pi}}(7\cos^2(\theta) - 3)\sin(\phi)\cos(\theta)\sin(\theta) \\
Y_{13}(\theta, \phi) &= \frac{3}{4}\sqrt{\frac{5}{\pi}}(7\cos^2(\theta) - 1)\cos(\phi)\sin(\phi)\sin^2(\theta) \\
Y_{14}(\theta, \phi) &= \frac{3}{4}\sqrt{\frac{35}{2\pi}}(3\cos^2(\phi) - \sin^2(\phi))\sin(\phi)\cos(\theta)\sin^3(\theta) \\
Y_{15}(\theta, \phi) &= \frac{3}{4}\sqrt{\frac{35}{\pi}}(\cos^2(\phi) - \sin^2(\phi))\cos(\phi)\sin(\phi)\sin^4(\theta)
\end{aligned}$$

Annexe B : calcul de l'ODF piquée

Calcul des $A'_{l(j)}(\alpha) = \int_{-1}^1 (1 - \alpha t^2)^{-1/2} P_{l(j)}(t) dt$ jusqu'à l'ordre 4 :

$$A'_0(\alpha) = \frac{1}{\alpha^{1/2}} 2 \arcsin(\sqrt{\alpha})$$

$$A'_2(\alpha) = \frac{-1}{2\alpha^{3/2}} [(-3 + 2\alpha) \arcsin(\sqrt{\alpha}) + 3\sqrt{1-\alpha}\sqrt{\alpha}]$$

$$A'_4(\alpha) = \frac{1}{32\alpha^{5/2}} [(105 - 120\alpha + 24\alpha^2) \arcsin(\sqrt{\alpha}) + (-105 + 50\alpha)\sqrt{1-\alpha}\sqrt{\alpha}]$$

Annexe C : calcul de la matrice de changement de base à l'ordre 4

Cette annexe donne l'expression de la matrice de changement de base entre la base des polynômes homogènes et la base des harmoniques sphériques à l'ordre 4. Tous les calculs ont été réalisés avec Matlab (calcul symbolique) et peuvent être étendus pour un ordre quelconque. Dans le cas présent, $\mathbf{M}$ est une matrice 15×15 qu'il faut calculer en remplaçant les Y_j par leurs expressions données en annexe-A. Rappelons également qu'intégrer sur Ω (l'angle solide) signifie intégrer sur la sphère. On peut donc remplacer l'intégration sur Ω par $d\Omega = \sin(\theta)d\theta d\phi$. On obtient finalement :

$$\mathbf{M}_4 = \begin{pmatrix} a\pi & 0 & b\pi & 0 & a\pi & 0 & 0 & 0 & 0 & b\pi & 0 & b\pi & 0 & 0 & a\pi \\ 0 & 0 & -c\pi & 0 & -c\pi & 0 & 0 & 0 & 0 & c\pi & 0 & 0 & 0 & 0 & c\pi \\ 0 & 0 & 0 & 0 & 0 & d\pi & 0 & d\pi & 0 & 0 & 0 & 0 & d\pi & 0 & 0 \\ e\pi & 0 & f\pi & 0 & -f\pi & 0 & 0 & 0 & 0 & f\pi & 0 & -e\pi & 0 & 0 & -f\pi \\ 0 & d\pi & 0 & d\pi & 0 & 0 & 0 & 0 & 0 & 0 & d\pi & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & d\pi & 0 & d\pi & 0 & 0 & 0 & 0 & d\pi & 0 \\ 0 & 0 & 0 & 0 & g\pi & 0 & 0 & 0 & 0 & 0 & 0 & -h\pi & 0 & 0 & g\pi \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -i\pi & 0 & 0 & 0 & 0 & j\pi & 0 & 0 \\ 0 & 0 & -k\pi & 0 & l\pi & 0 & 0 & 0 & 0 & k\pi & 0 & 0 & 0 & 0 & -l\pi \\ 0 & 0 & 0 & 0 & 0 & m\pi & 0 & -n\pi & 0 & 0 & 0 & 0 & -n\pi & 0 & 0 \\ o\pi & 0 & -p\pi & 0 & q\pi & 0 & 0 & 0 & 0 & -p\pi & 0 & r\pi & 0 & 0 & q\pi \\ 0 & m\pi & 0 & -n\pi & 0 & 0 & 0 & 0 & 0 & 0 & -n\pi & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & s\pi & 0 & -t\pi & 0 & 0 & 0 & 0 & -t\pi & 0 \\ 0 & 0 & 0 & -j\pi & 0 & 0 & 0 & 0 & 0 & 0 & i\pi & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -u\pi & 0 & 0 & 0 & 0 & u\pi & 0 \end{pmatrix} \quad (15.6)$$

Avec :

$$\begin{aligned} a &= 5081767996463981/22517998136852480 & b &= 5081767996463981/11258999068426240 \\ c &= 1230100176224677/4925812092436480 & d &= 1230100176224677/2462906046218240 \\ e &= 5681589342988187/19703248369745920 & f &= 5681589342988187/39406496739491840 \\ g &= 93950452832135/1477743627730944 & h &= 93950452832135/246290604621824 \\ i &= 996495034397243/1847179534663680 & j &= 996495034397243/5541538603991040 \\ k &= 213059600362057/738871813865472 & l &= 213059600362057/4433230883192832 \\ m &= 71740899149003/263882790666240 & n &= 71740899149003/351843720888320 \\ o &= 1905662998673993/22166154415964160 & p &= 1905662998673993/7388718138654720 \\ q &= 1905662998673993/59109745109237760 & r &= 1905662998673993/29554872554618880 \\ s &= 213059600362057/369435906932736 & t &= 213059600362057/2216615441596416 \\ u &= 93950452832135/369435906932736 \end{aligned} \quad (15.7)$$

Annexe D : Algorithme de recherche des zéros d'un polynôme par analyse par intervalles

LEXIQUE :

rechercheIntervalle : **fct** : recherche des zéros avec une optimisation par intervalles

entrée :

D : **Struct Intervalle** : domaine de recherche

f : **fct** : le polynôme dont on cherche les zéros

SIZE_MIN : **float** : taille minimale en dessous de laquelle une boîte n'est plus subdivisée

local :

l1 : **Liste ent** : liste courante

l2 : **Liste ent** : liste résultat

b : **Struct Intervalle** : boîte de recherche

pré :

fInclusion(f,b) : fonction d'inclusion qui permet de définir un encadrement de f sur b

subdivise(b) : fonction subdivisant la boîte b en quatre sous-boîtes de plus petites tailles et de même taille

INSTRUCTIONS :

Fonction rechercheIntervalle(e : D,f,SIZE_MIN) :

```

    l1 ← initListe()
    l2 ← initListe()
    l1.rajouter(D)
    Tant que ( non l1.estVide()) faire
        b ← l1.enlever()
        Si (0 in fInclusion(f,b)) Alors
            Si (b.size() ≤ SIZE_MIN) Alors
                l2.ajouter(b)
            Sinon
                l1.ajouter( subdivise(b) )
        Fin Si
    Fin Si
    Fait
    Retourner l2
Fin
```

Annexe E : Liste de mes publications

Articles dans des revues internationales à comité de lecture

A. Grigis, V. Noblet, F. Blanc, F. Heitz, J. de Seze, S. Kremer, J.-P. Armspach : “Longitudinal Change Detection : Inference on the Diffusion Tensor Along White Matter Pathways”, Medical Image Analysis, en cours de révision.

A. Grigis, V. Noblet, F. Heitz, F. Blanc, J. de Seze, S. Kremer, L. Rumbach, J.-P. Armspach : “Longitudinal Change Detection in Diffusion MRI using Multivariate Statistical Testing on Tensors”, NeuroImage 60(4), pp.2206–2221, 2012.

Conférences internationales avec comité de lecture et actes

A. Grigis, V. Noblet, F. Blanc, F. Heitz, J. de Seze, J.-P. Armspach : “Longitudinal Change Detection : Inference on the Diffusion Tensor Along White-Matter Pathways”, MICCAI 2011, Medical Image Computing and Computer Assisted Intervention, Toronto, Canada, pp. 1–8, Springer-Verlag, Lecture Notes in Computer Science, Vol. 6892, septembre 2011.

A. Grigis, F. Renard, V. Noblet, C. Heinrich, F. Heitz, J.-P. Armspach : “A new high order tensor decomposition : application to reorientation”, ISBI 2011, 8th IEEE International Symposium on Biomedical Imaging, Chicago, États-Unis, pp. 258–261, mars 2011.

A. Grigis, V. Noblet, F. Renard, F. Heitz, J.-P. Armspach, L. Rumbach : “Change Detection in Diffusion MRI using Multivariate Statistical Testing on Tensors”, MICCAI 2010, Medical Image Computing and Computer Assisted Intervention, Pékin, Chine, pp. 117–124, Springer-Verlag, Lecture Notes in Computer Science, Vol. 6362, Num. 2, septembre 2010.

F. Renard, V. Noblet, A. Grigis, C. Heinrich, S. Kremer : “Comparison of interpolation methods for angular resampling of diffusion weighted images”, IPTA 2011, International Conference on Image Processing Theory, Tools and Applications, Paris, France, pp. 207–211, juillet 2010.

Conférences nationales avec comité de lecture et actes

A. Grigis, V. Noblet, F. Heitz, J.-P. Armspach : “On the integration of spatial neighborhood information for detecting longitudinal changes in MRI sequences”, LIVIM 2011, Strasbourg, France, décembre 2011 (présentation orale).

A. Grigis, V. Noblet, F. Heitz, L. Rumbach, J.-P. Armspach : “Détection de changements en IRM du tenseur de diffusion : application au suivi longitudinal de la sclérose en plaques”, RITS 2011, Recherche en Imagerie et Technologies pour la Santé, Rennes, France, 874 pages, avril 2011 (présentation orale).

Approches statistiques pour la détection de changements en IRM de diffusion. Application au suivi longitudinal de pathologies neuro-dégénératives.

L'IRM de diffusion (IRMd) est une modalité d'imagerie médicale qui suscite un intérêt croissant dans la recherche en neuroimagerie. Elle permet d'apporter in vivo des informations nouvelles sur les micro-structures locales des tissus. En chaque point d'une acquisition d'IRMd, la distribution des directions de diffusion des molécules d'eau est modélisée par un tenseur de diffusion. La nature multivariée de ces images requiert la conception de nouvelles méthodes de traitement adaptées. Le contexte de cette thèse est l'analyse automatique de changements longitudinaux intra-patient avec pour application le suivi de pathologies neuro-dégénératives. Notre recherche a ainsi porté sur le développement de nouveaux modèles et tests statistiques permettant la détection automatique de changements sur des séquences temporelles d'images de diffusion. Cette thèse a ainsi permis une meilleure prise en compte de la nature tensorielle des modèles d'ordre 2 (tests statistiques sur des matrices définies positives), une extension vers des modèles d'ordre supérieur et une gestion plus fine des voisinages sur lesquels les tests sont menés, avec en particulier la conception de tests statistiques sur des faisceaux de fibres.

IRM de diffusion – tests statistiques – pathologies neuro-dégénératives.

Statistical approaches for change detection in diffusion MRI. Application to the longitudinal follow-up of neuro-degenerative pathologies.

Diffusion MRI is a new medical imaging modality of great interest in neuroimaging research. This modality enables the characterization in vivo of local micro-structures. Tensors have commonly been used to model the diffusivity profile at each voxel. This multivariate data set requires the design of new dedicated image processing techniques. The context of this thesis is the automatic analysis of intra-patient longitudinal changes with application to the follow-up of neuro-degenerative pathologies. Our research focused on the development of new models and statistical tests for the automatic detection of changes in temporal sequences of diffusion images. Thereby, this thesis led to a better modeling of second order tensors (statistical tests on positive definite matrices), to an extension to higher-order models, and to the definition of refined neighborhoods on which tests are conducted, in particular the design of statistical tests on fiber bundles.

Diffusion MRI – statistical tests – neuro-degenerative pathologies.