

N ° d'ordre : 1188



THÈSE

présentée pour obtenir le grade de
Docteur de l'Université de Strasbourg

École Doctorale : Mathématiques, Sciences de l'Information et de l'Ingénieur
Spécialité : Signal, Image, Automatique, Robotique
Traitement du signal et des images

Création et utilisation d'atlas en IRM de diffusion. Application à l'étude des troubles de la conscience.

Soutenue publiquement
le 29 septembre 2011
par

Félix RENARD

Membres du jury :

M.	Michel	DESVIGNES	Rapporteur	GIPSA-Lab, Grenoble
M.	Jean-François	MANGIN	Rapporteur	CEA, Saclay
Mme	Sophie	ACHARD	Examinatrice	GIPSA-Lab, Grenoble
Mme	Armelle	GUILLOU	Examinatrice	IRMA, Strasbourg
M.	Alexandre	VIGNAUD	Invité	Siemens, Saint-Denis
M.	Christian	HEINRICH	Directeur de thèse	LSIIT, Strasbourg
M.	Stéphane	KREMER	Directeur de thèse	LINC, Strasbourg

Équipes d'accueil :

Laboratoire des Sciences de l'Image,
de l'Informatique et de la Télédétection
UMR Uds-CNRS 7005

Laboratoire d'Imagerie
et de Neurosciences Cognitives
UMR Uds-CNRS 7237

Remerciements

Cette thèse s'est déroulée dans le cadre d'un projet de recherche commun à deux laboratoires, le Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT) et le Laboratoire d'Imagerie et de Neurosciences Cognitives (LINC). Un grand merci à Fabrice Heitz, directeur du LSIIT, et à Christian Kelche, directeur du LINC, pour m'avoir accueilli au sein de leur laboratoire respectif. Ce double rattachement a constitué un cadre stimulant et propice pour les recherches en traitement d'images médicales que j'ai menées. Je tiens aussi à remercier Jean-Paul Armspach, responsable de l'équipe Imagerie In Vivo du LINC, pour son accueil, ainsi que pour nos nombreuses discussions.

Cette thèse s'est déroulée dans le cadre d'une convention CIFRE. Je remercie la société Siemens et l'ANRT qui, par leur financement, ont rendu possible le déroulement de ces travaux. Je tiens à remercier les collaborateurs de la société Siemens, et plus particulièrement Alexandre Vignaud, Agnès Malgouyres et Isabelle Ginja, qui m'ont permis d'effectuer ma thèse dans de bonnes conditions.

J'exprime toute ma gratitude aux membres de mon jury de thèse, particulièrement aux rapporteurs, Michel Desvignes et Jean-François Mangin, pour le temps et les efforts consacrés à la lecture de ce mémoire, ainsi qu'à Sophie Achard, Armelle Guillou, et Alexandre Vignaud pour leur contribution. C'est un grand honneur de les compter comme membres de mon jury.

Je tiens à remercier les membres du jury pour les discussions et les différentes interactions scientifiques que j'ai eues avec eux.

Je tiens à remercier chaleureusement mes directeurs de thèse, Christian Heinrich et Stéphane Kremer, pour leur travail et leur encadrement. En plus de leurs compétences scientifiques qui m'ont permis d'avancer, ils m'ont toujours apporté leur enthousiasme et leur soutien, dans une bonne humeur sans faille. Je tiens particulièrement à remercier Christian Heinrich pour le temps consacré à la relecture de ce mémoire.

Un grand merci à Vincent Noblet pour le temps qu'il m'a consacré, ses bons conseils et l'aide qu'il m'a apportée lors des difficultés rencontrées.

Cette thèse a été un parcours riches en rencontres humaines et scientifiques. Je souhaite remercier ici les nombreuses personnes avec lesquelles j'ai collaboré : Simon Chatelin, Caroline Deck, Rémy Willinger, Nathalie Philippi, Frédéric Blanc, Daniel Roquet, Jack Foucher, Edouard Hirsch, Jérôme de Sèze, Izzie Namer, Antoine Grigis, Chantal Delon-Martin et Sophie Achard. Je souhaite aussi remercier le service de réanimation du CHU de Strasbourg, et plus particulièrement Maleka Schenck et Francis Schneider, pour leur disponibilité, leur apport scientifique et leur gentillesse.

Je n'oublie pas non plus l'ensemble des personnes que j'ai pu côtoyer au LINC et au LSIIT : les jeunes chercheurs permanents (Sylvain, François, Vincent M., Julien, Étienne et Nicolas), les personnels administratifs et techniques (Frédérique, Simone, Nathalie), les doctorants (Bessem, Hervé, Matthieu, Stéphanie, Steven, Farouk, Franklin, Giorgos, Antoine, Swati, Slim, Wissam, Mériam, Benjamin, Benoît, Alice, Akram, et Julien P.), les contractuels (Nicolas, Gilberto, François-Xavier et Stan) et les stagiaires. Je tiens à remercier Anthony pour le coup de main

donné à la toute fin.

Un remerciement spécial à Farouk Nouizi pour nos longues parties de dominos le dimanche après midi, ainsi que nos nocturnes pendant la rédaction de thèse. Je tiens aussi à remercier vivement Virginie Pelaèz Muñoz, Amine et Nino Rahmani, Nadia Mercurio, Franklin Tellier, Benoît Caldairou, Thibault Deloze, Sabrina Noël et Momo Lamri, pour tous les bons moments passés à Strasbourg autour d'un verre ou d'une flammekueche.

Je souhaitais remercier la famille Boivin pour sa gentillesse et son soutien. Je tiens à remercier vivement mes parents et ma famille pour l'éducation et le soutien qu'ils m'apportent depuis toujours.

Les derniers mots de ces remerciements sont pour Cécile. Merci de tout coeur d'avoir été à mes côtés pour m'encourager et me soutenir.

Table des matières

Remerciements	
Introduction	
1. Problématique scientifique	1
2. Contexte de la thèse	2
3. Contributions	3
4. Collaborations	3
5. Organisation du mémoire	4
I IRM de diffusion : de la physique aux modèles	7
1. Physique de la diffusion et lien avec le signal IRM de diffusion	11
1. Mouvement brownien et propagateur	11
2. Diffusions libre et restreinte	12
3. Lien entre le signal IRMd et la diffusion	12
2. Outils mathématiques	17
1. Les tenseurs	17
2. Les harmoniques sphériques	28
3. Lien entre tenseurs et harmoniques sphériques	30
4. Conclusion partielle	30
3. Modèles en IRM de diffusion	33
1. Modélisation par tenseur de diffusion	33
2. Modélisation par diffusion apparente	35
3. Modélisation par fonction de distribution de l'orientation	38
4. Conclusion partielle	39
4. Tractographie	41
1. Méthodes de tractographie	41
2. Limitations et validation des méthodes de tractographie	44
3. Conclusion partielle	45
5. Connectivité et connectomes	47
1. Connectivité	47
2. Connectome	49

3.	Conclusion partielle	57
II	Mise en correspondance en IRM de diffusion	61
6.	Recalage et réorientation des modèles locaux	65
1.	Recalage d'images et spécificité rencontrée en IRMd	65
2.	État de l'art	67
3.	Réorientation des ODFs modélisées par des tenseurs d'ordre supérieur	75
4.	Conclusion partielle	85
7.	Mise en correspondance des modèles semi-globaux et globaux	89
1.	Mise en correspondance pour les tractographies	89
2.	Mise en correspondance pour les connectomes	90
3.	Conclusion partielle	91
III	Création d'atlas en IRM de diffusion	95
8.	Atlas en imagerie par résonance magnétique	99
1.	Définition et création d'un atlas	99
2.	Méthodologies proposées pour les modèles locaux	100
3.	Méthodologies proposées pour les modèles semi-globaux	103
4.	Méthodologies proposées pour les modèles globaux	105
5.	Conclusion partielle	106
9.	Modèles et représentations compactes en imagerie médicale	107
1.	Méthodes et algorithmes pour les représentations compactes	107
2.	Modélisation d'une population de référence	113
3.	Conclusion partielle	117
10.	Atlas proposé	119
1.	Création de l'atlas	121
2.	Utilisation de l'atlas	129
3.	Applications	133
4.	Conclusion partielle	135
11.	Applications à l'IRM de diffusion	139
1.	Atlas basé sur la connectivité anatomique et sur les ROIs	139
2.	Atlas basé sur la connectivité anatomique et fonctionnelle	143
3.	Conclusion partielle	147
	Conclusion et perspectives	151

IV Bibliographie	155
Publications de l'auteur	157
Références bibliographiques	159

Table des acronymes

AAL	<i>Automatic anatomical labelling</i>
ADC	<i>Apparent diffusion coefficient</i>
ACP	Analyse par composantes principales
ADL	Analyse discriminante linéaire
CC	Coefficient de corrélation
DKI	<i>Diffusion kurtosis imaging</i>
DO	Différence d'orientations
DSI	<i>Diffusion spectrum imaging</i>
DTI	<i>Diffusion tensor imaging</i>
EEG	Electro-encéphalographie
FA	Fraction d'anisotropie
FS	<i>Finite strain</i>
GA	<i>Generalized anisotropy</i>
HARDI	<i>High angular resolution diffusion imaging</i>
HOT	<i>High order tensor</i>
HS	Harmonique sphérique
IRMd	Imagerie par résonance magnétique de diffusion
IRMf	Imagerie par résonance magnétique fonctionnelle
ODF	<i>Orientation distribution function</i>
f-ODF	<i>fiber-orientation distribution function</i>
d-ODF	<i>diffusion-orientation distribution function</i>
MD	<i>Mean diffusivity</i>
MDS	<i>Multidimensional scaling</i>
NNW	Noyau de type Nadaraya-Watson
PAS	<i>Persistent angular structure</i>
PDF	<i>Probability density function</i>
PM	Petit monde
PME	Propagateur moyen d'ensemble
PPD	<i>Preservation of principal direction</i>
ROI	<i>Region of interest</i>
SA	Stratégie affine
SH	Stratégie de Hilbert
SS	Stratégie spectrale
SVDD	<i>Support vector data description</i> <i>Support vector domain description</i>
SVM	<i>Support vector machine</i>
T2	Tenseur d'ordre 2
T4	Tenseur d'ordre 4
TBSS	<i>Tract based spatial statistics</i>

Introduction

Dans cette introduction, nous présenterons succinctement les travaux effectués dans le cadre de cette thèse. Nous introduirons le contexte médical et le contexte scientifique, ainsi que les différents projets et partenariats établis.

1. Problématique scientifique

Le but de cette thèse a été la création d'atlas pour l'imagerie par résonance magnétique de diffusion (IRMd). L'IRMd est une modalité permettant d'observer indirectement pour la première fois les fibres de substance blanche *in vivo*. L'IRMd est une modalité-clé en raison de son contenu informatif. Elle présente un intérêt en pratique clinique en tant qu'outil de diagnostic et de suivi thérapeutique pour de très nombreuses pathologies. Elle présente également un intérêt en neurosciences en tant qu'outil d'aide à l'étude du développement, du fonctionnement et du vieillissement cérébral.

Les atlas constituent également un outil-clé en imagerie médicale (voir par exemple le livre [TOGA & MAZZIOTTA 2002]). En effet, le développement de l'imagerie médicale conduit à des volumes de données de plus en plus importants et de plus en plus complexes, pour des groupes d'individus de plus en plus nombreux. Les atlas permettent de modéliser cette grande quantité d'information tout en rendant compte de la variabilité inter-individuelle existant au sein d'une population. Cette variabilité est prise en compte dans différents algorithmes de traitement d'images médicales. L'aide au diagnostic et le suivi thérapeutique sont rendus plus performants, par la prise en compte implicite d'une grande base d'images de référence.

La modalité d'IRMd délivre des signaux complexes permettant d'étudier aussi bien des phénomènes à l'échelle du voxel que des mécanismes globaux à l'échelle du cerveau. Dès lors, cette modalité a été largement étudiée au cours de cette dernière décennie pour analyser et comprendre les phénomènes apparaissant pour certaines pathologies. En raison de la complexité de ce type d'imagerie, ainsi que de la grande variabilité inter-individuelle, il a été nécessaire de développer des outils statistiques, basés sur des atlas, permettant de quantifier ces phénomènes pathologiques, et de les différencier des phénomènes normaux.

Le but de cette thèse a été la création d'atlas probabilistes dans un cadre général, et leur application en IRMd pour des patients atteints de troubles de la conscience. Plus précisément, la construction de cet atlas a consisté premièrement en la modélisation probabiliste d'une population de référence supposée saine, deuxièmement en l'identification des différences statistiquement significatives entre un nouvel individu présentant des troubles de la conscience et l'atlas probabiliste. La création de cet atlas doit permettre une meilleure compréhension des mécanismes physiopathologiques sous-tendant les troubles de la conscience.

2. Contexte de la thèse

Cette thèse a été effectuée 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 de traitement d'images médicales (TIM) du Laboratoire d'Imagerie et de Neurosciences Cognitives (LINC, UMR 7237, Université de Strasbourg - CNRS).

Les travaux de l'équipe LSIIT-MIV ont pour objet le développement de modèles, de méthodes et d'algorithmes pour le traitement d'images et de séquences d'images, dans des cadres applicatifs divers (médecine, biologie, astronomie, physique, métrologie). L'équipe LINC-TIM est spécialisée dans le traitement d'images médicales. Cette équipe comprend des chercheurs issus de milieux très variés : des médecins, des physiciens et des traiteurs d'images. La collaboration de ces deux structures, mise en place en 1994, permet le développement d'algorithmes innovants pour répondre à des problématiques médicales.

Cette thèse a également bénéficié d'une collaboration entre les services hospitaliers de réanimation et de radiologie de l'hôpital de Strasbourg-Hautepierre qui fournit les images des patients étudiés dans le cadre de cette thèse. Précisons que l'acquisition des données IRM d'un patient comateux est d'une plus grande complexité que pour n'importe quelle autre pathologie. En effet la réalisation des examens IRM pour un patient comateux exige la présence d'une équipe de réanimation.

Cette thèse et celle d'Hervé Boisgontier [BOISGONTIER 2010] ont été les premiers travaux en traitement de données d'IRMd chez l'homme dans les deux laboratoires, LSIIT et LINC. La thèse d'Hervé Boisgontier a porté sur le suivi longitudinal basé sur des modèles locaux. La présente thèse a pour but la création d'atlas pour la comparaison d'un individu par rapport à une population de référence pour différents types de modèles représentant des phénomènes aussi bien locaux que globaux. Actuellement, plusieurs doctorants et chercheurs travaillent sur ce type de modalité au LINC et au LSIIT.

Du point de vue de la construction d'atlas et de l'analyse des populations, cette thèse prolonge les travaux de Torbjørn Vik [VIK 2004] sur la création d'atlas probabilistes avec des modèles linéaires non gaussiens, et de Matthieu Brucher [BRUCHER 2008] sur l'analyse de formes avec des modèles non linéaires. Les travaux en analyse de formes sont actuellement poursuivis par Giorgos Sfikas dans le cadre de sa thèse.

Cette thèse propose d'élaborer un atlas grâce à des modèles probabilistes non linéaires non gaussiens pour l'étude de différents types de données rencontrés en IRMd.

Durant la période de cette thèse, une restructuration importante de la plate-forme de développement d'algorithmes de traitement de données et d'images médicales a été effectuée. De plus, peu d'outils étaient disponibles pour l'IRMd car les travaux débutaient et l'existant n'était que très partiellement utilisable tel quel en raison des spécificités des données d'IRMd.

Une partie de cette thèse a donc consisté à élaborer les outils nécessaires pour la construction d'atlas, comme l'implantation des modèles ou la mise en correspondance d'objets. Plus précisément, des chaînes de traitements ont été élaborées pour permettre l'étude de populations contenant un grand nombre d'individus.

Au cours de cette thèse, des outils d'aide au diagnostic ont aussi été élaborés. Plus précisément, ces outils permettent d'étudier des cartes statistiques obtenues précédemment à la suite de l'utilisation d'atlas, et d'associer à chaque zone détectée comme pathologique une étiquette anatomique pouvant être extraite de différents atlas anatomiques disponibles sur internet. Ces

outils ont été élaborés en collaboration avec Julien Lamy.

3. Contributions

Les contributions originales de ce travail sont multiples. D'un point de vue bibliographique, différents états de l'art ont été effectués sur des étapes clés de l'élaboration de l'atlas. Plus précisément, les outils mathématiques disponibles pour la modélisation des phénomènes locaux, la mise en correspondance des modèles locaux ainsi que la modélisation lors de la conception d'atlas ont été examinés de manière approfondie.

D'un point de vue méthodologique, nous avons élaboré un atlas original basé sur un modèle probabiliste non linéaire dans la problématique de la comparaison d'un individu par rapport à un groupe de référence. L'originalité de cet atlas est de pouvoir à la fois discriminer un individu (*i.e.* indiquer s'il est pathologique ou non) et analyser les régions pathogènes, grâce au même algorithme du *support vector domain description* (SVDD) [TAX & DUIN 1999]. Cet atlas peut être utilisé pour une grande variété d'objets mathématiques. Nous avons proposé différentes estimations robustes pour la paramétrisation automatique du SVDD. Ces estimations peuvent être utiles pour d'autres domaines que l'imagerie médicale nécessitant le SVDD (par exemple l'imagerie satellitaire [BANERJEE & DIEHL 2006]). Pour l'IRMd, nous avons proposé des améliorations aux algorithmes de mise en correspondance existants pour les modèles locaux.

D'un point de vue applicatif, nous nous sommes particulièrement intéressé à la connectivité anatomique, terme désignant l'étude des connections anatomiques à l'échelle du cerveau. Cette connectivité anatomique permet de rendre compte de divers mécanismes cérébraux, notamment ceux liés à la conscience. Nous avons donc utilisé l'atlas élaboré dans cette thèse pour étudier les troubles de la conscience chez des individus comateux ou ayant été dans le coma. L'originalité des travaux est de pouvoir considérer simultanément l'intégrité anatomique et la connectivité cérébrale.

4. Collaborations

Dans le cadre de cette thèse, nous avons participé à plusieurs collaborations.

La première collaboration a été réalisée avec Simon Chatelin, Caroline Deck et Rémy Willinger (Institut de Mécanique des Fluides et des Solides (IMFS), UMR 7507, Université de Strasbourg - CNRS). Dans le cadre de sa thèse [CHATELIN 2010], Simon Chatelin a modélisé, d'un point de vue mécanique sur le cerveau, les conséquences de chocs subis au niveau de la tête. Les modèles utilisés initialement séparaient le cerveau en parties relativement grossières (une partie crâne, une partie cerveau). Nous avons proposé d'inclure dans ce modèle une information, liée aux fibres de substance blanche, obtenue grâce à l'IRMd. L'intérêt est que les déformations dues aux chocs sont privilégiées dans la direction de ces fibres, ce dont le modèle pourra rendre compte. Les résultats de simulation seront donc plus fidèles à la réalité. Les premiers résultats issus de cette collaboration sont parus dans [CHATELIN et coll. 2011].

Dans le cadre de cette thèse, nous travaillons sur la connectivité anatomique grâce aux séquences d'IRMd. La connectivité anatomique est un domaine en pleine expansion. Parallèlement à ces études, de nombreux chercheurs étudient la connectivité fonctionnelle obtenue grâce à des séquences d'IRM fonctionnelle (IRMf) au repos. Ces deux types de connectivité sont complémentaires. Une collaboration a été mise en place avec Chantal Delon Martin (Grenoble Institut des Neurosciences, GIN - INSERM U836 - Grenoble) et Sophie Achard (Grenoble Images Parole Signal Automatique, GIPSA-Lab - UMR 5216 - Grenoble), dans le but d'une étude des troubles de la conscience chez le patient comateux grâce à la prise en compte con-

jointe des connectivités anatomique et fonctionnelle. Cette prise en compte conjointe permet d'améliorer la performance du diagnostic.

La connectivité fonctionnelle peut être aussi étudiée grâce à des algorithmes basés sur l'analyse en composantes indépendantes. Cette problématique a été abordée par Jack Foucher et Daniel Roquet (INSERM U666 - Strasbourg) toujours dans le cadre des troubles de la conscience. La collaboration avec Jack Foucher et Daniel Roquet a pour but d'étudier les réseaux de connectivité fonctionnelle à l'aide d'atlas. Il s'agit de déterminer les réseaux pathologiques par comparaison à des réseaux d'une population saine de référence. Dans un deuxième temps, le but sera de discriminer les régions anormales à l'intérieur d'un réseau pathologique.

Au cours de cette thèse, nous avons tenu à créer une méthodologie la plus générale possible. Nous collaborons en analyse de formes, et plus spécifiquement sur l'atrophie des hippocampes dans la maladie d'Alzheimer où nos modèles permettent de décrire une population d'intérêt. Ce travail est effectué en collaboration avec Giorgos Sfikas et Christophoros Nikou (Université de Ioannina, Grèce).

En plus des données quantitatives traditionnelles (par exemple l'intensité des voxels), des données qualitatives décrivant les capacités cognitives des individus ont été acquises. Une collaboration est en cours avec Frédéric Blanc et Nathalie Philippi (Service de Neuropsychologie, département de Neurologie, Hôpitaux Universitaires de Strasbourg) pour intégrer dans les atlas ces données supplémentaires. L'étude porte sur l'analyse des formes de structures anatomiques (les hippocampes et les amygdales) dans la maladie d'Alzheimer.

5. Organisation du mémoire

Le mémoire est organisé en trois parties : le choix des descripteurs lors de la création de l'atlas, la mise en correspondance des différents individus, et l'implémentation d'algorithmes de détection adaptés à l'étude de signaux complexes appartenant à des espaces non euclidiens.

La première partie du mémoire est dédiée à la description de différents modèles utilisés en IRM_d. En effet, créer un atlas nécessite avant tout de cibler les différents paramètres d'intérêt. Différents modèles permettent de rendre compte des structures neuronales observées indirectement en IRM_d. On peut classer ces modèles en trois grandes catégories : les descripteurs locaux, semi-globaux et globaux. Les modèles locaux sont à l'échelle du voxel et ne permettent de voir que des anomalies locales. Ils permettent l'étude de structures simples telle la diffusion moyenne comme l'étude de structures complexes tel le croisement de fibres. Pour ce faire, de nombreux modèles mathématiques ont été utilisés, en particulier les tenseurs. Dans le cadre de cette thèse, nous nous sommes d'abord intéressés à des descripteurs scalaires de la diffusion, puis à des modèles plus complexes correspondant à une direction de diffusion privilégiée (tenseurs d'ordre 2) et enfin à des modèles permettant de rendre compte de plusieurs directions de diffusion (tenseurs d'ordre supérieur). Les deux derniers modèles rendent compte respectivement d'un morceau de fibre de myéline et d'un croisement de morceaux de fibres de myéline. Les descripteurs semi-globaux, décrivant des paquets de fibres, permettent de rendre compte de la connexion entre différentes régions corticales du cerveau. Les descripteurs globaux, connus sous le nom de connectogrammes, permettent de rendre compte de phénomènes à l'échelle du cerveau. Ils sont une intégration de toute l'information fournie par les descripteurs semi-globaux. Ces descripteurs globaux peuvent renseigner sur l'intégrité d'une région corticale, ainsi que permettre de décrire des réseaux neuronaux complexes. Remarquons que la compréhension anatomique du cerveau d'un point de vue global permet de mieux comprendre certains phénomènes fonctionnels.

La deuxième partie du mémoire traite de la mise en correspondance en IRM_d. Plus précisément, une fois le choix du descripteur établi, la construction des atlas nécessite de mettre en correspondance spatiale (recaler) les différents jeux de données. Il faut tenir compte de la structure complexe du signal. Lors du recalage, en plus des déplacements spatiaux classiques, une étape de réorientation doit être effectuée. Les recalages et les étapes de réorientation ont été longuement étudiés dans le cadre des tenseurs d'ordre 2. Le recalage pour des tenseurs d'ordre supérieur est plus complexe. Seul le recalage affine a été proposé avec une étape de réorientation non adaptée à des déformations non linéaires tels le cisaillement ou l'étirement. Dans le cadre de cette thèse, nous décrivons un système complet de techniques de réorientation pour une transformation quelconque. Nous proposons principalement deux nouvelles méthodes, basées sur des décompositions tensorielles (la décomposition spectrale et la décomposition *via* le théorème de Hilbert sur les quartiques ternaires), permettant de se ramener à une somme de tenseurs d'ordre 2. Ces différentes techniques ont été comparées, sur des données simulées, aux deux techniques existantes.

Dans la troisième partie du mémoire, nous présentons tout d'abord les différentes méthodes de création d'atlas pour l'IRM_d. Dans un deuxième temps, nous décrivons plusieurs stratégies pour la modélisation probabiliste de population grâce des représentations compactes. Nous nous sommes intéressés spécifiquement à des études « un individu par rapport à un groupe ». En effet, pour les types de coma étudiés, l'étiologie des patients est tellement hétérogène (différents types d'accidents, acquisition de l'IRM suite à l'accident dans des délais variant d'une semaine à un mois,...) qu'une étude « groupe à groupe » n'était pas pertinente. Pour ce faire, nous nous avons considéré des méthodes à noyaux permettant d'apprendre le support d'un ensemble de données supposées normales, et de discriminer des données aberrantes. Ces méthodes sont connues sous le nom de *support vector domain description*. Une fois les données aberrantes détectées globalement, la localisation fine des zones aberrantes (pathologiques) dans les données est effectuée grâce à une approche originale.

Nous proposons pour ces études un modèle original permettant de prendre en compte aussi bien les descripteurs locaux que globaux en IRM_d. Ces données appartenant à des espaces multidimensionnels non euclidiens et de grandes dimensions, nous avons développé des outils statistiques spécifiques, basés sur des méthodes de réduction de dimension, afin de permettre une détection des données aberrantes plus robuste.

La fin de la troisième partie consiste en une présentation des résultats obtenus pour des patients comateux. D'un point de vue médical, cette pathologie reste encore mal connue. L'étude conjointe des connectivités anatomique et fonctionnelle de ces patients permet une meilleure compréhension des mécanismes cérébraux sous-tendant les troubles de la conscience.

Enfin, les conclusions de ce travail sont tirées et des perspectives sont proposées.

Première partie

IRM de diffusion : de la physique aux
modèles

Introduction

La première partie du mémoire a pour objet une brève description de la physique de l'IRMd, ainsi qu'une étude plus détaillée des différents modèles considérés. En effet, construire un atlas nécessite avant tout de cibler les différents paramètres d'intérêt. Différents modèles permettent de rendre compte des structures neuronales observées en IRMd. On peut classer ces modèles en trois grandes catégories : les descripteurs locaux, semi-globaux et globaux.

Les modèles locaux sont à l'échelle du voxel et ne permettent de voir que des anomalies locales. De nombreux modèles mathématiques ont été utilisés, en particulier les tenseurs. Ces différents modèles locaux peuvent être interprétés comme rendant compte de faisceaux de fibres de myéline ou de croisements de faisceaux de fibres.

Les descripteurs semi-globaux, décrivant des faisceaux de fibres dans leur globalité, permettent de rendre compte de la connexion entre différentes régions corticales du cerveau.

Les descripteurs globaux, connus sous le nom de connectogrammes, permettent de rendre compte de phénomènes à l'échelle du cerveau. Ils constituent une intégration de toute l'information fournie par les descripteurs semi-globaux. Ces derniers descripteurs peuvent renseigner sur l'intégrité d'une région corticale, ainsi que permettre de décrire des réseaux neuronaux complexes. On peut remarquer que la description anatomique du cerveau d'un point de vue global permet de mieux comprendre certains phénomènes fonctionnels.

Dans la première partie de cette thèse, et après un bref descriptif de la physique de la diffusion et de son utilisation en IRMd, nous nous intéresserons d'abord aux outils mathématiques, et plus précisément aux tenseurs et aux harmoniques sphériques, permettant de modéliser les phénomènes locaux. Dans un deuxième temps, nous décrirons ces modèles locaux et leurs différentes interprétations. Ensuite nous étudierons les descripteurs semi-globaux, ainsi que les différents algorithmes pour les obtenir. Finalement, dans une dernière partie, nous présenterons les descripteurs globaux.

L'objectif de cette première partie n'est pas de décrire de façon détaillée modèles et méthodes mis en œuvre en IRMd, mais de mettre en évidence les grandes tendances en IRMd.

Chapitre 1

Physique de la diffusion et lien avec le signal IRM de diffusion

1.	Mouvement brownien et propagateur	11
2.	Diffusions libre et restreinte	12
3.	Lien entre le signal IRMd et la diffusion	12

L'imagerie par résonance magnétique de diffusion (IRMd) est basée sur le mouvement aléatoire des molécules d'eau. Cette notion de déplacement est particulièrement importante en imagerie cérébrale car les molécules d'eau se meuvent de manière plus ou moins contrainte selon l'architecture neuronale sous-jacente.

Dans cette section, nous expliquerons de manière succincte le phénomène de diffusion des molécules d'eau, ainsi que les modèles physiques correspondants.

1. Mouvement brownien et propagateur

La diffusion microscopique correspond au processus par lequel toute molécule d'eau dans un fluide, soumise à une énergie thermique, est transportée de manière irréversible. Ce déplacement est connu sous le nom de mouvement brownien. D'un point de vue microscopique, décrire les déplacements d'un système de particules est impossible en raison du grand nombre de particules mis en jeu. Néanmoins, à l'échelle macroscopique, il est possible de décrire ces déplacements en termes de fonctions de densité de probabilité (*probability density function* (PDF)). Plus précisément, il est possible de connaître la densité de probabilité qu'une particule partant de l'endroit \mathbf{r} arrive à l'endroit \mathbf{r}' en un intervalle de temps t donné. La fonction se nomme propagateur et peut être écrite sous la forme $p(\mathbf{r}, \mathbf{r}', t)$ [CRANK 1980].

Sous ce formalisme, le phénomène de diffusion peut être décrit par l'équation de Fokker-Planck, correspondant aussi à la seconde loi de Fick [FICK 1855, RISKEN & FRANK 1996] :

$$\frac{\partial p(\mathbf{r}, \mathbf{r}', t)}{\partial t} = \nabla \cdot (\mathbf{D} \nabla p(\mathbf{r}, \mathbf{r}', t)), \quad (1.1)$$

avec ∇ l'opérateur gradient, t le temps de diffusion et \mathbf{D} le tenseur d'ordre 2 représentant la diffusion du milieu et appelé usuellement tenseur de diffusion. Nous reviendrons par la suite en détail sur les caractéristiques de ce tenseur.

Il faut noter que le modèle du propagateur décrit par la loi de Fokker-Planck est le plus général possible. On peut simplifier cette équation sous l'hypothèse que la diffusion est constante

localement :

$$\frac{\partial p(\mathbf{r}, \mathbf{r}', t)}{\partial t} = \mathbf{D} \nabla^2 p(\mathbf{r}, \mathbf{r}', t), \quad (1.2)$$

où ∇^2 correspond à l'opérateur laplacien. A noter que cette loi peut être retrouvée par analogie avec la première loi de Fick décrivant l'évolution de concentrations dans un milieu fluide, mais peut également être déduite directement du modèle élémentaire de la marche aléatoire.

Cette équation décrit l'évolution du propagateur pour une molécule d'eau. En pratique, nous nous intéressons à un ensemble de molécules. On considère alors le formalisme du propagateur moyen de l'ensemble des molécules (PME) ou *ensemble average propagator* :

$$P(\mathbf{R}, t) = \int_{\omega} p(\mathbf{r}, \mathbf{r}', t) \rho(\mathbf{r}) d\mathbf{r}, \quad (1.3)$$

avec $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ le vecteur de déplacement, ω le domaine considéré (dans notre cas le voxel) et $\rho(\mathbf{r})$ la densité initiale en molécules d'eau. Par la suite, nous ne garderons que ce formalisme.

Nous pouvons observer trois différentes formes de distributions de la probabilité de diffusion des molécules d'eau en présence de trois configurations anatomiques sous-jacentes différentes (voir la Fig. 1.1).

2. Diffusions libre et restreinte

La diffusion libre correspond au cas simple où l'on suppose que le milieu est homogène, infini et qu'il n'existe pas d'entrave au déplacement des molécules d'eau. Sous ces hypothèses, la solution de l'équation conduit à un propagateur moyen gaussien :

$$P(\mathbf{R}, t) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp\left(-\frac{\mathbf{R}^T \mathbf{D}^{-1} \mathbf{R}}{4t}\right). \quad (1.4)$$

On parle aussi de diffusion gaussienne. Cependant, la diffusion gaussienne est un modèle simpliste pour les milieux anatomiques rencontrés en IRMd. La relation entre la diffusion de l'eau et le milieu anatomique sous-jacent reste un problème ouvert [TUCH 2002]. Dans le cadre de la diffusion de molécules d'eau dans le cerveau, de nombreux modèles ont été proposés pour rendre compte de la structure complexe sous-jacente. En présence de ces structures, la diffusion n'est plus considérée comme libre, mais est considérée restreinte. Le lecteur intéressé par la diffusion restreinte ainsi que sur le lien existant entre structure anatomique et anisotropie du milieu pourra se référer à l'article [BEAULIEU 2002].

Bien que l'hypothèse de diffusion gaussienne soit simpliste, elle a été largement utilisée au cours de la dernière décennie, étant à la base de la modalité la plus utilisée : l'imagerie du tenseur de diffusion (*diffusion tensor imaging* (DTI)) [BASSER et coll. 1994].

3. Lien entre le signal IRMd et la diffusion

En pratique, le déplacement des molécules d'eau n'est pas observable directement. L'IRM conventionnelle permet de déterminer la densité des molécules d'eau par la mesure du spin qui, au même titre que la masse ou la charge de l'électron, est une caractéristique intrinsèque de la particule d'hydrogène H , composante de la molécule d'eau H_2O .

Apparue dans le milieu des années 80, l'IRM de diffusion est une extension de l'IRM conventionnelle [LE BIHAN & BRETON 1985]. De manière générale, il est ajouté à une séquence

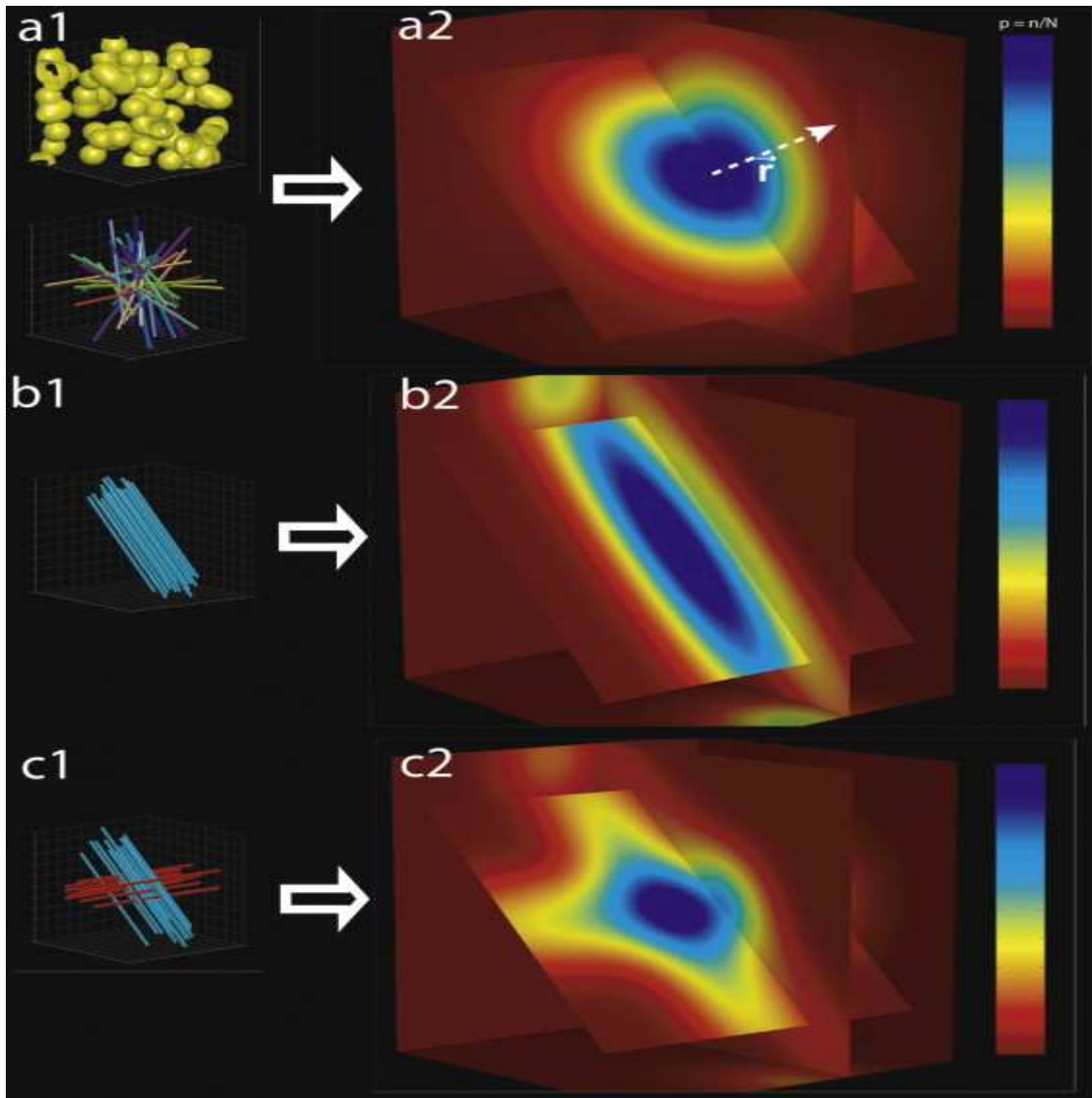


FIGURE 1.1 – A gauche : quatre configurations anatomiques correspondant respectivement à un milieu isotrope composé de cellules sphériques, un milieu composé d’une multitude de structures tubulaires orientées de manière aléatoire, un milieu structuré avec une direction privilégiée (*i.e.* un paquet de fibres de substance blanche), et un milieu structuré avec deux directions privilégiées (*i.e.* un croisement de paquets de fibres). A droite : la probabilité de diffusion des molécules d’eau correspondante. On observe que la diffusion isotrope (a2) peut être obtenue pour deux configurations anatomiques différentes (a1). Image extraite de [HAGMANN et coll. 2010].

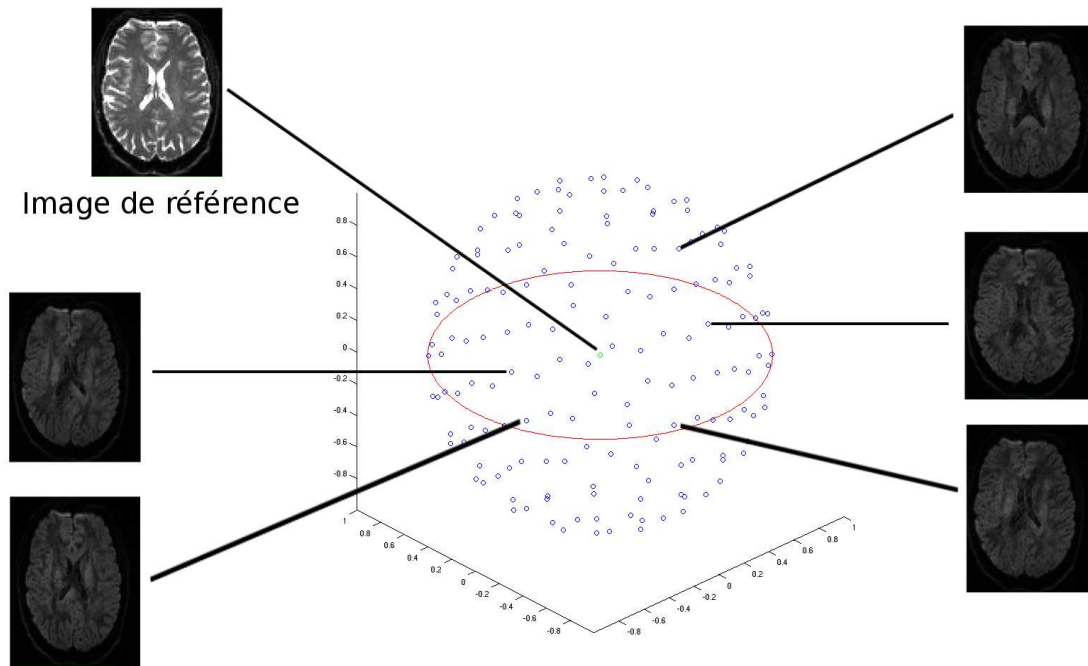


FIGURE 1.2 – Principe de l'imagerie de type HARDI. Les gradients de diffusion \mathbf{g} échantillonnent la sphère unité. Chaque point sur la sphère correspond à une image pondérée en diffusion. Le centre de cette sphère correspond à l'image de référence sans pondération ($b = 0$).

T_2 d'IRM conventionnelle un gradient de champ magnétique dans une direction spatiale \mathbf{g} . Ce gradient est communément appelé gradient de diffusion. L'image ainsi obtenue, appelée image pondérée en diffusion, correspond aux mouvements des molécules d'eau dans la direction \mathbf{g} . L'autre paramètre important de la séquence est la b -valeur, notée b , valeur proportionnelle au temps d'acquisition. Ce coefficient est crucial car la molécule d'eau parcourra plus ou moins de chemin selon le temps de diffusion. On observe ainsi différents phénomènes de diffusion selon cette valeur de b [ASSEMBLAL 2010, ASSEMBLAL et coll. 2011]. Plus précisément, dans une direction donnée, la décroissance du signal en fonction de b n'est pas gaussienne. On assiste à des alternances de phénomènes de diffusion et de diffraction. L'estimation de ce profil complexe nécessite l'emploi de plusieurs valeurs de b .

Le lecteur intéressé par une explication plus complète de l'acquisition en IRM de diffusion pourra se référer à la thèse de Cyril Poupon [POUPON 1999] ou à l'article de Ludovico Minati et Władysław Węglarz [MINATI & WĘGLARZ 2007].

Le signal $S(\mathbf{q}, t)$ délivré par l'IRM correspond à l'atténuation du signal de résonance magnétique (RM). De manière plus précise, il existe une relation entre l'atténuation du signal $S(\mathbf{q}, t)$ de RM et le propagateur moyen $P(\mathbf{R}, t)$ [STEJSKAL & TANNER 1965] :

$$S(\mathbf{q}, t) = S_0 \int_{\text{voxel}} P(\mathbf{R}, t) \exp(i2\pi\mathbf{q}^T \mathbf{R}) d\mathbf{R}, \quad (1.5)$$

avec S_0 le signal de référence acquis sans pondération (correspondant à $b = 0$), t le temps de diffusion moléculaire, $\mathbf{q} = \gamma\delta\mathbf{g}/(2\pi)$ le vecteur de déplacement réciproque (avec γ le ratio gyromagnétique des protons de l'eau et δ la durée d'application du gradient de diffusion).

On observe que $S(\mathbf{q}, t)$ correspond à la transformée de Fourier inverse du PME. L'espace de l'ensemble des vecteurs \mathbf{q} s'appelle le q-espace. Cet espace est à la base de l'*imagerie q-espace* (*q-space imaging*) [CALLAGHAN et coll. 1990, TUCH et coll. 2002]. Le PME contient toute l'information de diffusion contenue dans un voxel. Cependant, pour inverser l'équation 1.5 et estimer de manière fiable le PME, il est nécessaire de connaître $S(\mathbf{q}, t)$ en de nombreux vecteurs \mathbf{q} . Or les données $S(\mathbf{q}, t)$ acquises remplissent le q-espace de manière parcimonieuse. Le choix du modèle de diffusion va donc fortement dépendre du nombre d'acquisitions et de leurs positions relatives dans le q-espace. Pour compenser le manque d'acquisitions, on peut émettre des hypothèses fortes sur le propagateur. Sous l'hypothèse où la diffusion est gaussienne (voir Eq. 1.4), l'équation précédente (voir Eq. 1.5) peut s'écrire sous la forme :

$$S(\mathbf{g}, b) = S_0 \exp(-b \mathbf{g}^T \mathbf{D} \mathbf{g}), \quad (1.6)$$

avec \mathbf{g} le gradient diffusion normalisé correspondant à $\frac{\mathbf{q}}{\|\mathbf{q}\|}$, b la b -valeur égale à $t\|\mathbf{q}\|^2$, et \mathbf{D} le tenseur de diffusion. Cette dernière équation, plus connue sous le nom d'équation de Stejskal-Tanner, est à la base de l'imagerie du tenseur de diffusion [BASSER et coll. 1994]. Des acquisitions échantillonnant le q-espace de manière fine permettent de ne plus faire d'hypothèse sur le propagateur.

Les deux grands types d'acquisitions sont :

- un échantillonnage cartésien du q-espace permettant de calculer directement le PME par transformée de Fourier du signal. Ce type d'imagerie est appelé l'imagerie de diffusion spectrale [WEDEEN et coll. 2005] (*diffusion spectral imaging* (DSI)),
- un échantillonnage du q-espace sur une sphère de rayon la b -valeur b (voir Fig. 1.2), connu sous le nom d'imagerie de diffusion à haute résolution angulaire (*high angular resolution diffusion imaging* (HARDI)). Ce dernier type d'échantillonnage est à l'origine de nombreux modèles en IRMd.

Il faut noter que les échantillonnages sphériques sont plus largement utilisés. Cela provient du fait que l'acquisition cartésienne dure beaucoup plus longtemps que l'acquisition sphérique.

Chapitre 2

Outils mathématiques

1.	Les tenseurs	17
1.1	Formalisme	18
1.2	Tenseurs en IRM de diffusion	18
1.3	Le tenseur d'ordre 2	18
1.4	Le tenseur d'ordre 4	22
2.	Les harmoniques sphériques	28
2.1	Représentation	28
2.2	Caractéristiques	28
2.3	Métriques	29
3.	Lien entre tenseurs et harmoniques sphériques	30
4.	Conclusion partielle	30

Dans cette partie, nous décrivons les outils mathématiques nécessaires à la modélisation du signal d'IRM de diffusion. Cette description n'est pas exhaustive, et portera sur les deux outils les plus utilisés : les *tenseurs* et les *harmoniques sphériques*. Nous nous attacherons à montrer les spécificités de chaque outil ainsi que le lien unissant ces deux outils. Cette étude permettra de montrer au lecteur l'importance de choisir l'outil le plus adapté au traitement désiré.

Nous supposons que les fonctions sphériques considérées par la suite sont symétriques, définies, positives. Ces contraintes résultent des propriétés physiques de la diffusion.

1. Les tenseurs

Un tenseur est un objet mathématique permettant de représenter des applications multilinéaires. Il peut aussi être vu comme une généralisation des notions de scalaire et de vecteur. Le tenseur est un outil utilisé en mathématique en géométrie différentielle, ainsi qu'en physique notamment en mécanique et en théorie de la relativité. Il est largement utilisé pour une de ses propriétés : il ne dépend pas d'un référentiel donné, contrairement à ses coefficients. Cependant il existe des relations mathématiques assurant l'invariabilité du tenseur lors d'un changement de référentiel.

En raison de la complexité de ce modèle, nous nous restreindrons à une description utile en pratique pour l'IRM de diffusion. Les lecteurs désirant avoir plus d'informations sur les tenseurs peuvent se reporter au cours de Jean Garrigues [GARRIGUES 2007].

1.1 Formalisme

Le formalisme présenté ci-dessous est extrait du cours de Jean Garrigues. Soit \mathbf{E} un espace euclidien de dimension N . Un tenseur \mathbf{T} d'ordre p est une application p -linéaire de $\mathbf{E} \times \mathbf{E} \times \mathbf{E} \times \dots \times \mathbf{E}$ (p fois) dans \mathbb{R} .

Soit $\{\mathbf{e}_i\}$ une base de \mathbf{E} , et \mathbf{T} un tenseur d'ordre p . On a la relation suivante :

$$\mathbf{T}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p) = \sum_{i_1, i_2, \dots, i_p=1}^N T_{i_1 i_2 \dots i_p} X_{i_1}^1 X_{i_2}^2 \dots X_{i_p}^p, \quad (2.1)$$

où X_i^j est la i ème coordonnée du vecteur \mathbf{X}_j . Les différents $T_{i_1 i_2 \dots i_p}$ sont appelés les coefficients du tenseur \mathbf{T} dans la base $\{\mathbf{e}_i\}$. On appelle \mathbf{T} tenseur d'ordre p de dimension N .

1.2 Tenseurs en IRM de diffusion

Pour satisfaire aux contraintes de symétrie de la diffusion, le tenseur en IRM de diffusion doit être un tenseur d'ordre pair. Le tenseur le plus usité en IRM de diffusion est le tenseur d'ordre 2 [BASSER et coll. 1994], car c'est le modèle le plus simple et il permet d'obtenir des résultats satisfaisants. Les tenseurs d'ordres supérieurs (*high order tensors* (HOT)) voient peu à peu le jour dans les modèles [BASSER & PAJEVIC 2007, GHOSH et coll. 2008, BARMPOUTIS et coll. 2009, QI et coll. 2009].

Nous nous limiterons dans cette partie à l'étude des tenseurs d'ordre 2 et d'ordre 4. Ces deux types de tenseurs permettent de modéliser un bon nombre de phénomènes observés en IRM de diffusion.

Nous utiliserons la notation $\mathbf{g} = [g_1, g_2, g_3]^T$ pour décrire un vecteur dans \mathbb{R}^3 .

1.3 Le tenseur d'ordre 2

D'un point de vue analytique, le tenseur d'ordre 2 peut être considéré comme un polynôme homogène, c'est-à-dire comme un polynôme dont tous les monômes sont de même degré. Le tenseur d'ordre 2 correspond à l'expression ci-dessous :

$$d(\mathbf{g}) = \sum_{i,j=1}^3 D_{ij} g_i g_j. \quad (2.2)$$

Le tenseur d'ordre 2 en diffusion n'opère que sur un vecteur \mathbf{g} au lieu de deux. Formellement, on devrait écrire $d(\mathbf{g}, \mathbf{g})$ pour être en accord avec l'équation 2.1. Cependant, pour alléger la notation, nous noterons $d(\mathbf{g})$.

On peut remarquer qu'en principe, le tenseur d'ordre 2 est composé de neuf coefficients. Cependant il n'existe que six degrés de liberté. En effet, les couples de coefficients (D_{ij}, D_{ji}) avec i et j différents n'ont d'effet que par $(D_{ij} + D_{ji})$ dans l'équation 2.2. Par conséquent, même si le tenseur \mathbf{D} d'ordre 2 n'est pas symétrique (au sens où les coefficients D_{ij} et D_{ji} ne seraient pas égaux pour i et j différents), tout se passe comme s'il l'était effectivement et on peut considérer le tenseur \mathbf{D}^{sym} symétrique en posant

$$D_{ij}^{sym} = \frac{D_{ij} + D_{ji}}{2}.$$

Comme spécifié précédemment, nous nous limiterons à l'étude des tenseurs d'ordre 2 symétriques définis positifs. L'espace ainsi considéré est assimilable à l'espace des matrices symétriques

définies positives, noté généralement $GL_3^{+*}(\mathbb{R})$ pour groupe général linéaire. La matrice représentant le tenseur d'ordre 2 de diffusion s'écrit :

$$\mathbf{D} = \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{pmatrix}. \quad (2.3)$$

D'un point de vue algébrique, une matrice appartenant à cet espace est diagonalisable car, selon le théorème spectral, toute matrice symétrique réelle est diagonalisable avec λ_i ses valeurs propres et \mathbf{v}_i ses vecteurs propres. De plus, toutes ses valeurs propres sont réelles et strictement positives.

D'un point de vue géométrique, le tenseur d'ordre 2 peut être représenté par un ellipsoïde d'équation :

$$D_{11} g_1^2 + D_{22} g_2^2 + D_{33} g_3^2 + 2 D_{12} g_1 g_2 + 2 D_{13} g_1 g_3 + 2 D_{23} g_2 g_3 = C, \quad (2.4)$$

avec C une constante. Les demi-grands axes de cet ellipsoïde correspondent aux vecteurs propres \mathbf{v}_i . Plus rarement, on retrouve une représentation en forme de « cacahuète » (*peanut shape*). On l'obtient en fixant la norme du vecteur \mathbf{g} à 1, et en représentant $d(\mathbf{g})$.

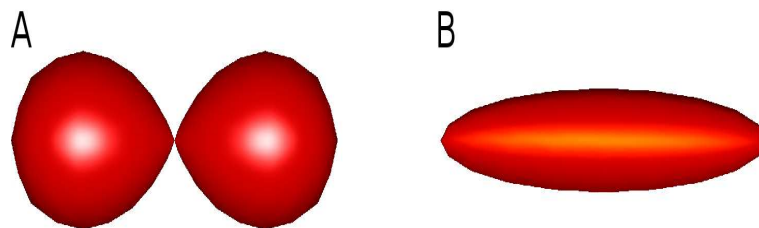


FIGURE 2.1 – A : surface paramétrée par un tenseur d'ordre 2 avec une représentation en forme de « cacahuète ». B : le même tenseur d'ordre 2 représenté par un ellipsoïde.

Métriques pour le tenseur d'ordre 2

La notion de métrique est importante en IRM de diffusion pour de nombreuses applications telles que l'estimation des tenseurs ou les statistiques. En effet, le choix de la métrique détermine la topologie de l'espace associé. Dans cette partie, nous présenterons les différentes métriques associées à l'espace des tenseurs de diffusion d'ordre 2. Nous discuterons dans un deuxième temps du choix de cette métrique.

On utilise classiquement la distance euclidienne pour comparer deux tenseurs d'ordre 2. Pour cela, on considère le produit scalaire usuel dans l'espace des matrices :

$$\langle \mathbf{D}_1 | \mathbf{D}_2 \rangle = \text{tr}(\mathbf{D}_1^T \mathbf{D}_2) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} D_{1;ij} D_{2;ij} \in \mathbb{R}, \quad (2.5)$$

où \mathbf{D}_1 et \mathbf{D}_2 sont des matrices, et $\text{tr}(\cdot)$ correspond à l'opérateur trace. La norme associée à ce produit scalaire est la norme de Frobenius :

$$\|\mathbf{D}\|_F = \left(\sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} D_{ij}^2 \right)^{\frac{1}{2}}. \quad (2.6)$$

C'est la norme canonique associée à l'espace des matrices. Elle est invariante par rotation. On peut en déduire la distance euclidienne telle que :

$$\begin{aligned} d_{euc}(\mathbf{D}_1, \mathbf{D}_2) &= \|\mathbf{D}_1 - \mathbf{D}_2\|_F, \\ &= \sqrt{\text{trace}\left((\mathbf{D}_1 - \mathbf{D}_2)^T (\mathbf{D}_1 - \mathbf{D}_2)\right)}. \end{aligned} \quad (2.7)$$

On utilise cette distance en IRM de diffusion en émettant l'hypothèse que l'espace $GL_3^{+*}(\mathbb{R})$ est un sous-espace de l'espace des matrices symétriques.

De nombreux travaux (par exemple [ARSIGNY et coll. 2006, LENGLET et coll. 2006]) ont montré que cette distance n'était peut-être pas la plus pertinente en IRM de diffusion. On peut remarquer que l'espace $GL_3^{+*}(\mathbb{R})$ a de moins bonnes propriétés algébriques que l'espace des matrices symétriques. Par exemple, $GL_3^{+*}(\mathbb{R})$ n'est pas un groupe pour l'addition (en effet, il ne possède pas d'élément neutre ni d'opposé). Un autre inconvénient existe : les matrices symétriques négatives peuvent être à des distances finies de matrices appartenant à $GL_3^{+*}(\mathbb{R})$.

Prenons un exemple concret : la matrice identité n'est qu'à une distance 3 de la matrice nulle qui n'a pas de réalité physique en IRMd. Pour résoudre ce problème, plusieurs métriques ont été proposées. Deux des métriques les plus connues sont la métrique riemannienne et la métrique log-euclidienne.

La métrique riemannienne a été proposée pour l'IRM de diffusion [LENGLET et coll. 2006, FLETCHER et coll. 2007]. Le principe est de considérer l'espace $GL_3^{+*}(\mathbb{R})$ comme une variété riemannienne, c'est-à-dire que l'espace n'est plus considéré comme euclidien mais comme un espace courbé [FLETCHER et coll. 2007]. La distance entre deux matrices correspond alors à la longueur du chemin le plus court en prenant en compte la topologie de l'espace. Ces chemins s'appellent des géodésiques.

La distance entre deux matrices est obtenue alors par la formule suivante :

$$d_{rie}(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{\frac{1}{2} \text{trace}\left(\log\left(\mathbf{D}_1^{-1/2} \mathbf{D}_2 \mathbf{D}_1^{-1/2}\right)^2\right)}. \quad (2.8)$$

L'inconvénient de cette métrique est sa complexité algorithmique. Par exemple, calculer la moyenne d'un ensemble de matrices appartenant à $GL_3^{+*}(\mathbb{R})$ ne correspond plus au calcul de la moyenne dans le cadre euclidien. Par exemple, dans le cadre riemannien, il faut s'assurer que la moyenne appartient à la variété, ce qui n'est pas nécessaire dans le formalisme euclidien. On considère alors la moyenne au sens de Fréchet, et son calcul nécessite une optimisation [LENGLET et coll. 2006] :

$$\mathbf{D}_{moy_{Fréchet}} = \arg \min_{\mathbf{D} \in GL_3^{+*}(\mathbb{R})} \sum_{i=1}^N d^2(\mathbf{D}, \mathbf{D}_i). \quad (2.9)$$

La deuxième métrique, appelée log-euclidienne, propose une alternative¹. L'idée est de transformer l'espace $GL_3^{+*}(\mathbb{R})$ en l'espace des matrices symétriques, qui a de bonnes propriétés, et d'effectuer les calculs dans ce dernier en le considérant comme un espace euclidien. La transformation utilisée est le logarithme, d'où le nom de cette métrique. Arsigny et coll. montrent qu'appliquer le logarithme à une matrice revient à appliquer le logarithme à ses valeurs propres [ARSIGNY et coll. 2006]. Les calculs reviennent à effectuer une décomposition spectrale et à

1. La métrique log-euclidienne est également riemannienne. On utilise les dénominations riemannienne (éq. 2.8) et log-euclidienne (éq. 2.10) pour les différencier.

considérer le logarithme des valeurs propres :

$$\begin{aligned} d_{\log\text{-euclidien}}(\mathbf{D}_1, \mathbf{D}_2) &= \|\log(\mathbf{D}_1) - \log(\mathbf{D}_2)\| \\ &= \|\mathbf{R}_1^T \log(\mathbf{\Lambda}_1) \mathbf{R}_1 - \mathbf{R}_2^T \log(\mathbf{\Lambda}_2) \mathbf{R}_2\|, \end{aligned} \quad (2.10)$$

avec $\|\cdot\|$ la norme euclidienne, et $\mathbf{R}_i^T \mathbf{\Lambda}_i \mathbf{R}_i$ la décomposition spectrale de \mathbf{D}_i .

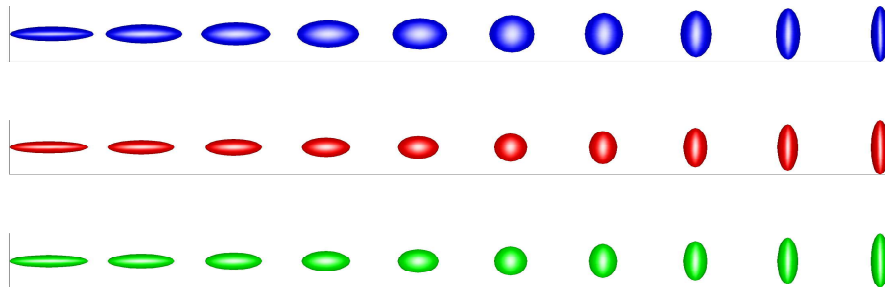


FIGURE 2.2 – Schéma montrant le phénomène de gonflement sur les tenseurs. On effectue une interpolation entre le tenseur le plus à gauche et celui le plus à droite avec différentes métriques : celle euclidienne (en haut), celle riemannienne (au milieu), et celle log-euclidienne (en bas). On observe un phénomène de gonflement lors de l’interpolation euclidienne, qui n’apparaît pas avec les deux autres interpolations.

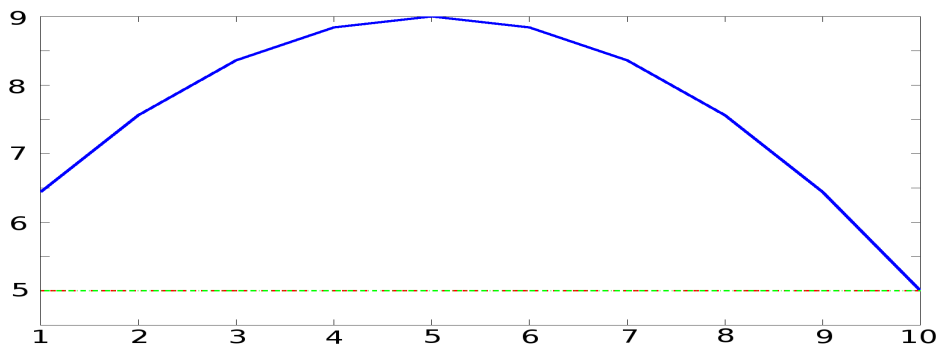


FIGURE 2.3 – Schéma montrant le phénomène de gonflement sur le volume des tenseurs. On visualise l’évolution du volume (en ordonnée) le long de l’interpolation (en abscisse). On estime le volume des tenseurs grâce au déterminant. On constate que le volume lors de l’interpolation euclidienne (en bleu) augmente, alors que le volume pour celle riemannienne (en rouge) et celle log-euclidienne (en vert) restent constantes.

Par exemple, la moyenne arithmétique dans $GL_3^{+*}(\mathbb{R})$ correspond à une moyenne géométrique en log-euclidien. Si on reprend l’exemple de la distance entre la matrice nulle et la matrice identité, la distance log-euclidienne est infinie entre ces deux matrices, alors qu’elle est de 3 pour une distance euclidienne. La distance log-euclidienne correspond donc mieux aux contraintes de positivité imposées en IRMd.

L’une des propriétés de ces deux distances est l’invariance à des effets d’échelle (*i.e.* la distance entre deux tenseurs ne varie pas s’ils sont tous les deux multipliés par le même facteur

d'échelle α). Plus précisément, la métrique riemannienne est affine invariante, et la métrique log-euclidienne est invariante par similitude. Le deuxième avantage, autre que celui de placer les tenseurs non définis positifs à des distances infinies, est de limiter le phénomène de gonflement (*swelling effect*). Ce phénomène est visualisé en figures 2.2 et 2.3. Ce phénomène se produit lorsqu'on essaie d'interpoler ou de moyenniser des tenseurs anisotropes possédant des axes principaux non colinéaires [ARSIGNY et coll. 2006, FLETCHER et coll. 2007]. La conséquence d'un tel phénomène est l'obtention de tenseurs avec des volumes plus importants que celui de chaque tenseur intervenant dans l'interpolation (voir par exemple [BATCHELOR et coll. 2005]).

Il existe différentes métriques autres que les métriques euclidienne, riemannienne et log-euclidienne. Nous renvoyons le lecteur aux thèses de Christophe Lenglet [LENGLET 2006], de Vincent Arsigny [ARSIGNY 2006] et de Pierre Fillard [FILLARD 2008], ainsi qu'à l'article [DRYDEN et coll. 2009] pour une étude plus approfondie des propriétés de ces métriques.

Devant ce grand nombre de métriques proposées, on peut s'interroger sur le choix de la métrique. Plus précisément, il n'existe pas de métrique idéale pour l'ensemble des traitements d'images (par exemple l'interpolation ou les tests statistiques). La métrique doit être sélectionnée en fonction du traitement à effectuer. Dans [PASTERNAK et coll. 2010]), les auteurs s'interrogent sur les variétés dans lesquelles résident les tenseurs de diffusion d'ordre 2, et soutiennent l'idée que le choix de la métrique ne doit pas être fait uniquement sur des considérations mathématiques, mais que des considérations physiques doivent être prises en compte. Ils introduisent la notion de mesure qualitative (associée aux métriques riemanniennes) et de mesure quantitative (associée à la métrique euclidienne). Dans [TARANTOLA 2005], Albert Tarantola insiste sur le fait que la plupart des mesures physiques sont positives et qualitatives. La distance associée à de telles grandeurs est le logarithme de leur rapport². En IRMd, certaines grandeurs sont également positives. On peut donc supposer que l'utilisation de distances du type de celle évoquée par Albert Tarantola est pertinente et justifiée.

Dans [PASTERNAK et coll. 2010], les auteurs soutiennent que la métrique euclidienne, bien que n'assurant pas la positivité, a de meilleures propriétés lors de l'estimation des tenseurs intra-voxel. Cependant, les auteurs reconnaissent aussi l'utilité des métriques affines invariantes lors de l'interpolation de tenseurs ayant la même anisotropie (par exemple des tenseurs appartenant à un même faisceau de fibres). Les auteurs introduisent une notion importante dans le choix de la métrique : la métrique doit être dépendante du contexte dans lequel on l'emploie.

Limitations du tenseur d'ordre 2

Le tenseur d'ordre 2 ne représente qu'une forme quadratique. De ce fait, ce modèle ne contient que peu d'informations. La limitation la plus importante est qu'il ne possède que trois directions privilégiées (*i.e.* les vecteurs propres), qui sont elles-mêmes contraintes à être orthogonales entre elles.

Dans les années 2000, la communauté de l'IRM de diffusion a commencé à introduire un outil généralisant le tenseur d'ordre 2, le tenseur d'ordre N avec N pair, et plus particulièrement le tenseur d'ordre 4.

1.4 Le tenseur d'ordre 4

Généralement en IRMd, un tenseur d'ordre 4 est une application multilinéaire de \mathbb{R}^3 dans \mathbb{R} . Il se peut que le tenseur d'ordre 4 soit considéré comme une application de \mathbb{R}^6 dans \mathbb{R}

2. « Most of the physical scalars are positive (mass, period, temperature...), and, typically, the distance is not related to the difference of values but, rather, to the logarithm of the ratio of values. » [TARANTOLA 2005, page 107].

[BASSER & PAJEVIC 2007, MOAKHER 2008, MOAKHER 2009]. Bien que les représentations matricielles soient identiques, les propriétés des tenseurs d'ordre 4 dans ces deux cas sont très différentes. Nous supposons tout au long de cette section que nous sommes dans le premier cas.

Symétries et notations mathématiques

A la manière de l'équation 2.2, le tenseur d'ordre 4 correspond à l'équation ci-dessous :

$$d(\mathbf{g}) = \sum_{i,j,k,l=1}^3 D_{ijkl} g_i g_j g_k g_l, \quad (2.11)$$

où D_{ijkl} correspond aux composantes du tenseur et g_i correspond à la i ème composante du vecteur \mathbf{g} . Dans le cadre de l'IRM de diffusion, nous nous limiterons à des vecteurs \mathbf{g} de norme unitaire. Cela revient à considérer que l'espace de départ est la sphère \mathbb{S}^2 . Formellement, nous devrions écrire $d(\mathbf{g}, \mathbf{g}, \mathbf{g}, \mathbf{g})$. Cependant, comme pour le tenseur d'ordre 2 (voir section 1.3), nous notons $d(\mathbf{g})$ pour alléger la notation.

De manière générale, un tenseur d'ordre 4 est composé de 81 coefficients ($81 = 3^4$). Si l'on peut permuter les deux premiers et deux derniers indices séparément ($D_{ijkl} = D_{jikl} = D_{ijlk} = D_{jilk}$) sans modifier le résultat de l'équation 2.1, le tenseur possède une symétrie dite « faible » ou « petite ». Si l'on peut permuter les deux premiers et deux derniers indices entre eux ($D_{ijkl} = D_{klij}$) sans changer le résultat de l'équation 2.1, le tenseur possède une symétrie dite « forte » ou « grande ». Le fait de posséder les deux symétries est parfois dénommé « super-symétrie ». Un tenseur d'ordre 4 doté de cette dernière propriété est paramétré par 15 coefficients.

On considérera que le tenseur d'ordre 4 de diffusion est un tenseur « super-symétrique ». Plus précisément, l'équation 2.11 peut se factoriser en 15 monômes. On remarque alors que les différents coefficients du tenseur s'expriment uniquement en groupe. Par exemple, considérons le coefficient D_{1113} qui ne s'exprime dans l'équation 2.11 que par le monôme ($D_{1113} + D_{1131} + D_{1311} + D_{3111}$) $\mathbf{g}_1^3 \mathbf{g}_3$. En pratique, même si le tenseur \mathbf{D} d'ordre 4 de l'équation 2.11 n'est pas « super-symétrique », tout se passe comme s'il l'était effectivement et on peut considérer le tenseur \mathbf{D}^{sym} « super-symétrique » :

$$D_{ijkl}^{sym} = \frac{\sum_{\sigma \in \mathfrak{S}} D_{\sigma(ijkl)}}{\mu},$$

où σ est une permutation de l'ensemble \mathfrak{S} des permutations possibles (on ne permute que les éléments discernables), et μ est la multiplicité du coefficient D_{ijkl}^{sym} (*i.e.* le cardinal de l'ensemble \mathfrak{S}). Pour l'exemple D_{1113} , la multiplicité est égale à quatre.

En considérant uniquement la « petite » symétrie, on est en présence de 36 coefficients indépendants. En utilisant la notation dite de Voigt, largement utilisée en mécanique, on a alors une matrice de taille 6×6 :

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

Il faut prendre en compte des coefficients de pondération dans la matrice \mathbf{D} pour que $d(\mathbf{g}) = \mathbf{g}^T \mathbf{D} \mathbf{g}$ (voir éq. 2.11) avec $\mathbf{g} = [g_1^2, g_2^2, g_3^2, g_1 g_2, g_2 g_3, g_1 g_3]^T$. Ces coefficients n'ont pas été inclus afin de simplifier les notations. On peut constater que cette notation, très utilisée en mécanique dans la représentation du tenseur élastique, commence à être employée en IRM de diffusion pour ses propriétés matricielles [BASSER & PAJEVIC 2007, GHOSH et coll. 2008, MOAKHER 2008, MOAKHER 2009]. Cependant ce formalisme est complètement différent entre la mécanique et l'IRM de diffusion. En effet, le vecteur d'application utilisé en mécanique appartient à \mathbb{R}^6 et ses coefficients ne sont pas liés, alors que par construction, le vecteur d'application $[g_1^2, g_2^2, g_3^2, g_1 g_2, g_2 g_3, g_1 g_3]$ utilisé en IRM de diffusion a les trois dernières composantes liés aux trois premières. Cette différence fondamentale fait qu'on ne peut pas utiliser les outils liés à la mécanique de manière aussi simple et directe en IRM de diffusion. Cette notation est introduite dans l'article [MOAKHER 2008]. Barmpoutis et coll. soulignent dans leur article [BARMPOUTIS et coll. 2009] les problèmes liés à cette notation. Cette discussion sera abordée plus largement dans la partie métrique.

En considérant la super-symétrie des tenseurs, on n'a plus que 15 coefficients indépendants, qui sont généralement regroupés sous forme vectorielle. Pour plus de clarté, Barmpoutis et coll. [BARMPOUTIS et coll. 2009] introduisent une notation plus adaptée à cette super-symétrie : D_{ijk} où i, j et k correspondent aux puissances de g_1, g_2 et g_3 respectivement. A chaque coefficient est associé un ordre de multiplicité. Par exemple, D_{400} correspondant à D_{1111} est de multiplicité 1 alors que D_{310} est de multiplicité 4 (car $D_{1112} = D_{1121} = D_{1211} = D_{2111}$). Cette multiplicité peut être incluse dans la notation de Barmpoutis, mais doit être prise en compte si on repasse sous forme matricielle.

Intéressons-nous maintenant au lien existant entre les représentations algébrique, analytique et géométrique du tenseur d'ordre 4 (T4)³.

Représentations algébrique, analytique et géométrique

D'un point de vue algébrique, un T4 super-symétrique peut être considéré comme une matrice symétrique grâce aux notations de Voigt vues précédemment. L'avantage est d'obtenir une représentation paramétrique simple d'un objet complexe. Le principal inconvénient est le manque de sens physique pour les différents paramètres. Par exemple, on comprend l'influence du terme D_{400} dans l'équation (ce paramètre agit uniquement dans la direction $(1, 0, 0)$), mais il est plus difficile de voir l'impact des termes couplés tels que D_{211} sur la forme géométrique.

D'un point de vue analytique, un tenseur d'ordre 4 correspond à une équation dite ternaire quartique, *i.e.* un polynôme homogène de degré 4 en trois variables. Un polynôme homogène de degré 4 est à un polynôme dont la somme des degrés de chaque monôme est égale à 4. Ces polynômes admettent des décompositions (décompositions dites de Hilbert) vues dans la prochaine section.

D'un point de vue géométrique, on peut représenter le T4 par une surface comme pour le T2, si on considère que l'espace d'application du T4 est une sphère (cas usuel en IRM de diffusion). L'étude de ces surfaces a été effectuée par Qi et coll. dans [QI 2006].

Le T4 avec restriction sur la sphère permet d'avoir au plus trois maxima locaux non colinéaires. Ce modèle permet donc plus d'oscillations que le T2, et est donc mieux adapté que ce dernier pour la modélisation de phénomènes complexes tel que le croisement de fibres.

3. On notera T2 le tenseur d'ordre 2.

Décompositions et valeurs remarquables

Le T4, et plus généralement les tenseurs d'ordre supérieur, sont des outils mathématiques complexes et de grandes dimensions. Les différentes communautés utilisant le T4 ont proposé des techniques pour synthétiser l'information. On peut citer comme exemple les travaux de Lieven De Lathauwer et coll. [DE LATHAUWER et coll. 2000] en traitement du signal, ou les travaux de Sandra Forte et Maurizio Vianello [FORTE & VIANELLO 1996] en mécanique.

Nous nous intéresserons ici aux décompositions puis aux valeurs remarquables utilisées en IRM de diffusion.

Dans leurs travaux sur l'estimation du T4 [BARMPOUTIS et coll. 2009], Barmpoutis et coll. ont introduit la décomposition basée sur un théorème de Hilbert, et définie dans la suite de ce mémoire comme décomposition de Hilbert. Le théorème est le suivant :

Théorème 1. *Une quartique ternaire d réelle positive peut être décomposée en une somme de trois fonctions quadratiques au carré.*

Ce théorème peut s'écrire de la manière suivante :

$$d(\mathbf{g}) = (\mathbf{v}^T \mathbf{q}_1)^2 + (\mathbf{v}^T \mathbf{q}_2)^2 + (\mathbf{v}^T \mathbf{q}_3)^2 = \mathbf{v}^T \mathbf{Q} \mathbf{Q}^T \mathbf{v} = \mathbf{v}^T \mathbf{G} \mathbf{v}, \quad (2.12)$$

avec \mathbf{v} un vecteur de monômes $[g_1^2, g_2^2, g_3^2, g_1 g_2, g_1 g_3, g_2 g_3]$, \mathbf{q}_i le vecteur contenant les coefficients de la i ème fonction quadratique. \mathbf{Q} correspond à la matrice 6×3 contenant les coefficients du polynôme. \mathbf{G} est une matrice de Gram.

Maintenant, considérons une matrice \mathbf{R} de rotation de taille 3×3 . On peut donc réécrire l'équation 2.12 comme suit :

$$\mathbf{v}^T \mathbf{G} \mathbf{v} = \mathbf{v}^T \mathbf{Q} \mathbf{Q}^T \mathbf{v} = \mathbf{v}^T \mathbf{Q} \mathbf{R} \mathbf{R}^T \mathbf{Q}^T \mathbf{v} = \mathbf{v}^T \mathbf{Q}' \mathbf{Q}'^T \mathbf{v}. \quad (2.13)$$

où \mathbf{Q}' correspond à une matrice 6×3 contenant d'autres coefficients que ceux contenus dans \mathbf{Q} . Cette équation est valable pour toutes les matrices de rotation de taille 3×3 . Ce dernier résultat montre la non-unicité de la décomposition [GHOSH et coll. 2008].

L'avantage de cette décomposition est la contrainte pour le signal d'être positif. L'inconvénient est que cette décomposition n'est pas unique.

La deuxième décomposition utilisée en IRM de diffusion est la décomposition spectrale [BASSER & PAJEVIC 2007, RENARD et coll. 2010]). Pour ce faire, on considère le T4 comme une matrice symétrique grâce aux notations de Voigt. Ensuite, *via* le théorème spectral, cette matrice peut se décomposer en 6 valeurs propres μ_i et vecteurs propres \mathbf{D}_2^i . Les valeurs propres μ_i sont appelées valeurs propres de Kelvin. Les vecteurs propres \mathbf{D}_2^i correspondent à des T2 sans contrainte de positivité ni même de rang. La décomposition spectrale du tenseur \mathbf{D} d'ordre 4 peut s'écrire sous la forme :

$$\mathbf{D} = \sum_{i=1}^6 \mu_i \mathbf{D}_2^i \mathbf{D}_2^{iT}. \quad (2.14)$$

L'avantage de cette décomposition est son unicité. L'inconvénient est l'absence de signification physique des vecteurs propres pour l'IRM de diffusion.

Cette décomposition introduit une notion de valeurs propres. Qi et coll. proposent des notions de valeurs propres construites sur des principes géométriques [QI et coll. 2009], nommées

D-valeurs propres :

$$\lambda =_{opt} \sum_{i,j,k,l=1}^3 D_{ijkl} g_i g_j g_k g_l , \quad (2.15)$$

sous contrainte $\|\mathbf{g}\| = 1$.

L'opérateur *opt* correspond à la recherche d'optima locaux et de points selles. Ces valeurs propres λ correspondent donc aux maxima et minima locaux, ainsi qu'aux valeurs de points selles de la surface correspondant au T4. Cela suppose que dans le cadre d'un T4 strictement positif, les λ sont strictement positifs.

D'un point de vue géométrique, elles sont plus proches de la notion de valeurs propres usuelles utilisées pour les T2 que les valeurs propres de Kelvin, et possèdent de bonnes propriétés géométriques.

Par exemple, si on considère λ_{max} et λ_{min} les deux D-valeurs propres maximale et minimale respectivement, et la valeur moyenne M_S du tenseur définie par

$$M_S = \frac{1}{4\pi} \int_{\mathbb{S}^2} \sum_{i,j,k,l=1}^3 D_{ijkl} g_i g_j g_k g_l d\mathbf{g},$$

on obtient l'inégalité :

$$\lambda_{min} \leq M_S \leq \lambda_{max}. \quad (2.16)$$

Dans leurs travaux, Qi et coll. montrent une inégalité entre les valeurs propres de Kelvin et les D-valeurs propres, justifiant l'utilisation de ces dernières au détriment de ces premières [Qi et coll. 2009]. Soient μ_{max} et μ_{min} respectivement les valeurs propres maximum et minimum de Kelvin. L'inégalité correspond alors à :

$$\mu_{min} \leq \lambda_{min} \leq \lambda_{max} \leq \mu_{max}. \quad (2.17)$$

Prenons l'exemple donné dans [BARMPOUTIS et coll. 2009] pour mettre en évidence l'importance de cette dernière inégalité. Soit $d(\mathbf{g}) = g_1^4 + g_2^4 + g_3^4$. Nous obtenons alors $\mu_{min} = 0$, $\lambda_{min} = \frac{1}{4}$ et $\mu_{max} = \lambda_{max} = 1$. Grâce à cet exemple, nous observons que les D-valeurs propres sont plus proches de la signification des valeurs propres du tenseur d'ordre 2 que les valeurs propres de Kelvin. Il est bon de rappeler que ces différences proviennent du fait qu'on modélise une application de \mathbb{R}^3 dans \mathbb{R} avec une matrice modélisant plus généralement une application de \mathbb{R}^6 dans \mathbb{R} .

A noter aussi que l'équation proposée dans [Qi et coll. 2009] est plus générale que le cadre proposé par l'équation 2.15.

Métriques

La métrique la plus évidente est la métrique canonique euclidienne, que l'on peut estimer à partir des T4 sous forme vectorielle ou matricielle. Le principal inconvénient de cette métrique est l'absence de prise en compte de la positivité des T4. Plus précisément, un T4 non positif est à une distance finie de T4 positifs.

A partir de ce constat, Maher Moakher dans [MOAKHER 2008, MOAKHER 2009] propose des métriques riemannienne et log-euclidienne basées sur la forme matricielle des T4. Ces métriques sont l'extension des métriques proposées dans le cadre des T2 considérés comme des matrices symétriques définies positives. Cependant, Maher Moakher considère les T4 comme

des applications multilinéaires de \mathbb{R}^6 dans \mathbb{R} . Ce cadre convient lorsqu'on travaille sur des tenseurs d'élasticité [MOAKHER 2008, MOAKHER 2009] mais n'est pas adapté aux T4 de \mathbb{R}^3 dans \mathbb{R} utilisés en IRMd. Si on reprend l'exemple introduit dans [BARMPOUTIS et coll. 2009] avec $d(\mathbf{g}) = g_1^4 + g_2^4 + g_3^4$, ce tenseur n'est pas représenté par une matrice définie positive. En conséquence, avec la métrique riemannienne de Moakher, ce tenseur est à une distance infinie du tenseur représentant la sphère unité⁴ ($d(\mathbf{g}) = g_1^4 + g_2^4 + g_3^4 + 2 g_1^2 g_2^2 + 2 g_1^2 g_3^2 + 2 g_2^2 g_3^2$) alors qu'une métrique adaptée devrait le placer à une distance finie de la sphère.

L'utilisation d'une métrique basée sur la géométrie semble dès lors plus intéressante. L'article [BARMPOUTIS et coll. 2009] introduit la distance de Hellinger définie par :

$$dist_{Hellinger}(\mathbf{D}_1, \mathbf{D}_2) = \int_{\mathbb{S}^2} (d_1(\mathbf{g}) - d_2(\mathbf{g}))^2 d\mathbf{g}, \quad (2.18)$$

avec d_i la fonction définie sur la sphère paramétrée par le tenseur \mathbf{D}_i d'ordre 4. Cette distance peut être calculée analytiquement avec la formule suivante :

$$\begin{aligned} dist_{Hellinger}(\mathbf{D}_1, \mathbf{D}_2) = & d(\Delta_{400} + \Delta_{040} + \Delta_{004} + \Delta_{220} + \Delta_{022} + \Delta_{202})^2 \\ & + d(\Delta_{211} + \Delta_{031} + \Delta_{013})^2 + d(\Delta_{121} + \Delta_{301} + \Delta_{103})^2 \\ & + d(\Delta_{112} + \Delta_{310} + \Delta_{130})^2 + (b-d)(\Delta_{400} + \Delta_{040} + \Delta_{004})^2 \\ & + (b-d)\left((\Delta_{310} + \Delta_{130})^2 + (\Delta_{301} + \Delta_{103})^2 + (\Delta_{031} + \Delta_{013})^2\right) \\ & + (c-d)\left((\Delta_{400} + \Delta_{220})^2 + (\Delta_{400} + \Delta_{202})^2\right) \\ & + (c-d)\left((\Delta_{040} + \Delta_{220})^2 + (\Delta_{040} + \Delta_{022})^2\right) \\ & + (c-d)\left((\Delta_{004} + \Delta_{202})^2 + (\Delta_{004} + \Delta_{022})^2\right) \\ & + (c-b)\left(\Delta_{310}^2 + \Delta_{301}^2 + \Delta_{130}^2 + \Delta_{031}^2 + \Delta_{103}^2 + \Delta_{013}^2\right) \\ & + (a-b-2(c-d))\left(\Delta_{400}^2 + \Delta_{040}^2 + \Delta_{004}^2\right) \\ & + (a-b-2(c-d))\left(\Delta_{220}^2 + \Delta_{022}^2 + \Delta_{202}^2\right), \end{aligned} \quad (2.19)$$

avec $a=1/9$, $b=1/105$, $c=1/63$, $d=1/315$, et $\Delta_{ijk} = \mathbf{D}_{2,ijk} - \mathbf{D}_{1,ijk}$. Cette distance ne prend pas en compte la positivité du T4.

Angelos Barmpoutis et Baba Vemuri [BARMPOUTIS & VEMURI 2009] proposent une métrique *ad hoc* prenant en compte la positivité. On retrouve aussi cette métrique, dans un cadre plus général, dans des travaux d'Albert Tarantola [TARANTOLA 2005]. Cette métrique s'écrit de la manière suivante :

$$dist_{Tarantola}(\mathbf{D}_1, \mathbf{D}_2) \simeq \frac{1}{N} \sum_{i=1}^N (\log(d_1(\mathbf{g}_i)) - \log(d_2(\mathbf{g}_i)))^2, \quad (2.20)$$

où les \mathbf{g}_i sont N vecteurs échantillonnant la sphère, et d_i sont les fonctions définies sur la sphère paramétrée par le tenseur \mathbf{D}_i d'ordre 4. A noter que l'équation 2.20 n'est qu'une estimation, basée sur un nombre limité d'échantillons, de la distance $dist_{Tarantola}$.

4. L'équation de la sphère unité s'écrit $d(\mathbf{g}) = g_1^2 + g_2^2 + g_3^2$. Si on élève cette dernière équation au carré, on obtient l'équation de la sphère unité pour un tenseur d'ordre 4.

2. Les harmoniques sphériques

Un autre outil mathématique couramment utilisé pour la représentation de fonctions définies sur la sphère est les harmoniques sphériques (HSs). La représentation d'une fonction en série de HSs peut être vue comme une généralisation de la transformée de Fourier sur les fonctions définies sur la sphère, avec une multitude d'applications physiques [GROEMER 1996]. Comme pour les tenseurs, nous présenterons le formalisme, puis certaines caractéristiques des HSs, et finalement certaines métriques associées.

2.1 Représentation

La base de HSs s'écrit usuellement de la manière suivante :

$$Y_\ell^m(\theta, \varphi) = \sqrt{\frac{(2\ell+1)(\ell-|m|)!}{4\pi(\ell+|m|)!}} P_{\ell,|m|}(\cos\theta) \exp(im\varphi), \quad (2.21)$$

avec $\theta \in [0, \pi]$ et $\varphi \in [0, 2\pi[$, et $P_{\ell,m}$ le polynôme de Legendre d'ordre m et de degré ℓ . Pour obtenir des fonctions symétriques, on doit considérer uniquement les degrés ℓ pairs.

Les premiers travaux utilisant les HSs en IRM de diffusion sont [ANDERSON 2005] et [HESS et coll. 2006].

Descoteaux et coll. [DESCOTEAUX et coll. 2007] ont introduit un formalisme couramment utilisé de nos jours pour les HSs. Ils ont formulé un nouvel indexage j de la base des HSs d'ordre ℓ_{max} en fonction de ℓ et m : $j(\ell, m) = \frac{\ell^2 + \ell + 2}{2} + m$, avec $\ell \in \{0, 2, \dots, \ell_{max}\}$ et $m \in \{-\ell, -\ell+1, \dots, 0, \dots, \ell-1, \ell\}$. La base des HSs d'ordre ℓ_{max} est composée de N harmoniques, avec $N = \frac{1}{2}(\ell_{max}+1)(\ell_{max}+2)$. Par exemple, pour ℓ_{max} égale à quatre, N correspond à quinze.

La base correspondante avec le nouvel indexage est :

$$Y_j = \begin{cases} \sqrt{2} \operatorname{Re}(Y_\ell^{|m|}) & \text{si } m > 0, \\ Y_\ell^0 & \text{si } m = 0, \\ \sqrt{2} (-1)^{m+1} \operatorname{Im}(Y_\ell^m) & \text{si } m < 0. \end{cases} \quad (2.22)$$

Cette base a été introduite pour ne plus avoir de fonction à partie imaginaire non nulle. On peut donc maintenant décomposer une fonction S définie sur la sphère sous la forme :

$$S(\theta, \phi) = \sum_{j=1}^N c_j Y_j(\theta, \phi). \quad (2.23)$$

2.2 Caractéristiques

Plusieurs propriétés caractérisent les HSs. Comme leur nom l'indique, les HSs sont des fonctions harmoniques vérifiant l'équation de Laplace $\Delta Y = 0$. Une autre propriété importante est que, à part le premier harmonique Y_1 , les HSs sont des fonctions à valeurs positives et négatives. Ceci peut être un désavantage dans le cadre de l'IRMd car on cherche à décomposer une fonction strictement positive en une somme de fonctions à valeurs positives et négatives. Plus précisément, une combinaison linéaire quelconque de telles fonctions n'est pas forcément à valeurs positives.

L'utilisation des HSs en IRM de diffusion s'explique car elles permettent une expression analytique simple de la transformée de Funk-Radon, ce qui a permis l'essor de l'imagerie *q-ball* qui n'était jusque là possible qu'à partir d'une résolution numérique complexe [TUCH 2004].

Plus précisément, cette transformée permet d'estimer la probabilité de présence des paquets de fibres à partir du signal contenu dans les images pondérées en diffusion.

Une autre caractéristique intéressante est le fait que la base des HSs forme une famille orthogonale de fonctions définies sur la sphère. On a donc :

$$\langle Y_i | Y_j \rangle = \int_{\mathbb{S}^2} Y_i(\mathbf{g}) Y_j(\mathbf{g}) d\mathbf{g} = \delta_{ij}, \quad (2.24)$$

où $\langle . | . \rangle$ est le produit scalaire sur l'espace des fonctions définies sur la sphère, δ_{ij} est le symbole de Kronecker.

2.3 Métriques

Il existe une métrique naturelle induite par le produit scalaire ci-dessus. Soient f_1 et f_2 deux fonctions définies sur la sphère de vecteurs de coefficients \mathbf{c}^1 et \mathbf{c}^2 respectivement. On peut considérer la distance telle que :

$$\begin{aligned} dist(d_1, d_2) &= \frac{1}{4\pi} \int_{\mathbb{S}^2} f_1(\mathbf{g}) f_2(\mathbf{g}) d\mathbf{g} . \\ &= \sum_{j=1}^N c_j^1 c_j^2 \end{aligned} \quad (2.25)$$

Sous hypothèse que la fonction définie sur la sphère soit normalisée (*i.e.* que la norme au sens précédent soit égale à 1), on peut observer que les coefficients satisfont l'équation d'une hypersphère d'ordre $N - 1$, où N est le nombre de coefficients des HSs. Il existe une métrique riemannienne pour des éléments appartenant à une hypersphère [CHENG et coll. 2009] :

$$dist(d_1, d_2) = \arccos \left(\sum_{j=1}^N c_j^1 c_j^2 \right). \quad (2.26)$$

Cette métrique prend en compte l'espace sous-jacent. Cependant la contrainte de normalisation autorise encore des éléments interdits (par exemple la fonction sphérique constante de valeur -1).

On peut aussi employer la distance de Kullback-Leibler symétrisée adaptée pour des mesures positives [CHIANG et coll. 2007]. Soient D_p et D_q deux mesures positives. On peut écrire :

$$D_{sKL}(p, q) = \frac{1}{2} \left(\frac{1}{gtr(D_p)} \int_{\mathbb{S}^2} D_p \log \left(\frac{D_p}{D_q} \right) d\Omega + \frac{1}{gtr(D_q)} \int_{\mathbb{S}^2} D_q \log \left(\frac{D_q}{D_p} \right) d\Omega \right), \quad (2.27)$$

avec $gtr(D_i)$ l'intégrale de la mesure D_i sur la sphère unité \mathbb{S}^2 .

On peut reformuler cette dernière équation directement à partir des coefficients des HSs. Soient $D_i(\theta, \phi) = \sum_{j=1}^N c_j^i Y_j(\theta, \phi)$ et $\log(D_i(\theta, \phi)) = \sum_{j=1}^N d_j^i Y_j(\theta, \phi)$. On peut écrire alors :

$$D_{sKL}(p, q) = \frac{1}{2} \left(\frac{1}{c_1^p} \left[\sum_{i=1}^N c_i^p d_i^p - c_i^p d_i^q \right] + \frac{1}{c_1^q} \left[\sum_{i=1}^N c_i^q d_i^q - c_i^q d_i^p \right] \right). \quad (2.28)$$

On peut constater que le coefficient c_1^i correspond au volume moyen de D^i .

3. Lien entre tenseurs et harmoniques sphériques

Il existe un isomorphisme entre les tenseurs symétriques d'ordre N et les harmoniques sphériques de degré pair du même ordre. Une preuve est donnée dans [DESCOTEAUX et coll. 2006, DESCOTEAUX 2008].

D'une part les HSs d'ordre ℓ correspondent aux polynômes homogènes d'ordre ℓ définis sur la sphère. D'autre part, les polynômes correspondant aux tenseurs de diffusion d'ordre ℓ constituent l'ensemble unique des monômes de degré ℓ définis sur la sphère. Ces deux derniers ensembles sont linéairement indépendants et constitués du même nombre d'éléments. L'idée de la démonstration est de montrer que ces deux derniers ensembles, pour un même ordre ℓ , forment une base d'un même espace de fonctions.

Il est évident que les HSs d'ordre 2 sont inclus dans les HSs d'ordre 4. En raison de l'isomorphisme entre les tenseurs symétriques d'ordre N et les HSs de degré pair du même ordre, les tenseurs d'ordre 2 sont inclus dans les tenseurs d'ordre 4. Néanmoins cette inclusion est moins évidente que pour les HSs. Nous nous proposons de montrer que les tenseurs d'ordre 2 sont inclus dans les tenseurs d'ordre 4.

On considère l'équation suivante pour exprimer un tenseur d'ordre 2 comme un tenseur d'ordre 4 :

$$g_1^2 + g_2^2 + g_3^2 = 1. \quad (2.29)$$

Plus précisément, on multiplie l'équation du tenseur d'ordre 2 avec cette dernière équation pour obtenir des monômes de degré 4. Considérons un tenseur $d(\mathbf{g}, \mathbf{g})$ d'ordre 2, et multiplions par l'équation 2.29 :

$$\begin{aligned} d(\mathbf{g}, \mathbf{g}) &= (g_1^2 + g_2^2 + g_3^2) \left(\sum_{i,j=1}^3 D_{ij} g_i g_j \right), \\ &= \left(\sum_{i,j=1}^3 D_{ij} g_i g_j g_1^2 \right) + \left(\sum_{i,j=1}^3 D_{ij} g_i g_j g_2^2 \right) + \left(\sum_{i,j=1}^3 D_{ij} g_i g_j g_3^2 \right), \\ &= d(\mathbf{g}, \mathbf{g}, \mathbf{g}, \mathbf{g}), \end{aligned} \quad (2.30)$$

où $d(\mathbf{g}, \mathbf{g}, \mathbf{g}, \mathbf{g})$ est un tenseur d'ordre 4.

Ce lien existant entre les tenseurs d'ordre ℓ et les HSs du même ordre nous autorise à utiliser le modèle mathématique le plus adéquat lors de traitements.

4. Conclusion partielle

Depuis plus d'une décennie maintenant, le T2 a été utilisé en IRM de diffusion. Bien que les recherches continuent dans ce domaine, il existe certains consensus tels que l'utilisation des valeurs propres et vecteurs propres du T2 comme descripteurs, ou l'utilisation de la métrique log-euclidienne, bien que ces choix restent subjectifs et fortement dépendants des applications.

Cependant, pour les modèles d'ordre supérieur, les outils universels n'existent pas. Cependant, la richesse des modèles permet d'obtenir de nombreux descripteurs et décompositions ayant chacun des caractéristiques spécifiques. Il s'agit alors d'utiliser l'outil adapté pour la tâche à effectuer.

De plus, le lien étroit entre les tenseurs et les harmoniques sphériques permet d'augmenter le nombre des représentations d'un même objet mathématique.

Il est aussi intéressant de noter que les formes paramétriques sont complexes à maîtriser, et que les coefficients de ces formes sont difficilement interprétables. Dès lors, les représentations géométriques semblent plus pertinentes à utiliser.

Chapitre 3

Modèles en IRM de diffusion

1.	Modélisation par tenseur de diffusion	33
2.	Modélisation par diffusion apparente	35
2.1	Modèles à plusieurs compartiments	36
2.2	Modèle de l'imagerie du tenseur de diffusion généralisé	36
2.3	Modèle de l'imagerie du kurtosis de diffusion	37
3.	Modélisation par fonction de distribution de l'orientation	38
3.1	Fonction de distribution de l'orientation de diffusion	38
3.2	Fonction de distribution de l'orientation de fibres	38
4.	Conclusion partielle	39

Dans cette section, nous nous intéresserons à la modélisation locale de la diffusion en IRM. Les acquisitions, de plus en plus complexes, ont donné lieu à l'élaboration d'une multitude de modèles durant cette dernière décennie. Des états de l'art complets ont été proposés dans la thèse de Maxime Descoteaux [DESCOTEAUX 2008] ainsi que dans la thèse d'Haz-Edine Assemlal [ASSEMLAL 2010]. Ces deux thèses classifient les différents modèles en deux niveaux :

- le premier niveau correspond à la complexité du modèle d'acquisition,
- le deuxième niveau correspond à la nature des modèles (avec ou sans hypothèse sur le modèle physique sous-jacent).

Les acquisitions sphériques HARDI se sont généralisées dans les acquisitions cliniques, contrairement aux acquisitions cartésiennes sur l'ensemble de l'espace q , trop longues pour les patients. Par conséquent, nous nous intéresserons particulièrement aux modèles relatifs aux acquisitions HARDI de type sphérique. Nous proposons une classification des modèles locaux en deux groupes basée sur la physique :

- modèles étudiant le coefficient de diffusion apparente (*apparent diffusivity coefficient* (ADC)),
- modèles reflétant le PME, connue sous le terme de fonction de distribution de l'orientation (*orientation distribution function* (ODF)), et par conséquent l'orientation des fibres sous-jacentes.

Nous commencerons par étudier le modèle historique du tenseur de diffusion.

1. Modélisation par tenseur de diffusion

La modélisation de la diffusion par un tenseur (DTI) est à ce jour la technique la plus répandue en IRM de diffusion. En effet, elle permet grâce à des acquisitions simples (au minimum six images pondérées en diffusion et une image de référence) de visualiser *in vivo* la structure neu-

ronale sous-jacente. Dans le cadre du DTI, on suppose que le maximum de diffusion correspond à la direction privilégiée d'un faisceau de fibres mono-directionnel.

Plus précisément, sous hypothèse de diffusion gaussienne, l'équation de Stejskal-Tanner (éq. 1.6) relie l'acquisition $S(\mathbf{g}, b)$ dans la direction \mathbf{g} et le tenseur de diffusion \mathbf{D} :

$$S(\mathbf{g}, b) = S_0 \exp(-b \mathbf{g}^T \mathbf{D} \mathbf{g}), \quad (3.1)$$

avec \mathbf{D} le tenseur de diffusion correspondant à un tenseur d'ordre 2 défini positif car il modélise une diffusion.

On suppose que le logarithme du signal $S(\mathbf{g}_i, b)$, délivré par l'IRM, pour la direction \mathbf{g}_i de gradient peut s'écrire de la manière suivante :

$$\ln(S(\mathbf{g}_i, b)) = \ln(S_0) - b \mathbf{g}_i^T \mathbf{D} \mathbf{g}_i + n_i, \quad (3.2)$$

où n_i correspond à un bruit suivant une loi gaussienne de moyenne nulle et de variance σ^2 . On suppose que σ est identique pour toutes les directions. Cette hypothèse sur le bruit est vérifiée quand le rapport signal sur bruit des images pondérées en diffusion est grand. Sous cette hypothèse de bruit gaussien, le tenseur \mathbf{D} peut être estimé au sens des moindres carrés. A noter que le coefficient $\mathbf{g}^T \mathbf{D} \mathbf{g}$ est aussi appelé coefficient de diffusion apparente (*apparent diffusion coefficient* (ADC)).

On pose $\mathbf{Y} = [ADC_1, ADC_2, \dots, ADC_N]^T$ avec

$$ADC_i = -\frac{1}{b} \ln\left(\frac{S(\mathbf{g}_i, b)}{S_0}\right),$$

$\mathbf{X} = [D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}]^T$, $\mathbf{n} = [n_1, n_2, \dots, n_N]^T$ et

$$\mathbf{A} = \begin{bmatrix} g_{1;x}^2 & g_{1;y}^2 & g_{1;z}^2 & 2 g_{1;x} g_{1;y} & 2 g_{1;x} g_{1;z} & 2 g_{1;y} g_{1;z} \\ & & & \vdots & & \\ g_{N;x}^2 & g_{N;y}^2 & g_{N;z}^2 & 2 g_{N;x} g_{N;y} & 2 g_{N;x} g_{N;z} & 2 g_{N;y} g_{N;z} \end{bmatrix},$$

avec $g_{i;j}$ la j^{eme} composante de la i^{eme} direction \mathbf{g}_i de gradient, et N le nombre de directions de gradient.

Le système linéaire peut alors s'écrire sous la forme :

$$\mathbf{Y} = \mathbf{A} \mathbf{X} - \frac{1}{b} \mathbf{n}. \quad (3.3)$$

Au sens des moindres carrés, la solution correspond à l'équation suivante :

$$\widehat{\mathbf{X}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Y}. \quad (3.4)$$

Cette équation est résolue pour chaque voxel pour obtenir l'image du tenseur de diffusion. On peut remarquer que l'inversion de la matrice $\mathbf{A}^T \mathbf{A}$ nécessite l'utilisation d'au moins six directions de gradient \mathbf{g}_i non colinéaires entre elles.

Des estimations plus robustes ont été proposées dans [FILLARD et coll. 2007] en prenant en compte la positivité des tenseurs grâce aux métriques riemannienne et log-euclidienne. Lorsque le rapport signal sur bruit est faible, l'hypothèse sur le bruit (éq. 3.2) n'est plus exacte. Il faut alors prendre en compte un bruit de type ricien. Des estimations prenant en compte ce bruit ricien ont été effectuées dans [FILLARD et coll. 2007, WIEST-DAESSLÉ et coll. 2008].

L'imagerie du tenseur de diffusion est une imagerie complexe à visualiser car vectorielle. C'est pourquoi l'utilisation de coefficients intermédiaires reflétant différentes caractéristiques

du tenseur de diffusion a vu le jour. On pourra relever l'utilisation de la diffusion moyenne, de la fraction d'anisotropie (respectivement *mean diffusivity* (MD) et *fractional anisotropy* (FA)), ainsi que de la FA colorée.

La MD se calcule de la manière suivante :

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}, \quad (3.5)$$

où les λ_i sont les valeurs propres du tenseur.

Le critère le plus souvent utilisé est la FA :

$$FA = \sqrt{\frac{1}{2} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}}{\sqrt{(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}}. \quad (3.6)$$

La FA permet de rendre compte de l'anisotropie du milieu sous-jacent. Sa valeur est comprise entre 0, pour des milieux supposés isotropes, et 1, pour des milieux totalement anisotropes. Ce coefficient permet donc une visualisation très simple des zones de matière blanche. Cependant il manque la notion de la direction principale du tenseur. Pour résoudre ce problème, la FA colorée a été créée [PAJEVIC & PIERPAOLI 2000] : on affecte un code de couleur à chaque direction de l'espace, puis on affecte en intensité la valeur de la FA du voxel considéré.

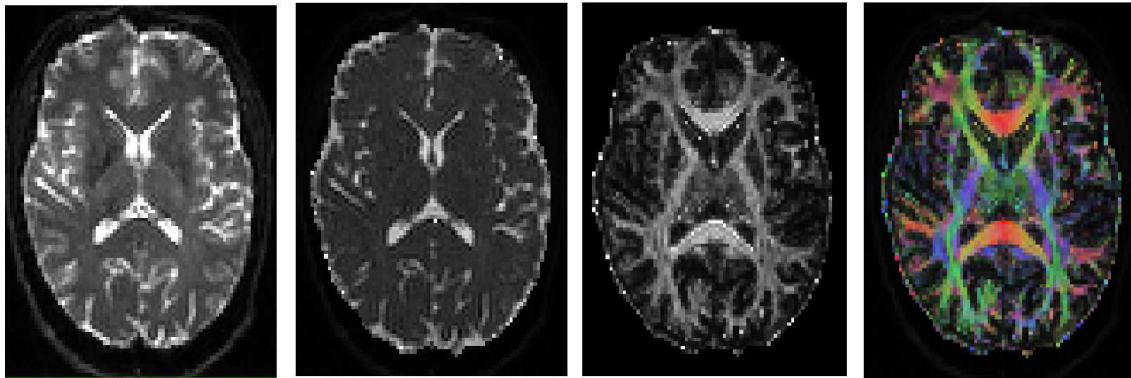


FIGURE 3.1 – De gauche à droite : l'image sans pondération, l'image MD, l'image de FA, et l'image de FA colorée.

De par leur simplicité et leur informativité, ces cartes sont largement utilisées en clinique ainsi que dans la communauté des chercheurs.

La plus grande limitation du tenseur de diffusion d'ordre 2 est de ne pouvoir modéliser qu'un faisceau de fibres mono-directionnel. En effet, si on cherche à modéliser un croisement de faisceaux de fibres par un tenseur de diffusion d'ordre 2, nous pouvons potentiellement obtenir un tenseur isotrope (*i.e.* un tenseur possédant la même valeur dans toutes les directions). Or un tenseur isotrope modélise aussi un milieu aqueux. On peut considérer que le tenseur de diffusion n'a pas assez de degrés de liberté pour distinguer un milieu aqueux d'un croisement de faisceaux fibres. A cause de cette indiscernabilité, de nouvelles stratégies ont vu le jour pour permettre de prendre en compte plus de degrés de liberté. Les deux stratégies principales sont une modélisation plus complexe de l'ADC et une modélisation reflétant plus le PME.

2. Modélisation par diffusion apparente

Une modélisation plus complexe de l'ADC peut être effectuée sous différentes hypothèses. Par la suite, nous nous intéresserons principalement à trois hypothèses donnant lieu à trois

modèles différents :

- les modèles à plusieurs compartiments,
- les modèles de généralisation du DTI,
- les modèles de l'imagerie du kurtosis de diffusion.

2.1 Modèles à plusieurs compartiments

L'hypothèse de ce modèle est que l'hétérogénéité de la diffusion intra-voxel provient principalement du fait qu'il existe une grande différence d'échelle entre la taille du voxel (quelques millimètres) et les différentes composantes ayant une influence sur la diffusion (cellules et axones ayant des dimensions de l'ordre du micromètre). L'idée est de modéliser alors chacune des composantes par un phénomène de diffusion homogène. Dans ce cadre, il existe principalement deux types de travaux : ceux séparant les diffusions intra- et extracellulaire [NIENDORF et coll. 1996], et ceux séparant les faisceaux de fibres ayant des directions différentes [TUCH et coll. 2002].

Nous intéresserons uniquement au deuxième cas. Supposons qu'il existe n fibres dans le modèle, et qu'il n'existe pas d'interaction entre ces n fibres. Le signal de diffusion peut alors être vu comme la somme des signaux de diffusion de chacune des n fibres :

$$S(\mathbf{g}) = \sum_{i=1}^n f_i \exp(-b\mathbf{g}^T \mathbf{D}_i \mathbf{g}). \quad (3.7)$$

Ce modèle est aussi connu sous le nom de mélange de gaussiennes. Il a été montré dans la littérature [TUCH et coll. 2002, KREHER et coll. 2005] la difficulté d'estimer ce type de modèle à compartiments. Récemment une estimation robuste des différents paramètres a été proposée [SCHERRER & WARFIELD 2010].

Dans [WEDEEN et coll. 2005, ÖZARSLAN et coll. 2006], les auteurs soulignent le fait que les maxima de diffusion apparente ne correspondent pas aux maxima de la probabilité de diffusion.

Nous pouvons visualiser dans la figure 3.2 différentes fonctions de diffusivité. On observe bien que le maximum de diffusion apparente ne correspond pas forcément au maximum de probabilité de diffusion.

On retrouve cette inadéquation dans les modèles de diffusion apparente d'ordre supérieur.

2.2 Modèle de l'imagerie du tenseur de diffusion généralisé

Un autre modèle, permettant de décrire des profils d'ADC plus complexes, a été proposé dans [ÖZARSLAN & MARECI 2003] et [ÖZARSLAN et coll. 2005]. Les auteurs présentent l'équation de Stejskal-Tanner généralisée :

$$S(\mathbf{g}) = S_0 \exp\left(-b \sum_{i_1, i_2, \dots, i_n=1}^3 D_{i_1 i_2 \dots i_n} g_{i_1} g_{i_2} \dots g_{i_n}\right), \quad (3.8)$$

où S_0 est le signal de référence, \mathbf{g} est le gradient de diffusion considéré, et $D_{i_1 i_2 \dots i_n}$ est le tenseur de diffusion d'ordre n et de dimension 3. Des estimations robustes ont été proposées prenant en compte la positivité du tenseur d'ordre 4 [GHOSH et coll. 2008, BARMPOUTIS et coll. 2009].

Nous avons montré dans la modélisation précédente que la diffusion apparente était capable de dissocier le cas du croisement de faisceaux de fibres du cas du milieu aqueux. Cependant cette fonction de diffusion apparente ne permet pas d'avoir l'information de probabilité de présence de fibres. On peut dès lors s'interroger sur l'utilité d'un tel modèle. L'un des avantages de l'imagerie du tenseur de diffusion généralisé est le calcul de l'anisotropie généralisée (*generalized anisotropy* (GA)), pendant de la FA en DTI. En effet, la GA permet d'avoir un meilleur contraste entre

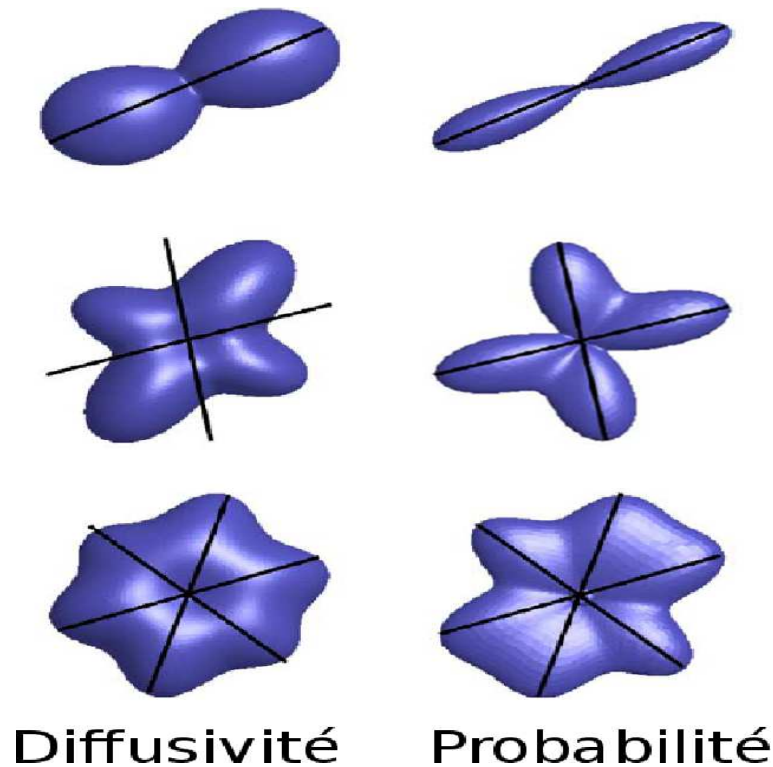


FIGURE 3.2 – La colonne de gauche représente des profils de diffusivité pour un, deux et trois paquets de fibres, tandis que la colonne de droite propose les PME's correspondants. On observe que les maxima des profils de diffusivité ne correspondent pas forcément aux maxima des PME's (voir par exemple la deuxième ligne). Schéma extrait de [ÖZARSLAN et coll. 2006].

les différents milieux. De plus, la GA a une valeur différente pour un milieu aqueux et pour un croisement de faisceaux de fibres. Plus précisément, la GA reflète la variance de la fonction définie sur la sphère modélisant le profil d'ADC.

2.3 Modèle de l'imagerie du kurtosis de diffusion

L'imagerie du kurtosis de diffusion (*diffusion kurtosis imaging* (DKI)) ne part plus de l'hypothèse que la diffusion est gaussienne.

Dans le cadre du DKI, un développement du logarithme du signal de diffusion est utilisé pour estimer de manière conjointe la diffusion apparente ainsi que le kurtosis apparent [JENSEN et coll. 2005, LU et coll. 2006, CHEUNG et coll. 2009]. On peut voir le kurtosis comme une mesure de la déviation d'une forme gaussienne.

On estime la diffusion apparente ainsi que le kurtosis apparent à partir de l'équation suivante :

$$\ln(S(\mathbf{g})/S_0) = -bD_{app}(\mathbf{g}) + b^2D_{app}(\mathbf{g})K_{app}(\mathbf{g}). \quad (3.9)$$

Une estimation robuste très récente a été proposée dans [BARMPOUTIS & ZHUO 2011].

Bien qu'on ne considère plus la diffusion comme gaussienne, on peut estimer les paramètres de ce modèle à partir d'une acquisition à b constant. Cependant la valeur de b doit être d'une valeur d'au moins $2000s/mm^2$ pour que la modélisation soit pertinente.

3. Modélisation par fonction de distribution de l'orientation

Dans la section précédente, nous venons de voir que l'ADC modélisé par un tenseur d'ordre supérieur permet de distinguer des croisements de faisceaux de fibres d'un milieu aqueux. Cependant les maxima de cette diffusion ne correspondent pas aux directions privilégiées des fibres. C'est pourquoi des modèles ont cherché à modéliser les fonctions de distribution de l'orientation des fibres (ODFs). Il existe deux grands type d'ODFs : les ODFs dites de diffusion (notées par la suite d-ODF) et les ODFs dites de fibres (notées par la suite f-ODF). Ces modèles sont souvent appelés modèles libres car ils ne se basent pas sur une équation physique, mais cherchent à modéliser le PME.

3.1 Fonction de distribution de l'orientation de diffusion

L'ODF de diffusion Ψ_{diff} peut être écrite formellement de la manière suivante :

$$\Psi_{diff}(\theta, \phi) = \int_0^\infty P(r, \theta, \phi) dr, \quad (3.10)$$

avec $\theta \in [0, \pi]$ et $\phi \in [0, 2\pi[$, et $P(r, \theta, \phi)$ le PME exprimé en coordonnées sphériques. On peut donc observer que l'ODF de diffusion est une surface reflétant le PME.

David Tuch [TUCH 2004] a proposé une première estimation numérique de cette ODF à partir d'une acquisition sphérique de l'espace q. Ce type d'imagerie porte le nom d'imagerie *q-ball*.

Des articles récents [ANDERSON 2005, HESS et coll. 2006] ont proposé une résolution analytique basée sur les HSs. Une obtention des coefficients de l'ODF de diffusion plus robuste est proposée dans [DESCOTEAUX et coll. 2007].

3.2 Fonction de distribution de l'orientation de fibres

L'ODF de fibres est la probabilité de localisation des fibres, estimée par deux catégories de techniques. La première stratégie [TOURNIER et coll. 2004, JIAN & VEMURI 2007] part du principe que le signal observé $S(\theta, \phi)$ (exprimé en coordonnées sphériques) est la convolution de l'ODF de fibres $\Psi_{fibre}(\theta, \phi)$ avec la modélisation de la diffusion $R(\theta)$ pour une fibre :

$$S(\theta, \phi) = \Psi_{fibre}(\theta, \phi) \otimes R(\theta). \quad (3.11)$$

La modélisation de la diffusion $R(\theta)$ pour une fibre correspond donc à son ODF de diffusion. La deuxième méthode, proposée dans [DESCOTEAUX et coll. 2007], part de cette constatation, et calcule l'ODF de fibres à partir de l'ODF de diffusion. Cette stratégie est mise en évidence dans la figure 3.3.

A noter que l'ODF de diffusion est une représentation floue de l'ODF de fibres. D'ailleurs le lien entre les deux ODFs est toujours une question ouverte. Plus précisément, l'influence des paramètres microscopiques des faisceaux de fibres tels que, par exemple, la densité des axones ainsi que leurs diamètres sur le phénomène de diffusion, reste mal connue.

On peut remarquer qu'il existe de nombreuses autres stratégies par exemple pour modéliser la probabilité de déplacement des molécules d'eau (par exemple celle appelée *persistent angular structure* (PAS) [JANSON & ALEXANDER 2003]) ou pour modéliser l'orientation des faisceaux de fibres (par exemple la technique appelée *diffusion orientation transform* (DOT) [ÖZARSLAN et coll. 2006]).

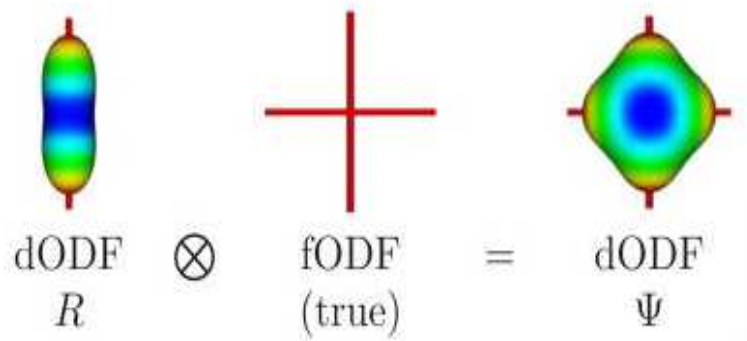


FIGURE 3.3 – Schéma modélisant la convolution de l’ODF de diffusion pour une fibre avec l’ODF de fibres réelle pour obtenir l’ODF de diffusion du voxel considéré. Schéma extrait de [LENGLET et coll. 2009].

4. Conclusion partielle

Au cours du temps, les études en IRM_d ont évolué vers des stratégies modélisant de mieux en mieux localement les structures anatomiques sous-jacentes.

Les nombreux modèles en IRM_d possèdent chacun des avantages et des inconvénients. Choisir un modèle dépend dès lors de l’application considérée.

Une attention particulière doit être accordée aux modèles permettant de distinguer les croisements de faisceaux de fibres. En effet, l’interprétation est différente entre l’ADC, attachée au signal de diffusion, et l’ODF, reliée à la structure des faisceaux de fibres sous-jacents.

Chapitre 4

Tractographie

1. Méthodes de tractographie	41
1.1 Algorithmes de tractographie déterministes	42
1.2 Algorithmes de tractographie probabilistes	43
1.3 Paramètres d'initialisation et critères d'arrêt	43
2. Limitations et validation des méthodes de tractographie	44
2.1 Limitations	44
2.2 Validation	45
3. Conclusion partielle	45

Dans les chapitres précédents, nous avons examiné l'influence des structures neuronales sur la diffusion locale à l'échelle du voxel et, plus précisément, l'influence de l'orientation des faisceaux de fibres de substance blanche. Les méthodes de tractographie correspondent aux différentes techniques permettant d'estimer ces faisceaux de fibres de substance blanche à l'échelle du cerveau. Bien que certaines modélisations représentent des fibres, le lecteur doit garder à l'esprit que les méthodes de tractographie ne peuvent représenter que des faisceaux de fibres. Cette limitation est due au fait que le voxel est à l'échelle du millimètre alors que le diamètre d'un axone est à l'échelle du micromètre.

La tractographie cérébrale est l'unique manière d'étudier des faisceaux de fibres de substance blanche *in vivo* [MORI et coll. 1999].

Cette possibilité d'étudier *in vivo* les fibres de substance blanche présente dès lors un grand intérêt clinique pour un nombre important de pathologies neurologiques et psychiatriques [CATANI 2006, YAMADA et coll. 2009] telles que l'épilepsie, la maladie d'Alzheimer, la sclérose en plaques et la schizophrénie.

Le but de ce chapitre est de fournir au lecteur un aperçu des différentes techniques de tractographie existantes. Dans un premier temps, nous nous attacherons à en montrer les principes de base. Nous discuterons ensuite de leurs limitations. Finalement, nous étudierons les méthodes de validation permettant de relier les résultats des tractographies à la réalité anatomique.

1. Méthodes de tractographie

De nombreux algorithmes de tractographie ont été proposés dans la littérature. Ils ont tous une structure commune.

Il faut d'abord choisir la classe de modèles locaux (DTI, ODF,...) qui va permettre de déterminer l'orientation des fibres sous-jacentes. Il faut ensuite choisir la méthode qui va générer

les trajectoires. Il existe deux grandes catégories d’algorithmes : les algorithmes déterministes et les algorithmes probabilistes. Les algorithmes déterministes créent une séquence de points en suivant de proche en proche les directions principales estimées à partir du modèle de diffusion. Le résultat d’une tractographie déterministe est donc une séquence de points. Les méthodes probabilistes sont basées sur les mêmes principes que les versions déterministes. La différence consiste à modéliser la direction principale par une loi probabiliste, permettant de prendre en compte, par exemple, l’erreur de reconstruction lors de l’estimation du modèle. On crée alors un grand nombre de séquences de points à partir d’une région de référence en prenant en compte la loi probabiliste. L’ensemble des séquences de points obtenues est difficilement interprétable telle quelle. On estime donc une image scalaire où chaque voxel contient le nombre de séquences de points passées à travers lui. Cette carte reflète de manière concise les connexions existant entre différents points de l’image et une région de référence.

Finalement, on doit choisir les zones d’initialisation et les critères d’arrêt pour chaque tractographie.

1.1 Algorithmes de tractographie déterministes

Les premières méthodes de tractographie ont été déterministes [CONTURO et coll. 1999] [MORI et coll. 1999, BASSER et coll. 2000, POUPON et coll. 2000]. Elles étaient basées sur les modèles de diffusion disponibles à l’époque, *i.e.* le tenseur de diffusion d’ordre 2.

Les premières méthodes sont basées sur la notion de ligne de courant (*streamline*). Les lignes de courant sont des familles de courbes qui sont tangentes en chaque point de mesure au champ vectoriel représentant l’écoulement. Dans le cadre de l’IRM de diffusion, l’écoulement correspond aux faisceaux de fibres, et les vecteurs composant le champ vectoriel aux vecteurs principaux des tenseurs d’ordre 2 (*i.e.* les directions privilégiées des faisceaux de fibres).

La première technique [MORI et coll. 1999], appelée *fiber assignment by continuous tracking* (FACT), consiste à ne prendre des points de mesure que lorsque la trajectoire sort du voxel courant. Plus précisément, les points de mesure sont pris à la frontière de chaque voxel. Aucune mesure n’est effectuée à l’intérieur du voxel. Cette technique très simple est à la base de toutes les tractographies.

Les premiers raffinements ont consisté à prendre des points de mesure à distance régulière, nécessitant alors une interpolation de la direction principale à chaque nouveau pas (voir par exemple [CONTURO et coll. 1999, BASSER et coll. 2000, POUPON et coll. 2000]).

Une seconde évolution a été de considérer l’ensemble du tenseur pour modifier la trajectoire [LAZAR et coll. 2003].

Les modèles de diffusion devenant plus complexes au cours du temps et permettant de mieux appréhender les phénomènes de croisement de faisceaux de fibres, les tractographies se sont basées sur eux pour estimer plusieurs directions privilégiées à l’intérieur d’un même voxel. Ces techniques ont été développées pour plusieurs modalités, tel que l’ODF [HAGMANN et coll. 2004, CAMPBELL et coll. 2005] et les mélanges de gaussiennes [KREHER et coll. 2005] [BERGMANN et coll. 2007].

D’autres techniques plus récentes reconstruisent la fibre en minimisant différents critères sur la globalité de la fibre [KREHER et coll. 2008, FILLARD et coll. 2009, REISERT et coll. 2011]. Ces dernières méthodes introduisent notamment des probabilités sur les directions principales. On les classe encore dans les tractographies déterministes car ces techniques fournissent des structures filaires et non des cartes de probabilité de présence comme les tractographies probabilistes. Ces dernières méthodes prenant en compte une information globale sont plus robustes au bruit dans les images.

1.2 Algorithmes de tractographie probabilistes

L'essor des algorithmes probabilistes provient de leur robustesse au bruit. Ils permettent donc une meilleure estimation de la substance blanche. Les premiers algorithmes sont apparus dès 2002 [KOCH et coll. 2002, PARKER & ALEXANDER 2003, JBABDI et coll. 2004] et ont connu depuis de nombreuses évolutions et développements [LAZAR & ALEXANDER 2005, RAMIREZ-MANZANARES & RIVERA 2006, JBABDI et coll. 2007, JBABDI et coll. 2008].

La différence fondamentale avec les algorithmes déterministes tient au fait que la tractographie ne se base plus uniquement sur une direction principale, mais prend en compte la fonction de densité de probabilité (*probability density function* (PDF)) de l'orientation de la fibre. Les principales différences dans les algorithmes probabilistes proviennent de la manière d'estimer ces PDFs [PARKER & ALEXANDER 2003, BEHRENS et coll. 2003, LAZAR & ALEXANDER 2005].

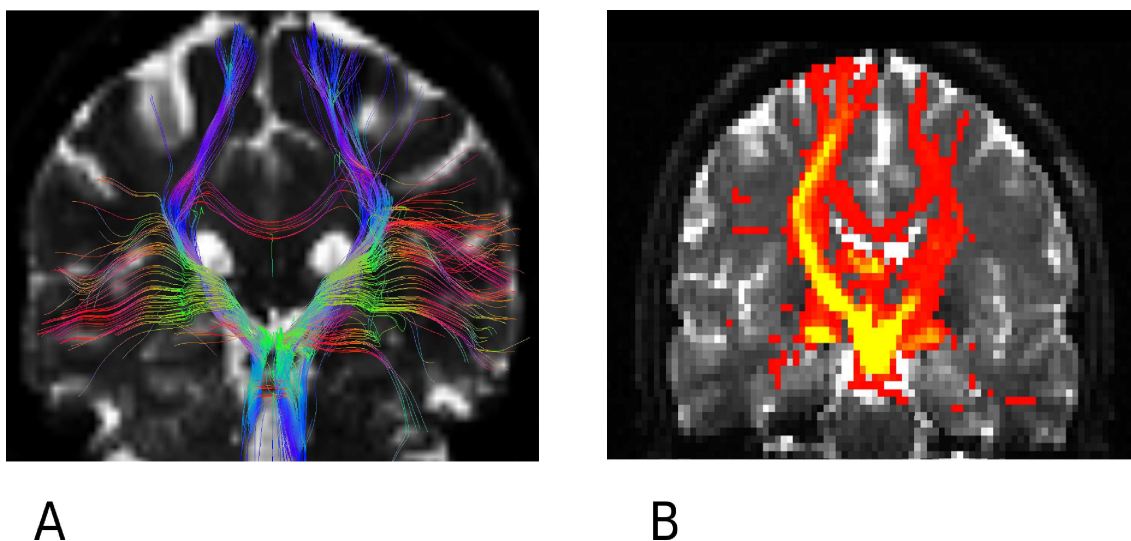


FIGURE 4.1 – A : tractographie déterministe d'après MedInria (<http://www-sop.inria.fr/asclepios/software/MedINRIA/>). B : tractographie probabiliste d'après FSL (<http://www.fmrib.ox.ac.uk/fsl/>). La région d'initialisation est placée dans le tronc cérébral pour chacune des tractographies. Nous pouvons observer la différence entre les deux tractographies : celle de gauche délivre des structures tubulaires, alors que celle de droite propose une carte de probabilité de présence.

1.3 Paramètres d'initialisation et critères d'arrêt

Une fois l'algorithme de reconstruction de la fibre choisi, il ne reste plus qu'à déterminer les paramètres d'initialisation et les critères d'arrêt de la tractographie.

Concernant les paramètres d'initialisation, tous les algorithmes de tractographie nécessitent le placement des « graines » (*seed*). Ces graines peuvent être contenues soit dans un voxel, soit dans une région d'intérêt (*region of interest* (ROI)). On peut dès lors penser à l'utilisation d'un atlas cérébral pour choisir ces ROIs. À noter qu'une tractographie déterministe peut avoir comme ROI l'ensemble du cerveau, contrairement à la tractographie probabiliste. En effet, dans le cadre de la tractographie probabiliste, on observe une carte de connectivité anatomique liée à une ROI de référence.

Il existe principalement deux critères d'arrêt. Le premier consiste à stopper l'algorithme dès lors que l'on rencontre un milieu avec une anisotropie trop faible pour correspondre à de la substance blanche, par exemple le liquide céphalorachidien des ventricules ou la matière grise. Le

deuxième critère correspond à la courbure de la fibre. Plus précisément, si la direction principale varie de façon trop importante, alors on émet l'hypothèse que le tractus suivi a changé. A ces deux critères d'arrêt, on peut ajouter certaines contraintes aux fibres afin qu'elles passent ou s'arrêtent dans des zones spécifiques du cerveau. En effet, il est parfois plus simple de placer des ROIs pour l'initialisation dans des zones où plusieurs faisceaux de fibres coïncident. La seule manière de ne conserver que le faisceau désiré est de placer une ROI complémentaire [WAKANA et coll. 2007].

2. Limitations et validation des méthodes de tractographie

Tout d'abord, il est bon de rappeler que les méthodes de tractographie ne permettent pas de représenter des fibres de substance blanche, mais des faisceaux de fibres. Dans cette section, nous allons nous intéresser tout d'abord aux limitations de ces méthodes pouvant conduire à des représentations erronées. Dans un deuxième temps, nous examinerons les différentes méthodes employées dans la communauté pour valider ces différentes techniques de tractographie.

2.1 Limitations

Intéressons-nous aux limitations des méthodes de tractographie déterministes puis probabilistes. Un article [JONES 2008] recense ces différentes limitations.

Les méthodes de tractographie déterministes présentent deux inconvénients. Premièrement ces techniques ne fournissent qu'une trajectoire par graine (un voxel peut contenir plusieurs graines), ce qui ne permet pas de visualiser des branchements de faisceaux de fibres (voir Fig. 4.2). En effet, on peut être en présence de phénomènes de volumes partiels, c'est-à-dire qu'il peut exister des faisceaux de fibres dans des directions différentes dans un même voxel [MORI et coll. 1999, WAKANA et coll. 2007]. Le second inconvénient est plus important : il n'existe aucune mesure de confiance sur la trajectoire suivie. A chaque itération lors de la construction de la trajectoire, cette dernière suit la direction privilégiée décrite par le modèle sans tenir compte de sa pertinence. Dans les zones où l'anisotropie est faible, l'estimation des directions privilégiées est nettement moins robuste au bruit. On peut donc obtenir des trajectoires erronées surtout dans les régions de substance grise ou proches des régions contenant du liquide céphalorachidien.

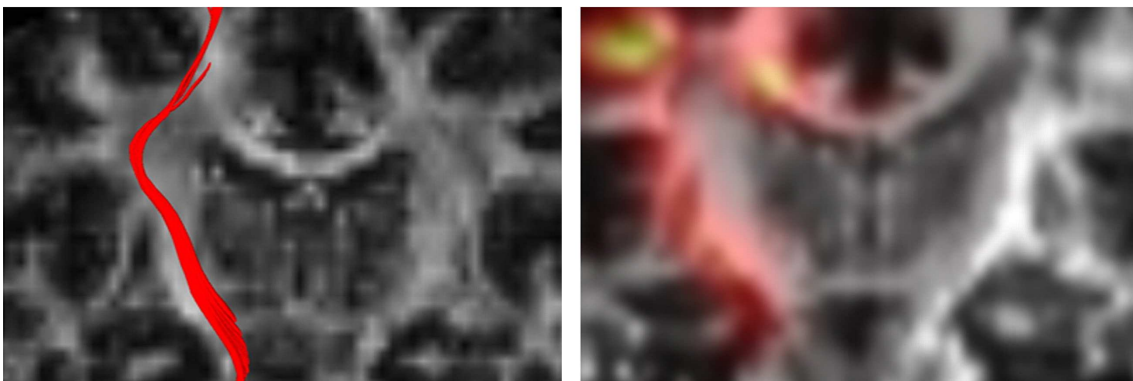


FIGURE 4.2 – Schéma représentant la tractographie déterministe (à gauche) et probabiliste (à droite). A gauche : ensemble de trajectoires obtenues avec un algorithme déterministe. A droite : une carte de probabilité de présence obtenue avec un algorithme probabiliste. On observe que la tractographie déterministe ne pourra pas rendre compte de branchements complexes contrairement à la tractographie probabiliste. D'après une figure de l'article [NUCIFORA et coll. 2007].

Les méthodes probabilistes prennent en compte l'incertitude sur le modèle et fournissent des intervalles de confiance sur le résultat. Cependant il existe toujours des erreurs lors de l'estimation des modèles, ce qui conduit, comme dans le cas déterministe, à des faux positifs et des faux négatifs.

Par ailleurs, l'interprétation des résultats fournis par les méthodes probabilistes pose un autre problème [JONES 2008]. Plus une trajectoire est longue, plus l'erreur de trajectoire accumulée le long de la trajectoire est grande. Il arrive donc que des régions correspondant à la fibre, mais éloignées de la graine de référence, aient la même probabilité d'être connectées à la graine de référence que des régions erronées mais très proches de la graine de référence. Les valeurs fournies par ces méthodes n'ont pas de réalité anatomique. Elles sont fortement dépendantes des graines de référence choisies.

2.2 Validation

Dès les premiers articles traitant de la tractographie (voir par exemple [MORI et coll. 1999, CONTURO et coll. 1999]), la validation a été une question abordée. Les premiers types de validation consistent en des comparaisons visuelles entre les anatomies de l'animal et de l'homme [MORI et coll. 1999, CONTURO et coll. 1999, POUPON et coll. 2000]. Ces types de validation sont en pleine expansion actuellement, et s'orientent vers une analyse conjointe avec des dissections anatomiques [DYRBY et coll. 2007, LAWES et coll. 2008].

Cependant, nous avons vu que la correspondance entre les modèles de diffusion et le substrat anatomique reste une question ouverte [TUCH et coll. 2002, PERRIN et coll. 2005]. Effectuer des comparaisons à l'échelle du cerveau semble complexe. Dès lors, des équipes de chercheurs ont mis en place des fantômes et ont effectué des acquisitions d'IRM de diffusion. Ces fantômes peuvent être synthétiques [PERRIN et coll. 2005, POUPON et coll. 2008] : ils consistent en des objets fibreux placés dans un milieu aqueux. Il existe aussi des fantômes anatomiques. Dans [CAMPBELL et coll. 2005], des moelles épinières de rat ont été extraites et placées dans un milieu aqueux. Les fantômes sont pourvus de croisements de fibres. Le développement de fantôme reste un sujet d'étude actuel [POUPON et coll. 2008, POUPON et coll. 2010].

Une « compétition » de tractographie a eu lieu durant la conférence MICCAI 2009 qui s'est déroulée à Londres. L'idée était de proposer un fantôme comportant tous les problèmes possibles rencontrés en tractographie (*crossing, kissing, demi-tour de la fibre ...*), et de comparer les différentes tractographies. Un article résume toutes les propriétés et les résultats de cette compétition [FILLARD et coll. 2011]. Il en ressort qu'en présence d'un signal faiblement bruité, les méthodes de tractographie déterministes et probabilistes fonctionnent correctement. Dans les zones où le rapport signal sur bruit est faible, les méthodes introduisant des *a priori* spatiaux sur les modèles de diffusion ou les faisceaux de fibres fonctionnent mieux.

3. Conclusion partielle

L'importance de l'IRMd réside dans la possibilité de déterminer *in vivo* les faisceaux de fibres de substance blanche. Par conséquent, la tractographie joue un rôle-clé dans cette modalité.

Au cours du temps, les algorithmes de tractographie se sont complexifiés pour rendre compte au mieux des structures anatomiques sous-jacentes, et plus particulièrement des croisements de fibres. Cependant, les stratégies de tractographie restent encore fortement dépendantes des paramètres employés. On peut observer qu'un véritable effort a été produit pour comparer et valider les différentes méthodologies.

Ces différentes méthodes de tractographie fournissent actuellement une quantité importante d'information sous forme d'objets difficiles à manipuler, à traiter et à interpréter tels qu'un

ensemble de trajectoires ou des cartes de probabilité de présence. Cependant l'analyse à l'échelle du cerveau est un enjeu majeur dans la compréhension du fonctionnement cérébral.

Chapitre 5

Connectivité et connectomes

1. Connectivité	47
1.1 Différentes échelles de connectivité	47
1.2 Différents types de connectivité	48
1.3 Conclusion partielle	49
2. Connectome	49
2.1 Historique	49
2.2 Création du connectome	50
2.3 Outils mathématiques : graphe ou matrice ?	53
3. Conclusion partielle	57

Nous avons vu dans le chapitre précédent les différentes techniques de tractographie permettant de relier des structures cérébrales entre elles. Dans ce chapitre, nous allons d'abord présenter le concept de connectivité cérébrale en imagerie. Dans un deuxième temps, nous nous intéresserons plus particulièrement à la connectivité anatomique et aux connectomes. Le connectome correspond à l'intégration de toute l'information fournie par la tractographie à l'échelle du cerveau. Finalement nous examinerons le lien entre les connectivités anatomique et fonctionnelle. Au travers de ce chapitre, ce sont les enjeux ainsi que les défis liés à l'étude de la connectivité que nous tenterons d'explicitier.

1. Connectivité

De manière très générale, la connectivité cérébrale renvoie à l'étude des liens anatomiques (connectivité anatomique), des dépendances fonctionnelles (connectivité fonctionnelle) ou des interactions fonctionnelles de cause à effet (connectivité effective) entre des entités distinctes à l'intérieur d'un système nerveux [SPORNS 2007]. Nous utilisons le terme d'entité car la connectivité peut s'étudier à l'échelle du neurone (échelle microscopique), d'une population de neurones (échelle mésoscopique) ou de régions cérébrales anatomiquement séparées (échelle macroscopique).

Nous allons présenter les différentes échelles d'étude possibles ainsi que les différents types de connectivité.

1.1 Différentes échelles de connectivité

Dans les systèmes nerveux évolués, la connectivité cérébrale peut être décrite à différentes échelles [KÖTTER 2007, SPORNS 2010, SPORNS 2010]

Il existe ainsi trois échelles dans la création du connectome. Premièrement considérons l'échelle microscopique où le neurone constitue l'élément de base. Il existe environ 10^{11} neurones dans le cerveau humain et 10^{15} connexions entre eux. De plus, d'un point de vue anatomique, des variations à l'échelle du neurone n'ont pas d'effet macroscopique notable. Les tâches cognitives humaines nécessitent l'activation et la co-activation d'un grand nombre de neurones [MOUNTCASTLE 1998]. D'ailleurs, les neurones et leurs inter-connexions sont sujets à des modifications liées à l'apprentissage des tâches fonctionnelles par exemple. Ce phénomène est plus connu sous le nom de plasticité cérébrale [ENGERT & BONHOEFFER 1999]. Du fait du grand nombre de neurones, de leur grande variabilité et de leur dynamique, il n'est pas envisageable de réaliser l'étude du cerveau complet à l'échelle microscopique.

Malgré ces inconvénients, le neurone a l'avantage d'être une entité simple à définir. Cependant, à l'échelle macroscopique, la question de la ségrégation des aires cérébrales reste ouverte. De plus, il a été montré que le choix de la parcellisation avait un réel impact sur l'étude du connectome [ZALESKY et coll. 2010]. On entend par parcellisation la division du cerveau en ROIs, correspondant souvent à des aires cérébrales précises [TZOURIO-MAZOYER et coll. 2002]. Bien qu'il n'existe pas de parcellisation idéale, on peut considérer que le cerveau peut être divisé en une centaine d'aires cérébrales. A noter aussi que selon les régions considérées, la parcellisation doit être plus ou moins fine. Par exemple, le thalamus peut être subdivisé en une dizaine de régions [BEHRENS et coll. 2003].

L'étude des connexions entre ces régions a débuté au travers de dissections ou de méthodes utilisant des traceurs axonaux. Les premières études *in vivo* sur les connectomes chez l'homme ont vu le jour grâce aux modalités d'IRM de diffusion ainsi qu'aux techniques de tractographie. Ces dernières méthodes donnent une idée de la pertinence de cette échelle macroscopique. En effet, les trajectoires obtenues par la tractographie entre deux régions corticales présentent une forte similarité avec l'anatomie sous-jacente. De plus, sur l'ensemble du cerveau, les trajectoires permettent d'effectuer une bonne ségrégation en différentes aires. La pertinence de cette parcellisation peut être confortée par le couplage du connectome avec des études fonctionnelles (voir [KOCH et coll. 2002] par exemple).

Si on se place dans le cas critique où chaque voxel de l'image correspond à une unité indépendante, le nombre d'éléments considérés serait de l'ordre de $10^4 - 10^5$ avec entre 10^5 et 10^7 connexions entre eux. Du point de vue du traitement informatique, cette quantité d'information est envisageable.

A ces deux échelles, micro et macroscopiques, s'ajoute une échelle intermédiaire appelée mésoscopique. En effet, l'échelle macroscopique ne permet pas à elle seule de comprendre l'ensemble du fonctionnement cérébral, et en particulier sa capacité à traiter l'information. Il existe une entité structurellement indépendante et correspondant à un fonctionnement fonctionnel propre : les mini-colonnes corticales [MOUNTCASTLE 1998]. Ces dernières structures contiennent environ de 80 à 100 neurones. Elles ont pour diamètre environ entre 30 et 50 micromètres. Le cerveau comprendrait de 10^7 à 10^8 de ces éléments. Ces structures sont trop fines et nécessitent une information macroscopique afin de connaître leur rôle et savoir où elles se situent dans le cerveau.

Dans la suite de cette thèse, nous nous intéresserons uniquement à l'échelle macroscopique qui est l'échelle pertinente en IRM de diffusion.

1.2 Différents types de connectivité

Il existe principalement trois types de connectivités : les connectivités anatomique, fonctionnelle et effective.

La connectivité anatomique correspond aux réseaux physiques entre les différentes aires

cérébrales. Ces réseaux apparaissent comme relativement stables sur des périodes très brèves (de l'ordre de la seconde à la minute) ainsi qu'à des échelles de temps plus grandes (de l'heure à plusieurs jours). A l'échelle macroscopique, la modalité d'IRM de diffusion est pour l'instant la seule permettant d'étudier *in vivo* la connectivité anatomique.

La détermination de la connectivité fonctionnelle se base sur des concepts statistiques. L'idée est de détecter des aires cérébrales qui ont une activité statistiquement dépendante au cours du temps. Cette dépendance peut être mesurée par des estimations de corrélation, de covariance, de cohérence spectrale... La connectivité fonctionnelle doit être étudiée pour chaque couple d'aires cérébrales sans *a priori* sur de possibles liens anatomiques. Contrairement à la connectivité anatomique, la connectivité fonctionnelle dépend fortement du temps. Plus précisément, on observe si les paires d'aires cérébrales ont des activations qui fluctuent dans le temps de manière similaire. Les variations des activations sont de l'ordre de la milliseconde à la seconde. Différentes modalités telles que l'IRM fonctionnelle ou l'électro-encéphalographie (EEG) permettent d'étudier l'activité cérébrale. La connectivité fonctionnelle ne permet pas de déterminer les possibles relations de cause à effet entre l'activation des différentes aires cérébrales.

La connectivité effective cherche à déterminer les effets directionnels des aires cérébrales les unes sur les autres. Plus précisément, la connectivité effective permet de décrire les phénomènes de cause à effet entre les régions (*i.e.* est-ce qu'une région **A** a activé une région **B**). L'étude de cette connectivité peut se faire grâce à des études conjointes de la connectivité anatomique et fonctionnelle. Ce type de connectivité ne sera pas abordé dans cette thèse.

1.3 Conclusion partielle

Les systèmes nerveux, et plus spécifiquement le système nerveux humain, sont des ensembles complexes pouvant être représentés par des réseaux. La compréhension complète de ces systèmes nécessite des études à différentes échelles. De plus, la seule connaissance anatomique ou fonctionnelle du cerveau ne permet pas d'appréhender tous les mécanismes du cerveau. Plus précisément, une étude conjointe de ces différentes connectivités laisse envisager une meilleure compréhension globale du cerveau.

Dans la suite de ce chapitre, nous nous intéresserons particulièrement à la connectivité anatomique à l'échelle macroscopique. Les liens entre connectivités anatomique et fonctionnelle seront abordés dans la partie modélisation et traitements statistiques.

2. Connectome

La description complète de la connectivité anatomique à l'échelle du cerveau est plus connue sous le nom de connectome ou de connectogramme. Dans cette section, nous proposons tout d'abord un bref historique du connectome. Nous détaillerons ensuite la construction de ce connectome. Nous décrirons dans une dernière partie certains outils mathématiques utilisés dans l'étude du connectome.

2.1 Historique

Le terme de connectome a été introduit en 2005 indépendamment par Olaf Sporns [SPORNS et coll. 2005] et par Patric Hagmann [HAGMANN 2005]. Ce terme a été choisi par analogie avec le génome, qui est la description complète de l'information génétique.

La volonté d'étudier les systèmes nerveux de manière globale est relativement ancienne. La première étude [WHITE et coll. 1986] a été effectuée sur un type de ver nommé *C. Elegans*, choisi en raison de la simplicité de son système nerveux. Une étude plus tardive, réalisée sur un

système nerveux plus complexe (celui d'un primate), a été proposée par [ESSEN et coll. 1992]. Celle-ci s'est focalisée sur le système visuel en étudiant à la fois les connectivités anatomique et fonctionnelle. On peut aussi noter l'étude sur le système thalamo-cortical du chat dans [SCANNELL et coll. 1999]. Cette dernière étude constitue le véritable précurseur des études de connectome. En effet, à partir de l'ensemble des connexions thalamo-corticales, Scannell et coll. sont parvenus à extraire la structure d'un réseau et à parcelliser le cerveau en aires anatomiques. L'une des premières études sur la connectivité anatomique chez l'homme date de 2002 [KOCH et coll. 2002]. Celle-ci montre le lien existant entre connectivités anatomique et fonctionnelle.

Avec l'apparition de l'IRM de diffusion et le développement des méthodes de tractographie, les études liées au connectome chez l'être humain se sont développées. Cependant, construire un connectome à partir d'acquisitions d'IRM de diffusion nécessite de nombreux pré-traitements. Certains laboratoires se sont donc regroupés en consortiums pour partager leurs compétences. Il en existe trois à l'échelle mondiale :

- le consortium américain *the Human Connectome Project* (<http://www.humanconnectomeproject.org/>),
- le consortium européen *CONNECT* (<http://www.brain-connect.eu>),
- le consortium asiatique *Brainnetome* (<http://www.brainnetome.org>).

Les thèmes abordés par ceux-ci sont aussi bien d'ordre médical que portés sur le traitement d'images médicales.

De par sa complexité, le connectome soulève donc de nombreux défis techniques. Une agence américaine (*the National Academy of Engineering* <http://www.nae.edu>) traitant de l'ingénierie a d'ailleurs considéré les problèmes scientifiques liés aux connectomes comme l'un des 14 *grands défis pour l'ingénierie* [ROYSAM et coll. 2009].

2.2 Création du connectome

La construction du connectome nécessite trois grandes étapes : la parcellisation, la tractographie et le calcul de la mesure de connectivité anatomique. La tractographie ayant été vue au chapitre précédent, nous nous intéresserons principalement au choix de la parcellisation et au choix de la mesure de connectivité. On peut représenter le connectome sous forme matricielle, où les coefficients de la matrice correspondent aux mesures de connectivité, et les lignes et les colonnes correspondent aux différentes ROIs étudiées.

La Fig. 5.1 représente le processus de création du connectome.

Parcellisation

Comme nous avons pu l'observer dans la section 1.1, il n'existe pas de parcellisation idéale du cerveau. Tout d'abord le nombre de régions considérées peut varier de la centaine au millier. Cette variation a un réel impact sur les propriétés des réseaux [ZALESKY et coll. 2010].

Actuellement, la parcellisation consiste à transporter un atlas déjà segmenté tel l'atlas *automatic anatomical labelling* (AAL) [TZOURIO-MAZOYER et coll. 2002]. Cette technique a pour avantage d'être relativement simple à implémenter. De plus, cette stratégie est reproductible à l'échelle d'une population. L'inconvénient majeur est que la parcellisation ne prend en compte que faiblement la variabilité inter-individuelle. Plus précisément, l'atlas correspond à une moyenne d'individus. Il se peut que la parcellisation ne soit pas adaptée pour un individu donné. Cette non-adéquation entre l'atlas et l'individu est d'autant plus apparente que la parcellisation est fine. Pour pallier ce défaut, le nombre de régions utilisées dans cette stratégie est relativement faible (90 régions pour l'atlas AAL).

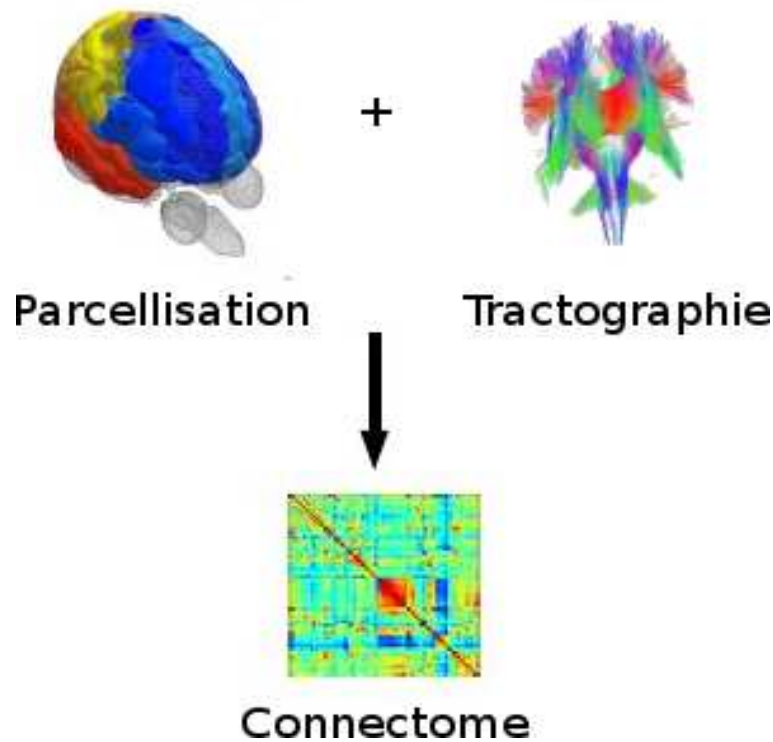


FIGURE 5.1 – Schéma représentant la création d'un connectome.

Hagmann et coll. [HAGMANN et coll. 2008] proposent de partir d'une parcellisation de 66 régions communes à l'ensemble des individus en recalant un atlas, puis de subdiviser pour chaque individu séparément ces premières régions pour obtenir 998 régions. Cette technique a l'avantage d'assurer une certaine reproductibilité entre les individus tout en possédant une parcellisation assez fine et adaptée à chaque individu.

On peut noter qu'il existe aussi des techniques de parcellisation de la matière grise grâce à la tractographie [ROCA et coll. 2009, PERRIN 2006], l'idée sous-jacente étant que des faisceaux de fibres similaires d'un point de vue géométrique appartiennent à une même aire corticale. Le problème de parcellisation consiste à effectuer une classification des différentes trajectoires obtenues par tractographie. Ces techniques semblent prometteuses car la parcellisation est définie spécifiquement pour chaque individu par construction.

Mesure de connectivité

Une fois la parcellisation et la tractographie effectuées, une grandeur doit être choisie. Cette grandeur, appelée mesure de connectivité, doit refléter les liens existant pour chaque couple de régions de la parcellisation. Il existe deux grandes catégories de mesures :

- les mesures quantitatives, reflétant le nombre de faisceaux de fibres passant d'une région à une autre,
- les mesures qualitatives, permettant de décrire les propriétés des faisceaux de fibres.

Mesures quantitatives : comme vu précédemment, la tractographie n'est qu'une représentation des faisceaux de fibres. De plus, les deux grands types de tractographie (déterministe et probabiliste) fournissent des résultats différents (ensemble de trajectoires ou carte de connectivité). Pour finir, chaque résultat est fortement dépendant des paramètres d'initialisation

(choix du nombre de graines dans le cadre déterministe, choix du nombre d'itérations dans le cadre probabiliste).

Pour pallier ces inconvénients, les mesures quantitatives doivent tenir compte de ces différents paramètres. Dans [YO et coll. 2009], les auteurs considèrent que la mesure de connectivité anatomique doit refléter l'influence de l'activité neuronale moyenne N_A d'une région **A** sur l'activité neuronale moyenne N_B d'une région **B**. Soient W_A et W_B la taille des régions de départ et d'arrivée respectivement (proportionnelle au nombre de neurones et au nombre de voxels), et F le nombre de faisceaux de fibres connectant ces deux régions (proportionnel au nombre de trajectoires ou au coefficient des cartes de connectivité). On peut décrire l'influence $C_{A \rightarrow B}$ de la région **A** sur la région **B** par le rapport N_B sur N_A . On considère que $C_{A \rightarrow B}$ quantifie la connectivité de la région **A** vers la région **B**. Ce coefficient ne peut pas être calculé directement, mais il est proportionnel au rapport suivant :

$$C_{A \rightarrow B} = \frac{N_B}{N_A} \propto \frac{F}{W_B}. \quad (5.1)$$

Dans le cadre d'une tractographie probabiliste, avant d'obtenir une carte de connectivité, on tire n trajectoires selon des lois de probabilités. Il faut alors prendre en compte le nombre n de ces itérations utilisées au cours de l'algorithme. Le rapport est alors proportionnel à :

$$C_{A \rightarrow B} \propto \frac{F}{nW_B}. \quad (5.2)$$

On peut donc observer que la connectivité de **A** sur **B** est différente de celle de **B** sur **A**. Prenons un exemple de deux ROIs **A** et **B** contenant respectivement peu et beaucoup de voxels. Supposons qu'il existe des connexions entre ces deux régions. Par principe, il y aura moins de paquets de fibres partant de la région **A** que de la région **B**. On observera donc des connectivités différentes entre deux régions selon que l'on considère l'une ou l'autre région (voir Fig. 5.2). De plus, on remarque que cette connectivité dépend fortement de la taille de cette région. Le coefficient $C_{A \rightarrow B}$ permet de prendre en compte ces variations de la taille des régions dans la parcellisation.

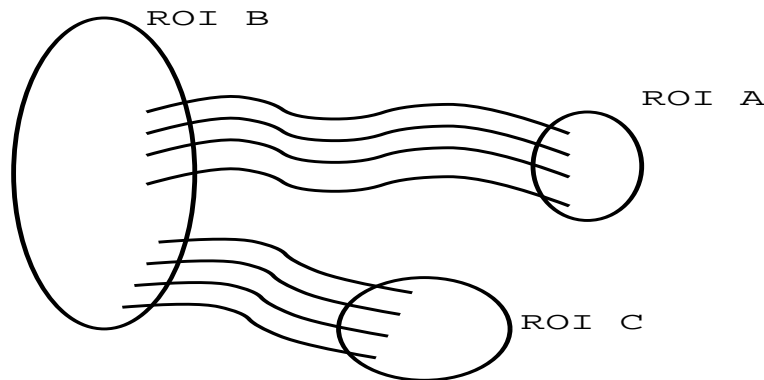


FIGURE 5.2 – Schéma représentant la connectivité entre trois ROIs **A**, **B** et **C**. On observe par exemple que la ROI **A** est petite, et que l'ensemble de ses paquets de fibres est relié à la ROI **B**. Par contre, la ROI **B** est plus grande, et ses paquets de fibres sont reliés à la fois à la ROI **A** et à la ROI **C**. La connectivité entre la ROI **A** et la ROI **B** est différente selon que l'on considère l'une ou l'autre région.

Il existe des variantes de ce coefficient [HAGMANN et coll. 2008], où l'orientation des mesures n'est plus considérée mais où la longueur des fibres intervient maintenant dans le calcul, péna-

lisant les longues fibres :

$$C_{A,B} = \frac{2}{W_A + W_B} \sum_{f \in F_{A,B}} \frac{1}{\ell(f)}, \quad (5.3)$$

où $F_{A,B}$ correspond à l'ensemble des fibres possédant une extrémité dans la région \mathbf{A} et l'autre extrémité dans la région \mathbf{B} , et $\ell(f)$ désigne la longueur de la fibre f .

Mesures qualitatives : des études récentes ont montré la pertinence d'observer le long des faisceaux de fibres des coefficients physiques, telle que la FA [COROUGE et coll. 2006], ou géométriques, tel que le vecteur propre principal [ZHU et coll. 2010].

Intégrer ces informations le long des fibres reliant deux régions \mathbf{A} et \mathbf{B} permet d'obtenir une mesure qualitative de connectivité entre les régions \mathbf{A} et \mathbf{B} :

$$C_{A,B}^f = \frac{1}{L} \int_0^L f(X(s)) ds, \quad (5.4)$$

avec s l'abscisse curviligne, $X(s)$ la position à l'abscisse s , et $f(X(s))$ le coefficient physique à la position $X(s)$. A noter que cette connectivité est définie pour une fibre reliant les régions \mathbf{A} et \mathbf{B} . La connectivité globale $C_{A,B}$ est la moyenne des différentes connectivités $C_{A,B}^f$:

$$C_{A,B} = \frac{1}{\text{Card}(F_{A,B})} \sum_{f \in F_{A,B}} C_{A,B}^f, \quad (5.5)$$

où $\text{Card}(F_{A,B})$ est le nombre de fibres reliant les régions \mathbf{A} et \mathbf{B} .

Les mesures locales peuvent correspondre à des mesures physiques, telles que la diffusivité moyenne ou l'anisotropie (la FA ou des coefficients basés sur les ODFs [PÉCHAUD et coll. 2009, ROBINSON et coll. 2010] ou sur les images pondérées en diffusion [DE BOER et coll. 2011]).

Une mesure récente, introduite dans [ZALESKY et coll. 2010], est basée sur la géométrie. L'idée est de considérer les faisceaux de fibres comme des câbles transportant de l'information, et de quantifier la capacité de ces câbles à convoier de l'information. Pour un faisceau de fibres, la capacité correspond à l'aire de la section transversale du faisceau.

Des études très récentes [WEE et coll. 2010, WEE et coll. 2011] ont pris en compte simultanément des mesures quantitatives et qualitatives pour obtenir une meilleure description de la connectivité.

2.3 Outils mathématiques : graphe ou matrice ?

Lors de la création du connectome, une quantité importante d'information est extraite des acquisitions d'IRMd. Si nous considérons une connectivité non directionnelle avec une parcellisation possédant N aires cérébrales, nous obtenons $\frac{N(N+1)}{2}$ coefficients. Si maintenant nous considérons une connectivité directionnelle avec la même parcellisation, nous obtenons N^2 coefficients.

Il existe deux grandes classes d'outils mathématiques pour contenir cette information : les matrices et les graphes. Dans cette section, nous décrirons ces outils mathématiques. Nous ne différencierons pas connectivité anatomique et connectivité fonctionnelle. En effet, ces outils sont employés dans les deux catégories de connectivité. De plus, ils possèdent les mêmes propriétés (voir par exemple [PARK et coll. 2008, SKUDLARSKI et coll. 2008, HONEY et coll. 2009], [ZALESKY & FORNITO 2009]).

Matrices d'association et d'adjacence

Les N régions de la parcellisation correspondent aux N colonnes et aux N lignes d'une matrice $N \times N$, notée \mathbf{A} . L'entrée a_{ij} de cette matrice correspond à la mesure de connectivité entre les régions i et j . Cette matrice, nommée matrice d'association, peut être ou ne pas être symétrique selon la directionnalité de la connectivité. On peut remarquer que certaines études [ROBINSON et coll. 2010, VAROQUAUX et coll. 2010] emploient uniquement ces matrices d'association dans leurs traitements.

Il peut être intéressant de calculer la force moyenne d'association (*mean strength of association*) pour une région i :

$$a_{i,\cdot} = \frac{1}{N} \sum_{j=1}^N a_{i,j}. \quad (5.6)$$

Il peut être aussi intéressant d'avoir un coefficient décrivant des propriétés de dispersion plus globales [LYNALL et coll. 2010] :

$$I = \frac{\lambda_1}{\sum_{i=2}^N \lambda_i}, \quad (5.7)$$

avec λ_i la i ème valeur propre de la matrice d'association.

On peut aussi appliquer un seuil τ sur ces matrices d'association pour obtenir des matrices binaires, dites matrices d'adjacence [BULLMORE & BASSETT 2011]. La valeur 1 correspond à l'existence d'un lien entre deux régions, et la valeur 0 à l'absence de lien. On peut s'interroger alors sur le choix de τ . Il existe différentes manières de procéder : soit choisir un seuil fixe (par exemple si on sait que les connexions sont significatives uniquement au dessus d'une certaine valeur) [ACHARD et coll. 2006, HE et coll. 2007], soit déterminer ce seuil pour obtenir par exemple un nombre fixé de connexions non nulles [ACHARD & BULLMORE 2007, BASSETT et coll. 2008]. A la suite de ce seuillage, on obtient la matrice binaire dite d'adjacence.

Graphes

On peut associer à une matrice d'adjacence de taille $N \times N$ un graphe G composé d'un ensemble V de nœuds (ou sommets) et d'un ensemble E d'arêtes. Plus précisément, il existe une arête (v_i, v_j) entre les nœuds v_i et v_j si le coefficient $a_{i,j}$ de la matrice d'adjacence est égal à 1.

Nous décrirons par la suite uniquement les graphes non pondérés (toutes les arêtes du graphe valent 1) et non directionnels. Il existe cependant une généralisation des propriétés énoncées ci-dessus pour des graphes pondérés et directionnels [RUBINOV & SPORNS 2010].

Différentes études ont analysé les propriétés topologiques et géométriques de ces graphes ([ACHARD et coll. 2006, ACHARD & BULLMORE 2007] pour la connectivité fonctionnelle, [HAGMANN et coll. 2008] pour la connectivité anatomique, et [SKUDLARSKI et coll. 2008] [PARK et coll. 2008, HONEY et coll. 2009, ZALESKY & FORNITO 2009] pour des études couplées anatomo-fonctionnelles). Des articles de synthèse ont pour objet l'étude des graphes en connectivité cérébrale (voir [BULLMORE & SPORNS 2009, RUBINOV & SPORNS 2010] [BULLMORE & BASSETT 2011]).

Dans la suite de cette partie, nous introduirons tout d'abord différents critères généraux associés aux graphes. Nous discuterons ensuite des propriétés des graphes obtenues en connectivité cérébrale.

Mesures topologiques et géométriques

Degré d'un nœud et distribution des degrés. Le degré d'un nœud est le nombre de connexions reliant ce nœud au reste du réseau. Les degrés de l'ensemble des nœuds forment une distribution caractéristique de la classe de graphe employé (réseaux aléatoires, réguliers, ...). D'autres mesures peuvent être définies à partir de la notion de degré comme l'*assortativité*, quantifiant la propriété que des nœuds de même degré sont reliés entre eux.

Longueur de chemin et efficacité. La longueur $\ell_{i,j}$ de chemin correspond au nombre minimum d'arêtes traversées pour aller d'un nœud v_i à un autre nœud v_j . Les notions d'efficacité globale E_{glob} et locale E_{loc} sont préférées à la longueur de chemin. On les calcule de la manière suivante [LATORA & MARCHIORI 2001] :

$$\begin{aligned} E_{glob} &= \frac{1}{N} \sum_{i=1}^N E_{glob}(i), \\ E_{glob}(i) &= \frac{1}{N-1} \sum_{j \neq i \in G} \frac{1}{\ell_{i,j}}, \\ E_{loc} &= \frac{1}{N} \sum_{i=1}^N E_{loc}(i), \\ E_{loc}(i) &= \frac{1}{N_{G_i}(N_{G_i}-1)} \sum_{u,v:u \neq v \in G_i} \frac{1}{\ell_{u,v}}, \end{aligned} \tag{5.8}$$

où $E_{glob}(i)$ et $E_{loc}(i)$ sont respectivement l'efficacité globale et l'efficacité locale du nœud i , G_i un sous-graphe comprenant les nœuds directement reliés au nœud i , et N_{G_i} le nombre de nœuds appartenant à G_i .

D'un point de vue pratique, l'efficacité correspond à la capacité de transporter de l'information d'un nœud à un autre. Plus précisément, E_{glob} mesure la capacité de transférer parallèlement de l'information dans le réseau, alors que E_{loc} mesure la capacité de transférer localement de l'information dans un voisinage proche. Ce coefficient est similaire à une mesure d'agrégation.

Concentrateur, centralité et modularité. Un concentrateur (*hub*) est un nœud possédant un degré élevé et un haut niveau de centralité. La centralité d'un nœud est le nombre de plus courts chemins passant par ce nœud. Un module de G est un sous-ensemble de nœuds qui sont connectés de manière plus dense entre eux qu'avec le reste du réseau. Les concentrateurs jouent un rôle clef du point de vue de la modularité. Plus précisément, il existe deux catégories de concentrateurs : les concentrateurs régionaux qui sont hautement connectés aux autres nœuds d'un même module, et les concentrateurs de connexions qui relient des modules entre eux. On peut visualiser un exemple en Fig. 5.3.

Il existe de nombreux autres coefficients pour décrire les propriétés topologiques et géométriques d'un graphe [STROGATZ 2001, ALBERT & BARABÁSI 2002]. Ces paramètres ont permis en neurosciences de mieux appréhender des phénomènes globaux, et ont eu dès lors une grande importance [BASSETT & BULLMORE 2006, BASSETT & BULLMORE 2009, WANG et coll. 2009, BULLMORE & SPORNS 2009, HAGMANN et coll. 2010, BULLMORE & BASSETT 2011].

Propriétés des graphes en connectivité cérébrale. Il est connu que les graphes modélisant des phénomènes naturels se structurent généralement de manière différente des graphes aléatoires [BULLMORE & BASSETT 2011]. Le graphe modélisant le cerveau présente des propriétés spécifiques appelées *petit monde* (PM) (*small world*) (voir [HAGMANN 2005] pour une des premières études anatomiques et [ACHARD et coll. 2006] pour une des premières études

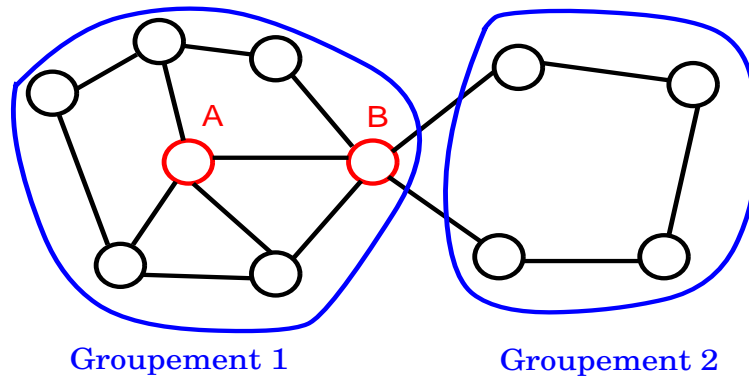


FIGURE 5.3 – Schéma représentant un graphe avec deux concentrateurs visualisés en rouge. On observe deux groupements de nœuds schématisés en bleu. On remarque que les deux concentrateurs n'ont pas le même rôle. Le concentrateur A relie les nœuds du groupement 1 (concentrateur régional), tandis que le concentrateur B relie les groupements 1 et 2 (concentrateur de connexions). Cependant les deux concentrateurs A et B possèdent un degré élevé et un haut niveau de centralité. Figure inspirée de [BULLMORE & SPORNS 2009].

fonctionnelles). La caractéristique d'un graphe PM est de présenter un haut niveau d'agrégation entre les nœuds à l'échelle locale ainsi que des courts chemins reliant l'ensemble des points du graphe (voir Fig. 5.4). L'organisation en PM est une organisation intermédiaire entre les graphes aléatoires, présentant des courts chemins entre l'ensemble de ces nœuds mais un faible niveau d'agrégation locale, et un graphe régulier, présentant un haut niveau d'agrégation mais de longs chemins pour relier deux nœuds quelconques. D'ailleurs, il n'existe pas de coefficient intrinsèque pour définir une organisation PM. Cette organisation est déterminée par rapport à une organisation aléatoire pour le même nombre de nœuds.

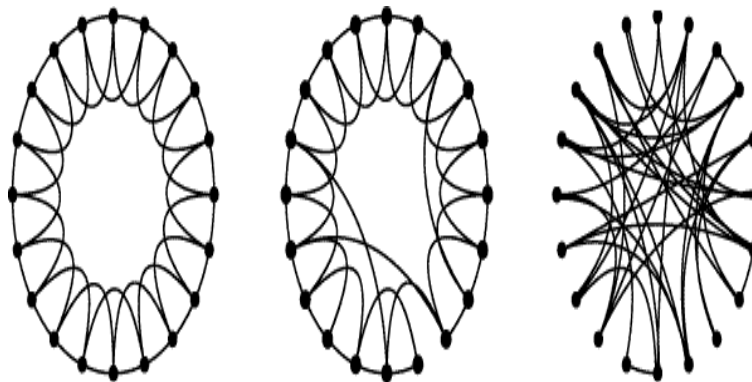


FIGURE 5.4 – Schéma représentant différentes structures de graphes. De gauche à droite : un graphe régulier, un graphe PM et un graphe aléatoire. On observe pour le graphe régulier une forte agrégation locale et une grande longueur de chemin, contrairement au graphe aléatoire qui a une faible agrégation locale mais une petite longueur de chemin. Le graphe PM possède à la fois une forte agrégation locale et une longueur de chemin petite (on observe quelques arêtes reliant des points distants, créant ainsi des raccourcis). Figure extraite de [WATTS & STROGATZ 1998].

Soit C la mesure de l'agrégation locale et L la longueur de chemin moyen pour un graphe G . On calcule les mêmes mesures C_r et L_r pour un graphe aléatoire. Pour ce faire, on génère

n graphes aléatoires ayant le même nombre de nœuds que le graphe G , et on moyenne les différentes mesures d'agrégation et de longueur de chemin obtenues à chaque réalisation. Dans [WATTS & STROGATZ 1998, HUMPHRIES & GURNEY 2008], les auteurs introduisent le scalaire σ dit « petit monde » :

$$\sigma = \frac{\frac{C}{C_r}}{\frac{L}{L_r}}. \quad (5.9)$$

On considère qu'un graphe possède des propriétés PMs si σ est supérieur à 1. En effet, on a le ratio $\gamma = \frac{L}{L_r}$ proche de la valeur 1 indiquant l'existence de chemins courts reliant l'ensemble des nœuds, et le ratio $\lambda = \frac{C}{C_r}$ supérieur à la valeur 1 indiquant une meilleure agrégation locale pour le graphe PM que pour un graphe aléatoire. Dans [BULLMORE & BASSETT 2011], les auteurs mettent en garde sur l'utilisation et l'interprétation du coefficient σ . Comme contre-exemple, ils proposent de considérer un graphe régulier possédant une bonne agrégation locale et une longueur moyenne de chemin grande (par exemple $\gamma = 3$ et $\lambda = 2$), qui possède une valeur σ supérieure à 1. Ils recommandent de présenter en plus du coefficient σ les deux ratios γ et λ .

On peut s'interroger sur la pertinence d'une structure PM pour le cerveau. D'une part, la nature PM du cerveau provient du fonctionnement même du cerveau (voir par exemple [BASSETT & BULLMORE 2006]) : il existe à la fois des zones spécialisées pour traiter l'information et à la fois une capacité de transmission rapide de l'information pour chaque couple d'aires cérébrales. Ces deux principes correspondent respectivement à un haut niveau de ségrégation et des longueurs de chemin faibles.

Actuellement, de plus en plus d'articles mettent en évidence les caractéristiques PM en connectivité anatomique [BASSETT & BULLMORE 2006, HAGMANN et coll. 2008], ainsi qu'en connectivité fonctionnelle [ACHARD et coll. 2006, ACHARD & BULLMORE 2007]. De plus, dans le cadre de pathologies (par exemple la maladie d'Alzheimer, la schizophrénie ou la sclérose en plaques), les caractéristiques PM sont altérées [LIU et coll. 2008, SUPEKAR et coll. 2008, HE et coll. 2009, BASSETT & BULLMORE 2009]. On peut donc faire de la valeur de la caractéristique PM un facteur diagnostique.

3. Conclusion partielle

L'étude de la connectivité cérébrale, et plus particulièrement celle anatomique, est en plein essor. Ces connectivités permettent d'analyser de manière globale le cerveau. Elles permettent aussi de comprendre les interactions entre mécanismes fonctionnels et anatomiques.

Du point de vue du traitement de données, la connectivité laisse beaucoup de questions ouvertes, comme par exemple le choix de la parcellisation et de la mesure de connectivité. Certaines études laissent entrevoir l'idée qu'il ne faut pas se limiter à un choix mais plutôt accumuler de l'information.

La connectivité cérébrale apporte une quantité d'information très importante, nécessitant l'utilisation d'outils spécifiques tels que les graphes. De plus, les propriétés de ces graphes ont mis en évidence des phénomènes communs à l'ensemble des catégories de connectivités, permettant une nouvelle approche dans la compréhension de pathologies.

Conclusion

Dans cette première partie, nous avons étudié les différents modèles aux échelles locale, semi-globale et globale, relatifs à l'IRMd. Ce qui ressort principalement de cette étude, c'est qu'il n'existe pas de modèle standard universel. Cependant, l'évolution des méthodes en IRMd tend vers des modèles toujours plus proches de la réalité anatomique sous-jacente.

Cette multitude de modèles possède une quantité d'information toujours plus importante, entraînant la création et l'utilisation d'outils toujours plus complexes. Dans cette première partie, on a observé qu'il n'existait pas non plus de standard universel pour ces outils mathématiques. Plus précisément, les outils dépendent du contexte dans lequel on veut les employer.

Cette multitude de modèles à des échelles différentes et d'objets mathématiques pour les représenter est une réelle opportunité. En effet, cette grande quantité d'information, pour peu qu'elle soit bien maîtrisée, permet d'étudier le cerveau d'un point de vue anatomique de diverses manières. L'idéal serait de coupler toutes ces informations et d'effectuer des études de manière conjointe.

Deuxième partie

Mise en correspondance en IRM de diffusion

Introduction

La deuxième partie de ce mémoire traite de la mise en correspondance en IRM_d. Plus précisément, une fois le choix du descripteur effectué, la conception des atlas nécessite tout d'abord de mettre en correspondance spatiale (recaler) les différents jeux de données entre eux. On doit tenir compte de la structure complexe du signal. Nous aborderons spécifiquement dans cette thèse la mise en correspondance des modèles locaux et globaux.

Lors du recalage de modèles locaux, en plus des déplacements spatiaux classiques, une étape de réorientation doit être effectuée. Les recalages et les étapes de réorientation ont été longuement étudiés dans le cadre des tenseurs d'ordre 2. Le recalage pour des tenseurs d'ordre supérieur est plus complexe. Les étapes de réorientation actuelles ne sont pas adaptées pour des déformations complexes tels le cisaillement ou l'étirement.

Dans le cadre des travaux de cette thèse, nous avons proposé un système de techniques de réorientation pour une transformation quelconque. Nous proposons principalement deux nouvelles méthodes basées sur des décompositions tensorielles (la décomposition spectrale et la décomposition *via* le théorème de Hilbert sur les quartiques ternaires) permettant de se ramener à une somme de tenseurs d'ordre 2. Ces différentes techniques ont été comparées aux deux techniques existantes sur des données simulées.

Dans un deuxième temps, nous décrirons la mise en correspondance des descripteurs globaux pour l'ensemble d'une population.

Chapitre 6

Recalage et réorientation des modèles locaux

1.	Recalage d'images et spécificité rencontrée en IRMd	65
2.	État de l'art	67
2.1	Mesure de similarité	67
2.2	Réorientation	71
3.	Réorientation des ODFs modélisées par des tenseurs d'ordre supérieur	75
3.1	Décompositions du tenseur d'ordre supérieur	77
3.2	Stratégies de réorientation	82
3.3	Expériences	83
3.4	Autre méthode de réorientation	84
3.5	Conclusion partielle	85
4.	Conclusion partielle	85

Dans ce chapitre, nous nous intéresserons au recalage des modèles locaux. Après une brève présentation des problèmes rencontrés en IRMd, nous ferons un état de l'art des méthodes de mise en correspondance dans les différents modèles d'IRMd. Finalement, nous proposerons des stratégies originales pour traiter l'un des problèmes rencontrés : la réorientation des tenseurs d'ordre supérieur.

1. Recalage d'images et spécificité rencontrée en IRMd

Le recalage d'images en IRMd présente des problèmes spécifiques liés à la complexité du signal. Nous rappellerons ici la méthodologie générale en recalage d'images [BANKMAN 2008], puis nous mettrons en évidence les spécificités liées à l'IRMd.

Tel n'importe quel algorithme de recalage, le recalage d'images en IRMd peut être vu comme un problème d'optimisation. Notons E l'espace du modèle local choisi (par exemple l'espace des tenseurs de diffusion d'ordre 2), N le nombre de voxels de l'image, $\Omega \subset \mathbb{R}^3$ le domaine de l'image, et deux images \mathbf{I}_s pour l'image à déformer, et \mathbf{I}_t pour l'image à atteindre. Ces deux images appartiennent à l'ensemble des images $\{\mathbf{I} | \mathbf{I} : \Omega \rightarrow E^N\}$.

Le problème de recalage, vu comme un problème d'optimisation, consiste à trouver la transformation optimale \hat{T} telle que \hat{T} appartenant à un ensemble \mathbf{T} de transformations admissibles minimise une fonction de coût C pour $C(\mathbf{I}_s(\hat{T}), \mathbf{I}_t)$. En d'autres termes, le problème con-

siste à trouver la transformation \hat{T} qui déforme au mieux l'image I_s sur l'image I_t au regard du critère C .

Il existe quatre paramètres à déterminer pour effectuer un recalage d'images :

- le type de transformation (par exemple rigide, affine, ou non linéaire),
- la fonction de coût C ou mesure de similarité (par exemple un critère quadratique, de corrélation ou d'information mutuelle),
- le type de données étudiées (une mesure voxel à voxel, l'utilisation d'amers (*landmarks*)...),
- la stratégie d'optimisation (par exemple une méthode de descente de gradient ou la méthode du simplexe).

Cette stratégie présentée est valable pour des recalages d'images scalaires. Dans le cadre de l'IRMd, l'image considérée est plus complexe : chaque voxel de l'image contient une fonction définie positive sur la sphère, paramétrée par un modèle (par exemple le tenseur de diffusion) ou non (les images pondérées en diffusion). Deux questions peuvent être soulevées. Premièrement : quelle mesure de similarité doit-on choisir ? Doit-on conserver l'ensemble du signal ou s'intéresser à des caractéristiques spécifiques telles que l'anisotropie ou l'orientation ? La deuxième question légitime est de savoir ensuite comment appliquer une transformation particulière à chaque modèle.

Observons deux exemples simples pour exposer les difficultés liées à cette deuxième question. Tout d'abord, on peut s'interroger sur la nécessité d'appliquer une transformation à cette fonction sphérique. Premièrement considérons un champ de tenseurs de diffusion (voir Fig. 6.1). On applique une rotation de 45° à ce champ sans appliquer la moindre transformation aux tenseurs. On constate que les tenseurs ne correspondent plus au substrat qu'ils devaient représenter. Cette expérience simple permet de montrer qu'il faut appliquer une transformation au champ de tenseur. On peut observer que l'on met cette anomalie en évidence grâce au contexte global.

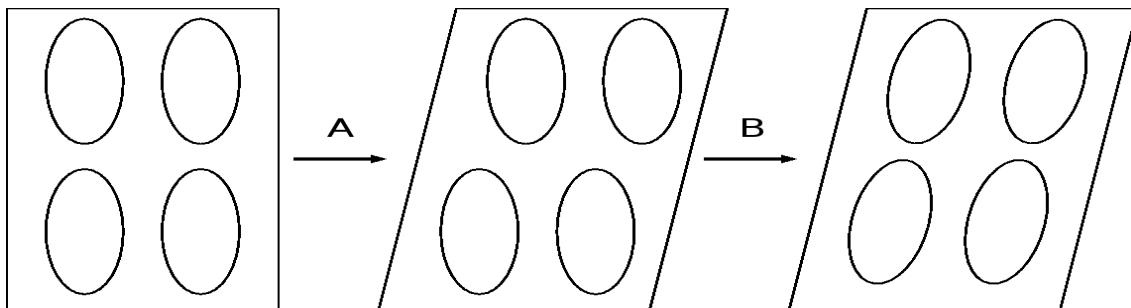


FIGURE 6.1 – A : on transforme quatre ODFs représentant des paquets de fibres verticales. On applique une rotation à ce champ d'ODFs sans appliquer aucune transformation aux ODFs. B : chaque élément est réorienté grâce à la transformation F .

La deuxième interrogation consiste à savoir quelle transformation appliquer à chaque tenseur de diffusion, et plus généralement à chaque modèle. Prenons l'exemple d'une fonction, définie sur la sphère, constante de valeur d , et appliquons lui directement un facteur d'échelle k issu d'une transformation géométrique. Nous obtenons alors une fonction définie sur la sphère, constante de valeur $d \times k$. La diffusion du milieu a donc été modifiée alors que l'eau avant et après transformation géométrique n'a pas subi de modification. Prenons maintenant une transformation modélisant un cisaillement, et appliquons la encore à une fonction constante définie sur la sphère. Nous obtenons alors une fonction définie sur la sphère, anisotrope, modélisant un milieu différent de l'eau. Nous venons encore de manière arbitraire de transformer un milieu aqueux en milieu non aqueux.

Ces expériences naïves ont montré une des difficultés du recalage en IRMd : il faut à la fois modifier le contenu du voxel en prenant en compte le contexte global tout en conservant la nature locale de ce contenu (*i.e.* préserver l'anatomie sous-jacente). Cette étape additionnelle au recalage d'image scalaire se nomme la réorientation. Cette étape provient du fait qu'il existe une notion d'orientation en IRMd, inexistante dans les autres modalités.

2. État de l'art

Dans un premier temps, nous aborderons la question du choix du critère de similarité en IRMd. Nous nous intéresserons par la suite aux méthodes de réorientation nécessaires dans le cadre de l'IRMd. Cet état de l'art n'est pas exhaustif, et a pour but de montrer au lecteur les spécificités du recalage pour différents modèles locaux en IRMd.

2.1 Mesure de similarité

Le recalage est à la base de nombreux traitements et études, par exemple des études de groupe. Dès lors, les algorithmes de recalage spécifiques à l'IRMd, et plus précisément au modèle historique du tenseur de diffusion, se sont fortement développés au cours de la dernière décennie. L'étude des critères de similarité est particulièrement intéressante. Dans un premier temps, nous donnerons les principales mesures de similarité pour des mesures scalaires. Nous nous intéresserons ensuite à l'évolution du recalage en imagerie du tenseur de diffusion, et plus spécifiquement aux mesures de similarité. Dans un troisième temps, nous nous intéresserons aux autres modalités.

Mesure de similarité pour des images scalaires

Le but de cette section est de donner au lecteur trois des différentes mesures de similarité qui seront par la suite utilisées et adaptées en IRMd.

Coefficient de corrélation : soient $\mathbf{I}_t(\mathbf{x})$ l'image de référence et $\mathbf{I}_s(T(\mathbf{x}))$ l'image source recalée sur $\mathbf{I}_t(\mathbf{x})$ au point \mathbf{x} . On peut considérer le coefficient de corrélation CC défini par :

$$CC = \frac{\sum_{\mathbf{x} \in \Omega} \left(\mathbf{I}_t(\mathbf{x}) - \bar{\mathbf{I}}_t(\mathbf{x}) \right) \cdot \left(\mathbf{I}_s(T(\mathbf{x})) - \bar{\mathbf{I}}_s(T(\mathbf{x})) \right)}{\sqrt{\sum_{\mathbf{x} \in \Omega} \left(\mathbf{I}_t(\mathbf{x}) - \bar{\mathbf{I}}_t(\mathbf{x}) \right)^2} \cdot \sqrt{\sum_{\mathbf{x} \in \Omega} \left(\mathbf{I}_s(T(\mathbf{x})) - \bar{\mathbf{I}}_s(T(\mathbf{x})) \right)^2}}, \quad (6.1)$$

avec $\bar{\mathbf{I}}_i$ la moyenne de l'image \mathbf{I}_i . Le CC quantifie la relation linéaire pour chaque couple d'intensités de voxels. L'utilisation de ce critère suppose qu'il existe une relation linéaire entre les intensités des deux images à recalier.

Somme des différences quadratiques : ce critère est défini par :

$$SDQ = \sum_{\mathbf{x} \in \Omega} \left(\mathbf{I}_t(\mathbf{x}) - \mathbf{I}_s(T(\mathbf{x})) \right)^2. \quad (6.2)$$

Ce coefficient suppose que les niveaux de gris des deux images sont en parfaite correspondance lorsque les images sont parfaitement recalées. De plus, ce coefficient, comme le CC , peut être grandement affecté par un petit nombre de voxels ayant des intensités très différentes entre les deux images.

Information mutuelle : soient $p(i_j)$ la probabilité de l'intensité i dans l'image j , et $p(i_t, i_s)$ la probabilité jointe des intensités des images \mathbf{I}_t et \mathbf{I}_s . La mesure d'information mutuelle IM est définie par :

$$IM(\mathbf{I}_t, \mathbf{I}_s) = \sum_{i_t, i_s} P(i_t, i_s) \log \left(\frac{P(i_t, i_s)}{P(i_t)P(i_s)} \right). \quad (6.3)$$

Le critère IM n'émet pas d'hypothèse sur la relation liant les intensités i_s et i_t entre les images. Ce critère est utile si on veut recalibrer des images de modalités différentes.

Mesure de similarité en imagerie du tenseur de diffusion

Outre les stratégies de réorientation, le recalage en imagerie du tenseur de diffusion connaît une autre difficulté importante [MÚÑOZ-MORENO et coll. 2009] : trouver une mesure de similarité adaptée aux tenseurs.

Pour ce faire, de nombreuses stratégies ont été développées. On peut diviser les méthodes en deux : les stratégies issues de mesures scalaires et celles prenant en compte l'ensemble du tenseur.

Tout d'abord, pour les mesures scalaires, on peut estimer la transformation à partir d'une image morphologique T_1 [XU et coll. 2003] ou à partir d'une image scalaire décrivant les propriétés de diffusion telles que la FA [JONES et coll. 2002].

D'autres études ont cherché à incorporer plus d'informations qu'une seule image scalaire, d'autant plus que cette mesure scalaire ne reflète souvent pas la notion d'orientation, importante en IRMd (par exemple la FA ne renseigne que sur l'anisotropie). Pour cela, ces études se sont basées, par exemple, sur des recalages dits multi-canaux (*i.e.* des recalages où la transformation est estimée à partir de plusieurs images). Dans [GUIMOND et coll. 2002, PARK et coll. 2003, ROHDE et coll. 2004], les différents auteurs ont utilisé une différence quadratique de différentes images scalaires : les six composantes du tenseur en rajoutant des données supplémentaires telles que l'image anatomique T_2 , la FA, la trace du tenseur... Ces différents articles ont montré que l'introduction d'information supplémentaire améliore le recalage. Dans [LEEMANS et coll. 2005], les auteurs ont considéré un critère d'information mutuelle sur l'ensemble des images pondérées en diffusion. Dans [ROHDE et coll. 2003, ROHDE et coll. 2004], les auteurs emploient une mesure d'information mutuelle multivariée sur les six composantes du tenseur et la valeur de référence S_0 , ainsi que sur la trace et l'anisotropie.

Dans [STUDHOLME 2007], l'auteur utilise l'information provenant du tenseur de diffusion pour recalibrer des images T_1 . L'idée sous-jacente est que l'imagerie du tenseur de diffusion permet de mieux décrire la matière blanche que la modalité T_1 . Pour cela, il emploie un critère d'IM multivariée sur les six composantes du tenseur en plus de l'image T_1 .

Dans [PARK et coll. 2003], les auteurs observent que le recalage sur les six composantes du tenseur, celui sur la trace des tenseurs et la FA, et celui sur les trois valeurs propres associées aux tenseurs sont relativement semblables. Ils suggèrent que l'information contenue dans le tenseur est redondante, et qu'elle n'est pas forcément utile pour améliorer le recalage, lorsque les déformations étudiées ont une faible résolution spatiale. Ils notent que l'utilisation de descripteurs plus spécifiques dans le critère de similarité telle la prise en compte de l'orientation serait plus pertinente pour des déformations plus importantes, par exemple entre deux individus.

Certaines mesures de similarité étudient des critères géométriques spécifiques aux tenseurs d'ordre 2. Premièrement, les auteurs de [ALEXANDER et coll. 1999, ALEXANDER & GEE 2000, CURRAN & ALEXANDER 2005] s'intéressent aux mesures prenant en compte la direction principale \mathbf{e}_1 . Pour ne considérer que les tenseurs « d'intérêt » (*i.e.* les tenseurs anisotropes), ils

pondèrent cette mesure par des coefficients $v(\mathbf{D})$ reflétant l'anisotropie du tenseur \mathbf{D} . Ces mesures sur la différence d'orientation (DO) entre deux tenseurs \mathbf{D}_1 et \mathbf{D}_2 peuvent s'écrire de la sorte :

$$DO(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{v(\mathbf{D}_1)v(\mathbf{D}_2)} \langle \mathbf{e}_1^{\mathbf{D}_1} | \mathbf{e}_1^{\mathbf{D}_2} \rangle, \quad (6.4)$$

avec $\mathbf{e}_1^{\mathbf{D}_i}$ la direction principale du tenseur \mathbf{D}_i , et $\langle . | . \rangle$ le produit scalaire entre deux vecteurs. Des critères similaires sont développés, rajoutant d'autres mesures. Par exemple, dans [DAVOODI-BOJD & SOLTANIAN-ZADEH 2008], les auteurs prennent en compte, en plus de la DO, la différence quadratique entre les images de FA. Dans [DUDA et coll. 2003], les auteurs proposent d'étudier simultanément la DO sur des régions ainsi que les valeurs propres des tenseurs.

Deuxièmement, des mesures basées sur la forme du tenseur ont vu le jour. En effet, le tenseur d'ordre 2 possède des formes caractéristiques telles qu'une forme en « assiette » ou en « cigare » (respectivement *plate* et *cigar shape*). Ces différentes formes peuvent être caractérisées par des coefficients scalaires [WESTIN et coll. 2002]. Les mesures de similarité basées sur les formes sont des combinaisons de ces coefficients [POLLARI et coll. 2006, POLLARI et coll. 2007, MÚÑOZ-MORENO & MARTIN-FERNANDEZ 2009].

A notre connaissance, les premiers travaux à considérer l'ensemble du tenseur, sans le considérer composante par composante, sont ceux d'Alexander et coll. [ALEXANDER et coll. 1999]. Dans cet article, les auteurs proposent plusieurs critères de similarité basés sur le tenseur tels que la métrique euclidienne canonique ou le produit scalaire. Les auteurs montrent que la métrique euclidienne se trouve être l'une des meilleures mesures. Dès lors, de nombreux articles utilisent cette métrique comme mesure de similarité [ALEXANDER & GEE 2000] [GEE et coll. 2002, RUIZ-ALZOLA et coll. 2002, CAO et coll. 2005, BARBIERI et coll. 2008]. Dans [RUIZ-ALZOLA et coll. 2002], les auteurs étendent les critères de similarité scalaire tel que le *CC* au cas tensoriel. Dans [MÚÑOZ-MORENO et coll. 2009], les auteurs mentionnent le fait que si on utilise le tenseur dans le critère de similarité, il faut prendre en compte l'étape de réorientation. Plusieurs articles [CAO et coll. 2005, ZHANG et coll. 2006, YEO et coll. 2008] prennent en compte cette étape, et proposent des calculs de gradient de leurs mesures pour améliorer l'optimisation.

En raison de la difficulté d'implémentation et de validation d'une méthode de recalage, les articles comparant différentes stratégies sont peu nombreux. Intéressons-nous à l'article [WANG et coll. 2011] proposant une comparaison entre six stratégies scalaires et une stratégie tensorielle. Les auteurs constatent que la méthodologie basée sur le tenseur semble donner de meilleurs résultats selon leurs critères de comparaison.

On peut aussi citer l'article [ZHANG et coll. 2007] qui compare des études statistiques réalisées à partir de recalage scalaire basé sur la FA et sur le tenseur. On peut constater dans l'article une nette amélioration des résultats grâce à l'utilisation du recalage basé sur les tenseurs.

Des critères de similarité basés sur la géométrie des fibres nerveuses ont aussi été développés. Par exemple, on peut citer la technique bien connue du *tract based spatial statistics* (TBSS) [SMITH et coll. 2006] qui extrait le squelette de l'image de FA au sens de la morphologie mathématique. Ce squelette correspond aux paquets principaux de fibres de substance blanche. Le TBSS met en correspondance les squelettes des différentes images, assurant ainsi une bonne mise en correspondance entre les paquets importants de fibres de substance blanche.

On peut aussi s'intéresser à l'article [LEEMANS et coll. 2006] où les auteurs vont extraire des coefficients spécifiques aux structures filaires : la courbure et la torsion. La mesure de similarité est basée sur la minimisation de la différence quadratique entre ces deux termes pour deux

paquets de fibres donnés.

En imagerie du tenseur de diffusion, l'utilisation de l'ensemble de l'information contenue dans le tenseur semble conduire à de meilleurs résultats. Cependant prendre en compte le fait que le tenseur n'est pas invariant par rotation, contrairement aux mesures scalaires telles que la FA ou la MD, rend les algorithmes plus complexes et plus coûteux en temps de calcul.

Mesure de similarité dans les autres modalités d'IRMd

Appliquer une transformation à des modèles complexes tel que le tenseur d'ordre 4 reste relativement difficile. En conséquence, le développement de méthodologie pour les autres modalités a connu un essor moindre que pour l'imagerie du tenseur de diffusion.

Un des premiers articles [TAO & MILLER 2006] a traité des images pondérées en diffusion. L'idée est ingénieuse car si on résout le problème du recalage pour les images « brutes », alors on est capable d'avoir des images recalées pour n'importe quelle modalité. La mesure de similarité correspond à une minimisation de l'erreur quadratique entre deux images pondérées en diffusion. L'étape de réorientation est incluse dans le critère de similarité.

Comme pour l'imagerie du tenseur diffusion, les critères de similarité pour les modalités plus complexes peuvent se classer en deux catégories : les critères basés sur des images scalaires représentant des propriétés du modèle comme l'anisotropie, et les critères basés sur l'ensemble du modèle.

Dans [BLOY & VERMA 2010], on estime la transformation à partir de coefficients invariants par rotation provenant d'un modèle d'ODF paramétré par des HSs. L'idée est d'employer plusieurs cartes scalaires invariantes par rotation, d'estimer une transformation pour chacune de ces cartes, et de proposer ensuite une moyenne de ces transformations comme transformation finale.

L'article [YAP et coll. 2011] utilise aussi des coefficients intermédiaires pour représenter une image d'ODFs. Les auteurs introduisent des paramètres locaux reflétant l'anisotropie, mais aussi des paramètres régionaux reflétant l'anisotropie ainsi que l'architecture de l'anatomie par le biais d'un détecteur de contours. De plus, les auteurs proposent une mise en correspondance hiérarchique en estimant successivement des ordres de HSs de plus en plus grands (de 2 à 6). Les auteurs proposent finalement d'estimer la transformation uniquement sur des voxels d'intérêt, c'est-à-dire des voxels possédant une forte anisotropie ainsi qu'un contour marqué.

Dans les critères de similarité utilisant l'ensemble du modèle, on peut notamment citer l'article [VERAART et coll. 2010] qui propose une estimation multi-canaux avec un critère IM multivariée pour un modèle DKI.

Dans le cadre des tenseurs d'ordre supérieur, on peut citer les travaux de Barmoutis et coll. [BARMPOUTIS et coll. 2007, BARMPOUTIS et coll. 2009] calculant une somme de distances entre deux champs de tenseur. Les distances utilisées sont spécifiques aux tenseurs d'ordre 4 (distance de Hellinger pour [BARMPOUTIS et coll. 2007] et distance appelée distance de Tarantola dans ce mémoire [BARMPOUTIS et coll. 2009], voir la section 1.4 pour une description de ces métriques).

Une distance euclidienne est utilisée pour un mélange de gaussiennes dans [CHENG et coll. 2009] et pour des HSs dans [GENG et coll. 2009, RAFFELT et coll. 2011]. Pour ce dernier, une méthode basée sur le CC est aussi employée.

Pour les méthodes des articles [BARMPOUTIS et coll. 2007, BARMPOUTIS et coll. 2009] [CHENG et coll. 2009, RAFFELT et coll. 2011], la réorientation est prise en compte dans le critère

de similarité.

Parmi les articles présentés précédemment, nombreux [BARMPOUTIS et coll. 2007] [BARMPOUTIS et coll. 2009, CHENG et coll. 2009, VERAART et coll. 2010, YAP et coll. 2011] [RAFFELT et coll. 2011] sont ceux qui proposent une comparaison avec des recalages basés sur des cartes scalaires représentant l'anisotropie ainsi que le modèle du tenseur de diffusion. Les études montrent que l'utilisation de modèles plus complexes permettent une meilleure mise en correspondance.

Dans [RAFFELT et coll. 2011, YAP et coll. 2011], les auteurs montrent que, pour un même modèle (ici les HSs), l'utilisation d'ordre plus élevé améliore la mise en correspondance.

2.2 Réorientation

La réorientation est une étape essentielle pour appliquer une transformation à un modèle non scalaire en IRMd. Nous nous intéresserons tout d'abord aux travaux effectués en imagerie du tenseur de diffusion. Ces travaux ont permis de comprendre les principes de base de l'étape de réorientation. Nous verrons dans un deuxième temps comment ces principes ont été appliqués dans des cas plus complexes.

Réorientation en imagerie du tenseur de diffusion

Tout d'abord, on considère la transformation T telle que $\mathbf{x}' = T(\mathbf{x})$. On peut alors utiliser une matrice de transformation affine \mathbf{F} pour exprimer localement T . Si la transformation T est affine ou rigide, alors la matrice \mathbf{F} ne dépend pas de \mathbf{x} . Si la transformation T est non linéaire, on peut calculer une matrice \mathbf{F} , en chaque point \mathbf{x} , telle que :

$$\mathbf{F} = \mathbf{I} + \mathbf{J}_{\mathbf{u}}, \quad (6.5)$$

où \mathbf{I} est la matrice identité et $\mathbf{J}_{\mathbf{u}}$ est la matrice jacobienne du champ vectoriel \mathbf{u} au point \mathbf{x} [ALEXANDER et coll. 2001]. Dans le reste du mémoire, nous utiliserons toujours la matrice \mathbf{F} sans mentionner si la transformation T est linéaire ou non.

Dans l'article de référence sur l'étape de réorientation [ALEXANDER et coll. 2001], les auteurs proposent trois types de réorientation.

La première stratégie consiste à ne pas appliquer de réorientation sous hypothèse que la déformation appliquée au champ tensoriel est faible.

Dans la section 1., nous avons observé que la transformation appliquée au tenseur ne devait pas déformer ce dernier. Pour cela, une seule classe de transformation est admissible : les rotations. Les deux autres stratégies proposent de déterminer la rotation à appliquer au champ tensoriel.

La deuxième méthode, appelée *finite strain* (FS), est inspirée de la mécanique des milieux continus pour les déformations. Elle a pour but d'extraire d'une transformation affine \mathbf{F} une rotation \mathbf{R} .

Pour déterminer la matrice de rotation \mathbf{R} , on utilise la décomposition polaire matricielle de \mathbf{F} :

$$\begin{aligned} O_n(\mathbb{R}) \times S_n^{++}(\mathbb{R}) &\rightarrow GL_n(\mathbb{R}) \\ (\mathbf{R}, \mathbf{U}) &\mapsto \mathbf{R}\mathbf{U}, \end{aligned} \quad (6.6)$$

avec $O_n(\mathbb{R})$ l'ensemble des matrices orthogonales, et \mathbf{U} appartenant à l'ensemble $S_n^{++}(\mathbb{R})$ des matrices symétriques strictement positives. Cette décomposition est unique. On a les relations

suivantes :

$$\begin{aligned} \mathbf{F} &= \mathbf{R}\mathbf{U} \\ &= \mathbf{L}\mathbf{M}\mathbf{N}^T \\ &= (\mathbf{L}\mathbf{N}^T)(\mathbf{N}\mathbf{M}\mathbf{N}^T), \end{aligned} \quad (6.7)$$

avec $\mathbf{L}\mathbf{M}\mathbf{N}^T$ la décomposition en valeurs singulières de \mathbf{F} , $\mathbf{L}\mathbf{N}^T$ une matrice orthogonale et $\mathbf{N}\mathbf{M}\mathbf{N}^T$ une matrice symétrique définie positive. Par identification, on conclut que $\mathbf{R} = \mathbf{L}\mathbf{N}^T$.

Cette réorientation ne dépend pas du tenseur considéré. Ceci peut engendrer des erreurs dans l'orientation lors de certaines transformations tel que le cisaillement. Prenons un exemple pour mieux appréhender ce problème. Considérons un tenseur possédant une seule direction privilégiée et une transformation consistant en un cisaillement dans cette même direction. La rotation \mathbf{R} extraite de cette transformation sera différente de la matrice identité, et par conséquent la direction privilégiée du tenseur sera modifiée. Cependant, appliquer un cisaillement dans la direction du tenseur ne doit pas modifier cette direction.

Pour réussir à dépasser cette limitation, Alexander et coll. dans [ALEXANDER et coll. 2001] proposent une autre manière d'estimer cette matrice de rotation en prenant en compte l'ensemble de la transformation tout en conservant la géométrie du tenseur. Cette technique est appelée la méthode de *preservation of principal direction* (PPD). L'idée est d'appliquer l'ensemble de la transformation au tenseur, puis de normaliser le tenseur pour qu'il conserve sa forme initiale.

Cette stratégie se décompose en trois étapes. La première consiste à appliquer la transformation à la direction principale \mathbf{e}_1 (*i.e.* le vecteur propre dont la valeur propre est la plus grande), puis à normaliser le résultat :

$$\mathbf{n}_1 = \frac{\mathbf{F} \mathbf{e}_1}{\|\mathbf{F} \mathbf{e}_1\|}, \quad (6.8)$$

où \mathbf{n}_1 est le nouveau vecteur principal réorienté.

La deuxième étape consiste à appliquer la transformation au second vecteur propre \mathbf{e}_2 , tout en prenant en compte que le nouvel axe doit être orthogonal à \mathbf{n}_1 :

$$\mathbf{n}_2 = \frac{\mathbf{F} \mathbf{e}_2 - (\mathbf{n}_1^T \mathbf{F} \mathbf{e}_2) \mathbf{n}_1}{\|\mathbf{F} \mathbf{e}_2 - (\mathbf{n}_1^T \mathbf{F} \mathbf{e}_2) \mathbf{n}_1\|}, \quad (6.9)$$

où \mathbf{n}_2 est le nouveau vecteur principal réorienté.

Pour obtenir le dernier vecteur propre \mathbf{n}_3 , il suffit de considérer que les trois vecteurs \mathbf{n}_1 , \mathbf{n}_2 et \mathbf{n}_3 forment une base orthogonale directe :

$$\mathbf{n}_3 = \mathbf{n}_1 \wedge \mathbf{n}_2. \quad (6.10)$$

Cette méthodologie constitue le standard universel pour les techniques de réorientation. Cependant, en imagerie du tenseur de diffusion, il existe d'autres méthodes. Citons notamment les travaux de [XU et coll. 2003]. Les auteurs considèrent un voisinage local pour estimer la direction principale et ensuite la réorientation.

Dans l'article [BARBIERI et coll. 2008], l'effet de cisaillement est estimé séparément de la matrice de rotation. Les effets de cisaillement nécessitent cependant d'être faibles.

Les différentes stratégies de réorientation de l'imagerie du tenseur de diffusion proposées permettent d'appliquer correctement une transformation. Néanmoins ces stratégies, basées sur les directions principales, connaissent une limite principale. En présence d'un croisement de paquets de fibres, la forme du tenseur peut représenter une sphère, et la réorientation ne sera pas adaptée sachant que les directions principales ne correspondent plus à la réalité sous-jacente.

Dans [GEE & ALEXANDER 2006], les auteurs soulignent le fait que l'utilisation de la rotation, préconisée dans [ALEXANDER et coll. 2001], n'est plus adaptée pour réorienter les croisements de paquets de fibres. Pour traiter le cas du croisement de paquets de fibres, il faut alors utiliser un autre type d'imagerie que le DTI avec une technique de réorientation adaptée.

Réorientation dans les autres modalités d'IRMd

Il existe plusieurs différences entre les réorientations de l'imagerie du tenseur de diffusion et les autres types d'imagerie.

Tout d'abord, à la différence de l'imagerie du tenseur de diffusion, l'obtention des directions principales dans les autres types d'imagerie proposés en IRMd est souvent complexe. Prenons maintenant un exemple pour comprendre la deuxième grande différence. Soit un tenseur représentant la distribution d'orientation des fibres lors d'un croisement orthogonal de paquets de fibres. Supposons maintenant un cisaillement dans la direction de l'une des fibres. Cette dernière direction ne doit pas être réorientée. Cependant le paquet de fibres orthogonal subira l'effet du cisaillement. Le croisement réorienté ne sera plus orthogonal. Dans le cadre d'imageries permettant de visualiser des phénomènes complexes tels que les croisements de paquets de fibres, les transformations de réorientations admises seront plus complexes que les rotations.

En IRMd, on peut séparer les méthodes de réorientation en deux catégories : la première qui propose une réorientation locale, et la deuxième qui essaie de se ramener aux différentes stratégies proposées dans [ALEXANDER et coll. 2001], et qu'on nommera réorientation globale. Comme pour l'imagerie du tenseur de diffusion, constatons aussi que ce sont les directions des paquets de fibres qui sont généralement utilisées pour estimer les réorientations. Les types d'imagerie représentant les distributions d'orientation des fibres ont été dès lors particulièrement étudiées. On notera f la fonction définie sur la sphère, à réorienter, et S_f la surface associée.

Réorientation locale. Les stratégies locales proposent de réorienter chaque direction séparément. Une des premières méthodes a été introduite par [TAO & MILLER 2006] sur les images pondérées en diffusion. Les auteurs appliquent la première étape du PPD sur chaque gradient de diffusion \mathbf{g} :

$$\mathbf{g}^{reo} = \frac{\mathbf{F}\mathbf{g}}{\|\mathbf{F}\mathbf{g}\|}. \quad (6.11)$$

L'application de la première étape du PPD permet de prendre en compte l'ensemble de la transformation (par exemple des effets de cisaillement). Cependant l'échantillonnage des gradients va être modifié. Cette technique nécessite donc une étape d'interpolation des images pondérées en diffusion [RENARD et coll. 2010].

Une idée similaire a été appliquée dans [RAFFELT et coll. 2011, YAP et coll. 2011]. Ces articles considèrent une distribution d'orientation des fibres paramétrée par des HSs. L'idée de la technique est de décomposer le signal f en une base de HSs élémentaires δ_i . Chaque élément δ_i de cette base modélise une direction spécifique (θ_i, ϕ_i) , et on lui associe un poids w_i . Ensuite, chaque direction est réorientée selon l'équation 6.11. On définit alors une nouvelle base $\{\delta_i^{reo}\}$, de directions associées $\{(\theta_i^{reo}, \phi_i^{reo})\}$. On associe les poids $\{w_i\}$ aux $\{\delta_i^{reo}\}$. Chaque élément des bases δ_i et δ_i^{reo} possède le même volume. Le fait de conserver les anciens poids $\{w_i\}$ contraint alors de conserver les volumes entre les fonctions f et f^{reo} .

On voit dès lors une nouvelle condition apparaître, autre que la conservation de la forme : la conservation du volume. Plus précisément, la distribution d'orientation des fibres peut être considérée comme une probabilité. Dès lors, l'intégrale de f sur l'ensemble de la sphère unité

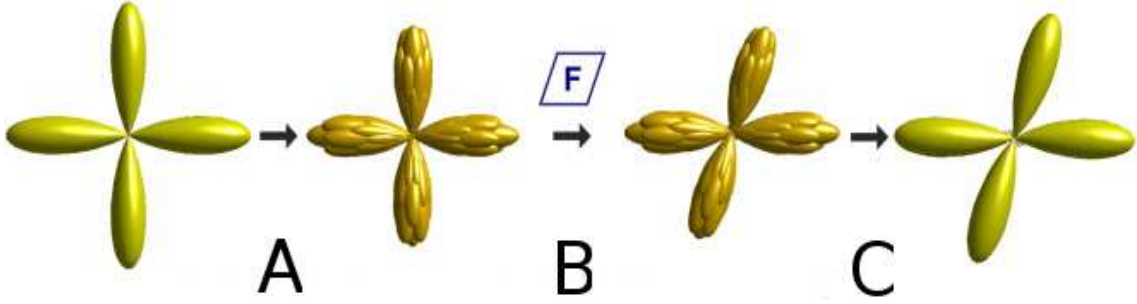


FIGURE 6.2 – A : l'ODF est décomposée en éléments de base. B : chaque élément est réorienté grâce à la transformation F . C : les éléments sont rassemblés pour reconstruire l'ODF réorientée. Schéma extrait de [RAFFELT et coll. 2011].

\mathbb{S}^2 doit être égale à 1. Cette dernière condition est moins contraignante et autorise un ensemble de transformations plus large que celui des rotations.

Cependant, cette technique peut avoir des inconvénients dans les zones isotropes avec des transformations possédant un fort effet de cisaillement. Considérons une sphère représentant un milieu isotrope. Décomposons le signal de cette sphère en choisissant des directions uniformément réparties sur la sphère. Chaque HS élémentaire associé à une direction est supposé égal à une valeur donnée. Si nous appliquons une transformation T non linéaire sur l'ensemble de ces directions, la distribution de ces directions sur cette sphère ne sera plus uniforme. Dès lors, si nous recomposons le signal, certaines HS élémentaires se seront rapprochées, et la fonction recomposée ne correspondra donc plus à une sphère.

La conservation du volume est un critère nécessaire lors de la réorientation, cependant il ne peut être considéré comme optimal.

Une autre technique conserve la valeur de cette intégrale en s'intéressant aux propriétés de l'aire d'un élément de surface élémentaire avant et après la réorientation [HONG et coll. 2009]. On veut mettre en correspondance le domaine $d\Omega$ localisé en (θ, ϕ) avec celui $d\Omega'$ localisé en (θ', ϕ') par la transformation J . Les probabilités de présence de fibres $P(\theta, \phi)$ et $P^{reo}(\theta, \phi)$ sont reliées par l'équation suivante :

$$P(\theta, \phi) d\Omega = P^{reo}(\theta', \phi') d\Omega', \quad (6.12)$$

avec $d\Omega = |\sin(\theta)|d\theta d\phi$ et $d\Omega' = |\sin(\theta')|d\theta' d\phi'$ les surfaces élémentaires avant et après réorientation. Le lien entre $d\Omega$ et $d\Omega'$ peut s'obtenir de la manière suivante :

$$d\theta' d\phi' = |\det(J_\Omega)|d\theta d\phi, \quad (6.13)$$

avec $\det(J_\Omega)$ le déterminant de la matrice suivante :

$$J_\Omega = \begin{pmatrix} \frac{\partial \theta'}{\partial \theta} & \frac{\partial \theta'}{\partial \phi} \\ \frac{\partial \phi'}{\partial \theta} & \frac{\partial \phi'}{\partial \phi} \end{pmatrix}. \quad (6.14)$$

Nous obtenons finalement la formule suivante pour réorienter localement la probabilité de présence des fibres :

$$P^{reo}(\theta', \phi') = P(\theta, \phi) \left| \frac{\sin(\theta)}{\sin(\theta')} \right| \frac{1}{|\det(J_\Omega)|}. \quad (6.15)$$

La relation 6.12 assure que la valeur de l'intégrale est conservée après la réorientation. Ces différentes techniques n'ont pas *a priori* sur les directions privilégiées du modèle.

Réorientation globale. Se ramener à la stratégie du PPD est aisé lorsqu'on est en présence de tenseurs d'ordre 2, comme dans le cas du mélange de gaussiennes [CHENG et coll. 2009]. Cependant la stratégie du PPD n'est pas applicable telle quelle aux imageries utilisant les tenseurs d'ordres supérieurs comme les HSs. En effet, le PPD sous-entend qu'on ait extrait les directions principales correspondant aux directions des paquets de fibres et qu'on modifie chacune de ces directions séparément.

Cependant les modèles de tenseur et de HSs ne permettent pas d'obtenir de manière simple les directions principales et de modifier les structures séparément. Dès lors, les stratégies utilisées sont basées sur le FS. Cependant, appliquer une réorientation à un tenseur d'ordre supérieur ou à des HSs reste complexe. Dans [GENG et coll. 2009, BLOY & VERMA 2010], les auteurs ont recours à des matrices de Wigner pour appliquer la réorientation.

Dans le cadre des tenseurs, Barmpoutis et coll. dans [BARMPOUTIS et coll. 2009] proposent d'appliquer la rotation \mathbf{R} estimée dans le cadre du FS à l'ensemble des vecteurs \mathbf{g} de la manière suivante :

$$D^{reo}(\mathbf{g}) = \sum_{i,j,k,l}^3 D_{i,j,k,l} (\mathbf{R} \mathbf{g})_i (\mathbf{R} \mathbf{g})_j (\mathbf{R} \mathbf{g})_k (\mathbf{R} \mathbf{g})_l, \quad (6.16)$$

avec $(\mathbf{R} \mathbf{g})_i$ la i ème composante du vecteur $\mathbf{R} \mathbf{g}$.

Il existe d'autres techniques de réorientation globale utilisant d'autres transformations que la rotation ou d'autres stratégies que le FS. Par exemple, nous pouvons observer celle décrite dans [BARMPOUTIS et coll. 2007] où les auteurs remplacent la matrice de rotation \mathbf{R} par la matrice affine \mathbf{A} modélisant localement la transformation T dans l'équation 6.16. Nous appellerons par la suite cette méthode la stratégie affine (SA). L'utilisation d'une matrice affine a l'avantage d'utiliser l'ensemble de l'information contenue dans la transformation T . Cependant l'intégrale sur la sphère de la nouvelle fonction obtenue n'est pas garantie d'être identique à la valeur de l'intégrale avant réorientation.

Nous pouvons aussi citer l'article [CHIANG et coll. 2007]. Les auteurs estiment la direction principale en discrétisant la surface paramétrée par l'ODF, puis en effectuant une analyse en composantes principales (ACP), et en conservant la première composante obtenue par l'ACP (*i.e.* celle qui explique au mieux la variance de l'ensemble des points). Cette direction est supposée être la direction privilégiée des fibres. Les auteurs estiment ensuite la rotation \mathbf{R} au sens de la première étape du PPD. Ils appliquent ensuite cette rotation \mathbf{R} à l'ensemble des directions. Cette stratégie est intéressante car elle prend en compte l'anatomie sous-jacente en plus de l'information de la transformation. Cependant cet algorithme ne prend pas en compte les croisements de paquets fibres car on n'estime qu'une direction principale.

Nous proposons de récapituler les différentes méthodes dans le tableau 2.2. Nous pouvons observer qu'à ce jour aucun consensus, à la manière du PPD pour le tenseur de diffusion, n'a été trouvé pour les autres modalités d'IRMd.

3. Réorientation des ODFs modélisées par des tenseurs d'ordre supérieur

Dans la section précédente, nous avons observé que les méthodes basées sur les tenseurs d'ordre supérieur, ainsi que sur les HSs, n'emploient jamais l'information de direction principale. Plus précisément, les deux types de stratégies sont soit l'application d'une rotation estimée à partir du FS, soit l'application de la première étape du PPD sur un ensemble de vecteurs préalablement échantillonnés uniformément sur la sphère.

Articles	Grandeur modélisée	Paramétrisation du modèle	Méthodes de réorientation
[ALEXANDER et coll. 2001]	ADC	tenseur d'ordre 2	FS et PPD
[XU et coll. 2003]	ADC	tenseur d'ordre 2	utilisation d'un voisinage local
[TAO & MILLER 2006]	Atténuation du signal d'IRMd	Néant (images brutes)	étape 1 du PPD sur chaque direction (L)
[RAFFELT et coll. 2011]	ODF	HS	conservation de l'intégrale sur la sphère (L)
[YAP et coll. 2011]	ODF	HS	conservation de l'intégrale sur la sphère (L)
[HONG et coll. 2009]	ODF	HS	conservation de l'intégrale sur la sphère (L)
[CHENG et coll. 2009]	ODF	mélanges de gaussiennes	PPD sur chaque gaussienne (G)
[GENG et coll. 2009]	ODF	HS	FS (G)
[BLOY & VERMA 2010]	ODF	HS	FS (G)
[BARMPOUTIS et coll. 2009]	ODF	tenseur d'ordre 4	FS (G)
[BARMPOUTIS et coll. 2007]	ADC	tenseur d'ordre 4	appliquer l'ensemble de la transformation affine (G)
[CHIANG et coll. 2007]	ODF	HS	étape 1 du PPD sur la direction principale pour estimer la rotation à appliquer (G)
[VERAART et coll. 2010]	ADC	tenseurs d'ordre 2 et d'ordre 4 (DKI)	PPD sur le tenseur d'ordre 2, application de la même rotation sur le tenseur d'ordre 4 (G)

TABLE 6.1 – Récapitulatif des différentes méthodologies de réorientation en IRMd. (L) et (G) dans la dernière colonne signifient que les méthodes sont locales et globales respectivement.

Dans le cadre de cette thèse, nous proposons deux méthodes originales de réorientation pour les ODFs paramétrées par des tenseurs d'ordre supérieur. Nous émettons l'hypothèse que lors d'un croisement de paquets de fibres, chaque paquet peut évoluer de manière indépendante de l'autre. Nous intéresserons spécifiquement aux tenseurs d'ordre 4 car ils permettent de modéliser jusqu'à trois paquets de fibres non colinéaires entre eux, ce qui est suffisant pour représenter les croisements de faisceaux de fibres de substance blanche dans le cerveau.

L'idée principale des deux stratégies est d'utiliser des décompositions de tenseurs d'ordre supérieur en tenseurs d'ordre 2, et d'appliquer ensuite le PPD à chaque tenseur. Dans un premier temps, nous étudierons les deux décompositions ainsi que leurs propriétés. Dans un deuxième temps, nous décrirons les méthodes de réorientation. Finalement, nous proposerons différents tests avec les méthodes existantes pour les tenseurs d'ordre supérieur.

3.1 Décompositions du tenseur d'ordre supérieur

Les deux décompositions étudiées sont la décomposition spectrale et la décomposition de Hilbert. Ces deux décompositions ont déjà été décrites brièvement dans le chapitre 2. Dans cette section, nous nous intéresserons plus précisément aux propriétés de ces deux décompositions, et nous nous interrogerons plus particulièrement sur la capacité de ces décompositions à refléter le substrat anatomique sous-jacent. Plus clairement, chaque décomposition permet de décrire une ODF en somme de tenseurs d'ordre 2. Le but est de savoir si ces tenseurs correspondent aux paquets de fibres.

Décomposition spectrale

Si nous considérons le tenseur \mathbf{D}_4 d'ordre 4 comme une matrice 6×6 grâce à la notation de Voigt, nous pouvons décomposer cette matrice grâce à la décomposition spectrale. On obtient alors six vecteurs propres appartenant à \mathbb{R}^6 . Chacun de ces vecteurs \mathbf{v}_i peut être vu comme un tenseur d'ordre 2, appelé parfois tenseur propre \mathbf{D}_2^i . A chacun de ces tenseurs propres est associée une valeur propre λ_i (voir éq. 2.14).

La première constatation est qu'on possède six tenseurs d'ordre 2 pour décrire un maximum de trois paquets de fibres. Dès lors, on peut s'interroger sur ce que représente chacun de ces tenseurs d'ordre 2. La question est non triviale. Basser et Pajevic [BASSER & PAJEVIC 2007] proposent une visualisation de cette base de tenseurs propres \mathbf{D}_2^i . Cependant le tenseur d'ordre 4 décrit dans l'article modélise une covariance et non une ODF. Nous pouvons aussi observer dans l'article que les tenseurs propres ne sont pas définis symétriques positifs. Néanmoins, lorsque nous recomposons le signal, nous multiplions les tenseurs propres par eux-mêmes, ce qui les rend strictement positifs.

Le cas où les tenseurs propres sont strictement positifs ou strictement négatifs n'est donc pas problématique. Cependant, lorsqu'un tenseur d'ordre 2 possède des valeurs propres positives et négatives, le tenseur propre élevé au carré représente alors un croisillon (voir Fig. 6.3).

Intéressons-nous maintenant aux valeurs propres. Nous avons vu dans le chapitre 2 que la matrice 6×6 représentant le tenseur d'ordre 4 n'était pas obligatoirement définie positive. Cette propriété entraîne que nous pouvons obtenir des valeurs propres λ_i négatives ou nulles. Rappelons-nous que le tenseur d'ordre 4 représente une surface strictement positive. Avoir des valeurs propres négatives correspond donc à ôter un volume.

Cette décomposition semble donc peu intéressante pour décomposer le tenseur d'ordre 4 en tenseurs d'ordre 2 correspondant à quelque chose de physique. Cependant, lors de simulations consistant en la décomposition spectrale de différents tenseurs d'ordre 4, nous pouvons observer qu'un grand nombre de valeurs propres sont petites, rendant l'influence du tenseur propre

correspondant faible (voir Fig. 6.3). De plus, nous pouvons remarquer que l'apparition des croisillons est peu fréquente. Nous pouvons aussi souligner le fait que les directions privilégiées des tenseurs d'ordre 4 correspondent globalement aux directions des tenseurs propres associés.

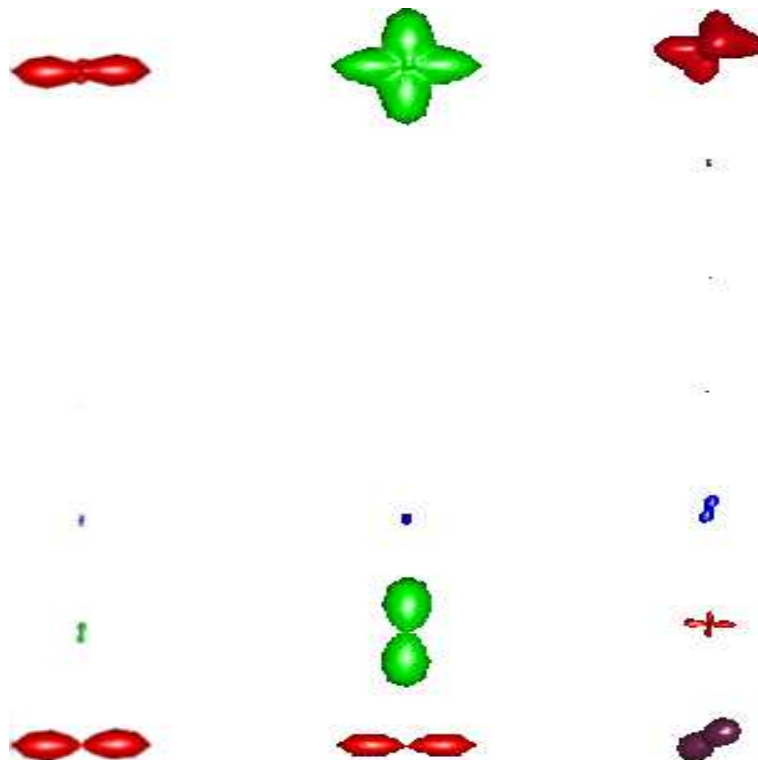


FIGURE 6.3 – La première ligne correspond au tenseur d'ordre 4. Les colonnes correspondent à la décomposition spectrale. La décomposition est représentée sur six lignes. Les tenseurs d'ordre 2 de la décomposition sont classés par valeurs propres croissantes. Tous les tenseurs d'ordre 2 sont à la même échelle. On n'observe rien sur les premières lignes car les valeurs propres sont trop petites. De gauche à droite, nous pouvons observer un paquet de fibres mono-directionnel, un croisement de paquets orthogonaux de fibres, et un croisement de paquets de fibres plus complexe. On peut observer que peu de tenseurs propres ont un poids important lors de la décomposition.

Nous allons quantifier ces trois propriétés pour différentes configurations : un paquet de fibres mono-directionnel, un croisement de paquets orthogonaux de fibres, et un croisement de paquets de fibres plus complexe. Intéressons-nous tout d'abord à l'écart entre les directions principales du tenseur d'ordre 4 calculées selon l'article [QI et coll. 2008] et les directions principales des différents tenseurs d'ordre 2. Nous ne considérons que les tenseurs d'ordre 2 ayant une importance dans la décomposition. Pour cela, nous introduisons un coefficient original K_i pour chaque tenseur propre \mathbf{D}_2^i , calculé de la manière suivante :

$$K_i = \frac{|\lambda_i|}{\sum_{j=1}^6 |\lambda_j|}, \quad (6.17)$$

où λ_i est la valeur propre associée au tenseur propre \mathbf{D}_2^i , et $|\cdot|$ est l'opérateur valeur absolue. Nous conserverons uniquement les tenseurs ayant un coefficient K_i supérieur à 20%. Pour voir si la décomposition est robuste, nous ajoutons pour chacune des trois configurations 10% de bruit, et nous effectuons 1000 tirages. Les résultats sur les erreurs des directions sont présentés dans le tableau 6.2. Nous pouvons constater que l'erreur entre les directions principales du tenseur

	Paquet de fibres mono-directionnel	Croisement de paquets orthogonaux de fibres	Croisement de paquets de fibres plus complexe
erreur moyenne en degrés	4.4	7.5	12.7
écart type	7.7	5.7	7.9

TABLE 6.2 – Récapitulatif des résultats sur les coefficients K_i et L des tenseurs d'ordre 2 obtenus par décomposition spectrale.

d'ordre 4 et celles de ses tenseurs propres restent relativement faibles, et augmentent lors des couplages plus complexes.

Étudions maintenant la présence des valeurs propres négatives. Nous avons généré 1000 réalisations de chacune des trois configurations précédentes en ajoutant 10% de bruit à chaque réalisation. Nous nous sommes intéressés à la pondération des valeurs propres négatives par rapport aux autres valeurs propres. Pour chaque valeur propre négative, nous avons donc calculé le coefficient K_i défini à l'équation 6.17. Nous représentons le résultat dans le tableau 6.3. On observe que ce coefficient K_i reste faible pour les valeurs propres négatives.

La dernière expérience portera sur l'impact des tenseurs propres possédant à la fois des valeurs propres négatives et positives. Comme pour les deux simulations précédentes, nous générerons 1000 réalisations de chacune des trois configurations en ajoutant 10% de bruit à chaque simulation. Nous introduisons un coefficient original permettant de mesurer l'écart d'un tenseur comportant des valeurs propres positives et négatives à un tenseur strictement positif ou négatif :

$$L = \min \left(\frac{\lambda^+}{\lambda_{tot}}, \frac{\lambda^-}{\lambda_{tot}} \right), \quad (6.18)$$

avec λ^+ et λ^- la somme des valeurs propres positives et celle des valeurs absolues des valeurs propres négatives respectivement, et λ_{tot} la somme des valeurs absolues des valeurs propres. Ce coefficient L est compris entre 0 et 0.5, 0 étant le cas du tenseur strictement positif ou négatif. Plus L est grand, plus le tenseur d'ordre 2 aura une forme de croisillon. Pour chaque tenseur ayant une valeur L supérieure à 0.2, nous calculons le coefficient K_i associé pour voir son importance dans la décomposition. Les résultats sont présentés dans le tableau 6.3. Nous pouvons observer qu'il existe beaucoup de tenseurs possédant à la fois des valeurs propres négatives et positives mais que les valeurs propres des tenseurs d'ordre 2 correspondants sont faibles. Par conséquent, l'influence de ces tenseurs dans la décomposition reste faible.

Bien que la décomposition spectrale semblait inadéquate pour modéliser les structures des ODFs, la simulation révèle que les phénomènes non désirés telles que la présence de valeurs propres négatives ou la présence de tenseurs d'ordre 2 possédant à la fois des valeurs propres négatives et positives, restent relativement faibles. De plus il existe de nombreux algorithmes robustes et rapides pour effectuer la diagonalisation d'une matrice.

Décomposition de Hilbert

La décomposition de Hilbert, obtenue à partir du théorème de Hilbert (voir la section 1.4 du chapitre 2), permet de décomposer le tenseur d'ordre 4 en trois tenseurs d'ordre 2. Ce nombre de tenseurs correspond au nombre maximum de directions privilégiées possibles pour un tenseur d'ordre 4. De plus, cette décomposition correspond à l'addition de trois composantes strictement positives. Dès lors cette décomposition semble plus intéressante que la décomposition spectrale.

	Paquet de fibres mono-directionnel	Croisement de paquets orthogonaux de fibres	Croisement de paquets de fibres plus complexe
K_i moyen des valeurs propres négatives	0.005	0.006	0.01
Pourcentage de tenseurs possédant à la fois des valeurs propres positives et négatives	86%	83%	83%
L moyen	0.27	0.29	0.29
K_i moyen associé au $L > 0.2$	0.025	0.05	0.06

TABLE 6.3 – Récapitulatif des résultats obtenus pour les coefficients K_i et L dans le cadre de la décomposition spectrale. On remarque dans la dernière ligne du tableau que les tenseurs de la décomposition ayant de mauvaises propriétés ($L > 0.2$) ont des valeurs propres associées faibles (*i.e.* des valeurs de K_i faible).

Cependant la décomposition de Hilbert n'impose aucune contrainte sur les tenseurs d'ordre 2. Comme pour le cas de la décomposition spectrale, nous pouvons obtenir des tenseurs strictement positifs ou strictement négatifs, ce qui n'est pas déroutant d'un point de vue conceptuel car nous nous intéressons au tenseur d'ordre 2 au carré. Le cas où on est en présence simultanément de valeurs propres négatives et positives est plus complexe. En effet, la surface alors modélisée correspond à un croisillon (problème identique à celui rencontré lors de la décomposition spectrale).

Un deuxième grand inconvénient est la non unicité de cette décomposition. Rappelons ici l'équation modélisant cette décomposition :

$$\mathbf{v}^T \mathbf{G} \mathbf{v} = \mathbf{v}^T \mathbf{Q} \mathbf{Q}^T \mathbf{v} = \mathbf{v}^T \mathbf{Q} \mathbf{R} \mathbf{R}^T \mathbf{Q}^T \mathbf{v} = \mathbf{v}^T \mathbf{Q}' \mathbf{Q}'^T \mathbf{v}, \quad (6.19)$$

où \mathbf{v} est un vecteur de monômes $[g_1^2, g_2^2, g_3^2, g_1 g_2, g_1 g_3, g_2 g_3]$, et les matrices \mathbf{Q} et \mathbf{Q}' contiennent les coefficients des tenseurs d'ordre 2. Nous pouvons aussi observer que cette décomposition dépend d'une matrice de rotation \mathbf{R} de taille 3×3 . Nous avons donc une infinité de triplets de tenseurs d'ordre deux qui correspondent au même tenseur d'ordre 4.

Nous proposons d'examiner l'influence de cette matrice \mathbf{R} sur la décomposition. Tout d'abord, nous proposons de visualiser une série de décompositions où nous faisons varier linéairement l'angle de rotation de cette matrice \mathbf{R} (voir Fig. 6.4). Nous nous rendons compte que la décomposition n'est pas unique, mais que cependant les tenseurs d'ordre 2 participant à cette décomposition semblent cohérents avec les directions privilégiées du tenseur d'ordre 4.

Nous pouvons observer dans un deuxième temps différentes décompositions pour des tenseurs d'ordre 4 aléatoires.

Comme dans le cadre de la décomposition spectrale, nous pouvons nous interroger sur la correspondance entre les directions principales du tenseur d'ordre 4 et celles des tenseurs d'ordre 2. Premièrement, comme pour la décomposition spectrale, nous considérons trois configurations particulières d'ODFs : le cas mono-directionnel, le croisement orthogonal, et le croisement présentant un couplage plus complexe, puis nous tirons aléatoirement 1000 matrices de rotation. Nous comparons alors les directions privilégiées de ces différentes décompositions avec celles du tenseur d'ordre 4, toujours obtenues grâce à l'article [Q1 et coll. 2008]. Nous introduisons la notion de fraction de volume $frac_{vol}^i$ du tenseur i d'ordre 2 \mathbf{D}_2^i . Ce coefficient original est défini

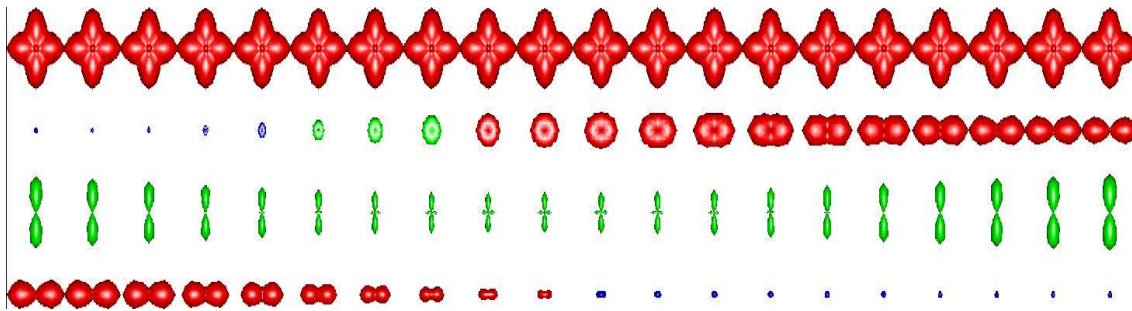


FIGURE 6.4 – Décomposition d'un tenseur d'ordre 4 en somme de tenseurs d'ordre 2 (décomposition de Hilbert). La première ligne correspond au tenseur d'ordre 4 à décomposer. Chaque colonne correspond à une décomposition de Hilbert pour une matrice de rotation \mathbf{R} donnée. Plus précisément, les lignes 2, 3 et 4 correspondent aux tenseurs \mathbf{D}_2^i obtenus à partir de la décomposition de Hilbert. On fait évoluer de manière linéaire l'angle de rotation de la matrice \mathbf{R} . On observe que les directions principales des tenseurs d'ordre 2 correspondent bien aux directions principales du tenseur d'ordre 4 quelle que soit la rotation \mathbf{R} utilisée.

	Paquet de fibres mono-directionnel	Croisement de paquets orthogonaux de fibres	Croisement de paquets de fibres plus complexe
erreur moyenne en degrés	1.4	7.7	10.8
écart type	1.2	5.7	7.2

TABLE 6.4 – Récapitulatif des résultats sur l'erreur (en degrés) des directions principales entre le tenseur d'ordre 4 et ses tenseurs d'ordre 2 obtenus par la décomposition de Hilbert.

comme suit :

$$frac_{vol}^i = \frac{\det(\mathbf{D}_2^i \times \mathbf{D}_2^i)}{\sum_{j=1}^3 \det(\mathbf{D}_2^j \times \mathbf{D}_2^j)}, \quad (6.20)$$

avec $\det(\cdot)$ le déterminant. Considérons $\mathbf{D}_2^i = \mathbf{R}^T \mathbf{\Delta}_i \mathbf{R}$, avec \mathbf{R} une matrice de rotation et $\mathbf{\Delta}_i$ la matrice de valeurs propres de \mathbf{D}_2^i . Le calcul de $\det(\mathbf{D}_2^i \times \mathbf{D}_2^i)$ revient à calculer $\det(\mathbf{\Delta}_i^2)$. Ceci nous permet de nous affranchir des problèmes liés aux valeurs propres positives et négatives. Le déterminant $\det(\mathbf{\Delta}_i^2)$ est proportionnel au carré du volume de l'ellipsoïde modélisé par le tenseur d'ordre 2 \mathbf{D}_2^i .

Nous considérons pour le calcul des directions uniquement les tenseurs d'ordre 2 ayant une fraction de volume supérieure à 0.2. Nous calculons ensuite la moyenne et l'écart-type des erreurs sur les directions. Nous observons que l'erreur moyenne est relativement faible, et qu'elle augmente avec la complexité du modèle.

Dans un deuxième temps, nous comptons le nombre de tenseurs possédant à la fois des valeurs propres négatives et positives. Nous utilisons le coefficient L introduit dans l'équation 6.18. Pour chaque L supérieur à 0.2, nous observons la fraction de volume associée à chaque tenseur d'ordre 2. Nous pouvons observer que la proportion des tenseurs d'ordre 2 ayant des valeurs propres positives et négatives varie fortement selon les configurations. Néanmoins ces tenseurs s'éloignent peu du cas idéal où le tenseur est strictement positif ou négatif. De plus, même si on s'écarte de ce cas idéal, la contribution de ce tenseur dans la décomposition reste faible.

Nous pouvons observer que dans la multitude de décompositions possibles, certaines dé-

	Paquet de fibres mono-directionnel	Croisement de paquets orthogonaux de fibres	Croisement de paquets de fibres plus complexe
Pourcentage de tenseurs possédant à la fois des valeurs propres positives et négatives	23%	71%	31%
L moyen	0.16	0.14	0.21
$frac_{vol}$ moyen associé au $L > 0.2$	0.13	0.3	0.24

TABLE 6.5 – Récapitulatif des résultats obtenus pour certains coefficients décrivant les propriétés des tenseurs d'ordre 2 intervenant dans la décomposition de Hilbert. On observe en dernière ligne du tableau que les tenseurs d'ordre 2 ayant de mauvaises propriétés ($L > 0.2$) ont des volumes associés peu importants (*i.e.* $frac_{vol}$ associée faible).

compositions coïncident avec les propriétés recherchées, *i.e.* des tenseurs d'ordre 2 modélisant l'agencement des croisements de fibres. Contrairement à ce que l'on peut penser, la non unicité de la décomposition permet d'obtenir des cas proches des cas désirés. Plus précisément, il est possible de trouver une matrice de rotation \mathbf{R} telle que la décomposition ait de bonnes propriétés.

En théorie, sélectionner une décomposition optimale revient à résoudre un problème d'optimisation sur les trois angles de rotation paramétrant la matrice de rotation \mathbf{R} . Une fonction de coût possible serait l'erreur quadratique entre les positions des maxima du tenseur d'ordre 4 et ceux des tenseurs d'ordre 2 de la décomposition. Cependant cette optimisation est complexe. De plus, en pratique, la sélection de la décomposition doit être rapide car elle est utilisée pour un grand nombre de voxels. Nous choisissons alors de tirer aléatoirement un grand nombre de matrices de rotation, et de sélectionner celle qui conduit à la valeur de la fonction de coût la plus faible.

3.2 Stratégies de réorientation

Les stratégies de réorientation proposées ici sont une adaptation du PPD aux tenseurs d'ordre 4.

Stratégie spectrale

La stratégie spectrale (SS) utilise la décomposition spectrale. L'idée est de décomposer le tenseur d'ordre 4 selon la décomposition spectrale, de réorienter chaque tenseur d'ordre 2, puis de recomposer le signal.

Chaque tenseur propre \mathbf{D}_i^2 est réorienté selon le PPD. Cependant le PPD a été conçu pour des tenseurs strictement positifs. Pour les tenseurs strictement négatifs et les tenseurs possédant des valeurs propres à la fois positives et négatives, nous choisissons comme direction principale le tenseur ayant la plus grande valeur propre en valeur absolue. Pour la deuxième direction, nous choisissons la deuxième plus grande valeur propre en valeur absolue.

Stratégie de Hilbert

A la différence de la décomposition spectrale, la décomposition de Hilbert permet d'obtenir plusieurs décompositions d'un même tenseur d'ordre 4. La solution idéale serait d'obtenir la décomposition telle que les maxima du tenseur d'ordre 4 correspondent au maxima des tenseurs d'ordre 2 de la décomposition. Le problème reviendrait à résoudre un problème d'optimisation paramétré par les angles de rotations de la matrice \mathbf{R} . D'une part, le calcul des maxima du tenseur d'ordre 4 est complexe, et d'autre part, résoudre un problème d'optimisation complexe pour un grand nombre de voxels serait très long et non utilisable pour une méthode de réorientation.

L'idée est alors de tirer de manière aléatoire un grand nombre de matrices \mathbf{R} de rotations, de calculer les décompositions correspondantes, et de conserver uniquement celle qui possède les critères désirés, *i.e.* des tenseurs d'ordres 2 avec de faibles coefficients L . Le coefficient L est beaucoup plus rapide à estimer que les maxima du tenseur d'ordre 4. Plus précisément, on effectue une optimisation stochastique en fixant le nombre de décompositions testées.

Une fois la décomposition désirée obtenue, on applique la même réorientation que pour la stratégie spectrale.

On peut remarquer que bien que l'on ait choisi une « bonne » décomposition, on peut être encore en présence de tenseurs d'ordre 2 possédant à la fois des valeurs propres positives et négatives. Néanmoins la contribution de ces tenseurs dans la décomposition est limitée.

3.3 Expériences

Premièrement nous effectuerons des tests sur une ODF, puis sur un champ d'ODFs synthétiques. Nous comparerons nos deux stratégies aux stratégies existantes pour les tenseurs d'ordre 4, c'est-à-dire la FS et la SA.

Les différentes stratégies utilisées dans les expériences sont rappelées Fig. 6.5.

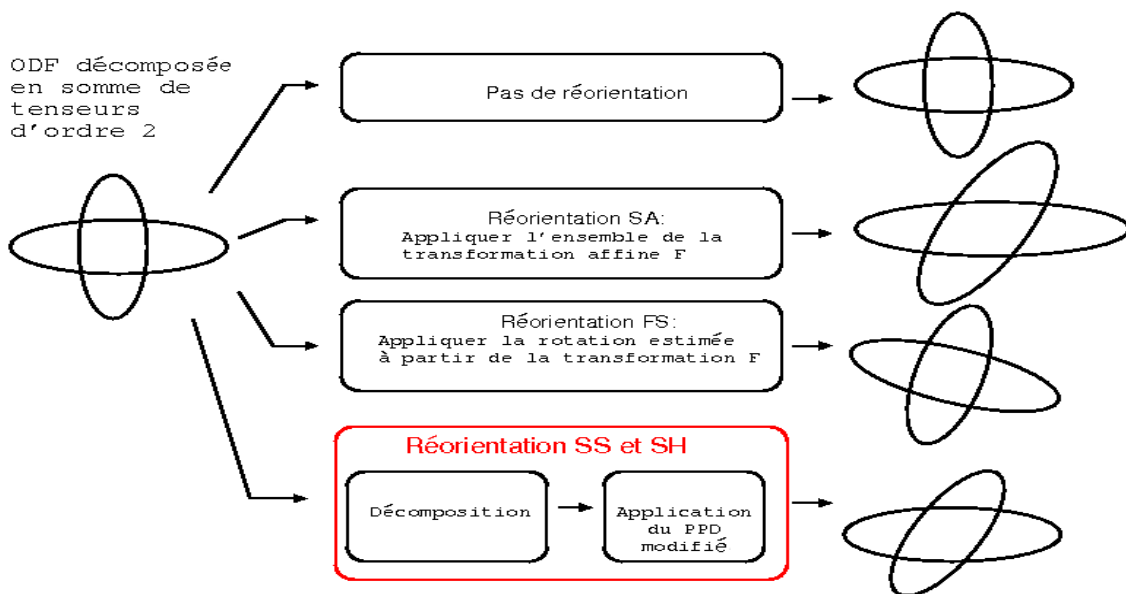


FIGURE 6.5 – Schéma récapitulant les différentes stratégies de réorientation utilisées en tenseur d'ordre 4.

	SA	FS	SS	SH
erreur moyenne (en degrés)	7.4	12.2	9	5.2
erreur maximum (en degrés)	17.1	22	20.3	13.96

TABLE 6.6 – Récapitulatif des résultats sur l’erreur (en degrés) des directions principales du tenseur d’ordre 4 réorienté.

Tests synthétiques sur un tenseur

La première simulation porte sur la stratégie de Hilbert (SH). Le but est de tester pour un même tenseur d’ordre 4 et pour une même transformation F mais avec 1000 différentes matrices de rotation \mathbf{R} , l’écart entre les directions principales du T4 réorientées et celle obtenues par la SH (voir Fig. 6.6 et 6.7). Pour obtenir les directions principales du T4 réorientées, on estime les directions principales du tenseur d’ordre 4 [QI et coll. 2008] et on applique directement la transformation sur ces directions [GRIGIS et coll. 2011]. On obtient en moyenne 4.16 degrés d’écart, et des écarts maximum de 17 degrés. Lorsque nous effectuons plusieurs tirages de la matrice de rotation pour obtenir des valeurs de L plus faibles, nous obtenons un écart de 3.66 degrés, et des écarts maximum de 12 degrés. Grâce à cette simulation, on se rend compte que la décomposition va avoir un rôle prépondérant dans la réorientation. De plus, la non unicité de la décomposition nous permet d’avoir des erreurs maximum plus faibles.

On s’intéresse maintenant à la comparaison entre les différentes méthodes SA, FS, SS et SH. On simule un croisement de paquets de fibres, puis on estime 1000 matrices F . On utilisera des transformations avec un fort effet de cisaillement. Les résultats sont présentés dans le tableau 6.6. On peut observer que le meilleur résultat correspond à la SH. On peut aussi constater que les autres stratégies ont des erreurs relativement faibles.

Tests synthétiques sur un champ de tenseurs

Nous avons simulé un croisement de paquets de fibres, et nous avons appliqué une transformation avec un fort déplacement. Nous avons comparé les différentes techniques SA, FS, SS, et SH (voir Fig. 6.8)

Nous pouvons observer que les deux méthodologies respectant le mieux l’anatomie sous-jacente correspondent aux méthodes SS et SH. La méthode FS ne permet pas de déplacer séparément les paquets de fibres. La méthode SA ne conserve pas la forme des ODFs dans le cadre mono-directionnel. De plus, en présence d’une grande déformation comme c’est le cas ici, cette stratégie va fortement déformer les ODFs.

3.4 Autre méthode de réorientation

Cette étude sur les tenseurs d’ordre supérieur a donné lieu à d’autres travaux [GRIGIS et coll. 2011]. La stratégie présentée propose de déterminer les maxima de l’ODF correspondant aux directions privilégiées des paquets de fibres. L’idée ensuite est d’effectuer une optimisation pour obtenir une somme de tenseurs d’ordre 2 orientés dans les directions privilégiées et définis positifs. Finalement on applique le PPD successivement aux différents tenseurs d’ordre 2 comme pour la SS et la SH, puis on recompose le tenseur d’ordre 4.

L’avantage de cette méthode est d’obtenir réellement des tenseurs d’ordre 2 orientés dans la direction des paquets de fibres. De plus, on contraint lors de l’optimisation que les tenseurs soient définis positifs, ce qui n’est pas le cas lors des décompositions spectrale et de Hilbert.

L'inconvénient est que la recherche de direction principale est complexe. De plus, lors de l'optimisation, les tenseurs d'ordre 2 possibles sont contraints d'avoir les deux plus petites valeurs propres égales. Cela entraîne que le tenseur reconstruit n'est pas exactement le même que celui avant décomposition.

3.5 Conclusion partielle

Dans le cadre de cette thèse, nous avons proposé deux nouvelles stratégies de réorientation pour les ODFs. Ces deux stratégies se basent sur une décomposition du signal en signaux élémentaires plus aisés à réorienter. Avec ces méthodes, nous proposons de combler un vide existant dans les stratégies existantes pour les tenseurs d'ordre supérieur, qui considéraient jusque là soit des rotations, soit des fortes déformations du volume. Le second apport est la meilleure compréhension de ces décompositions pouvant être réutilisées dans d'autres applications comme la tractographie par exemple. Ces stratégies sont basées sur le tenseur d'ordre 4, mais peuvent aussi être utilisées avec des HSs grâce au lien unissant les deux modèles. De plus, la stratégie spectrale peut être théoriquement effectuée pour des tenseurs d'ordre supérieur à 4.

4. Conclusion partielle

La mise en correspondance en IRMd pour les modèles locaux est complexe. Cette complexité s'explique notamment par le fait que les algorithmes de mise en correspondance scalaire ont dû être adaptés aux signaux multivariés. Ces algorithmes ont dû le plus souvent extraire l'information pertinente des modèles pour être de plus en plus performants.

De plus, la complexité de la mise en correspondance de signaux multivariés peut être aussi expliquée par l'introduction d'une étape intermédiaire devant être ajoutée dans la chaîne de traitement : la réorientation.

Cette partie consacrée à la mise en correspondance a mis en évidence plusieurs points. Premièrement l'information contenue dans les modalités d'IRMd contenant des signaux multivariés est plus riche que celle contenue dans des images scalaires. Cependant l'utilisation de plusieurs paramètres reflétant l'anisotropie, ainsi que la forme du modèle, permet généralement une mise en correspondance acceptable.

Deuxièmement l'information sur l'orientation des modèles est très importante. Plus généralement, certains descripteurs géométriques possèdent autant d'information que la paramétrisation canonique du modèle et sont plus facilement interprétables.

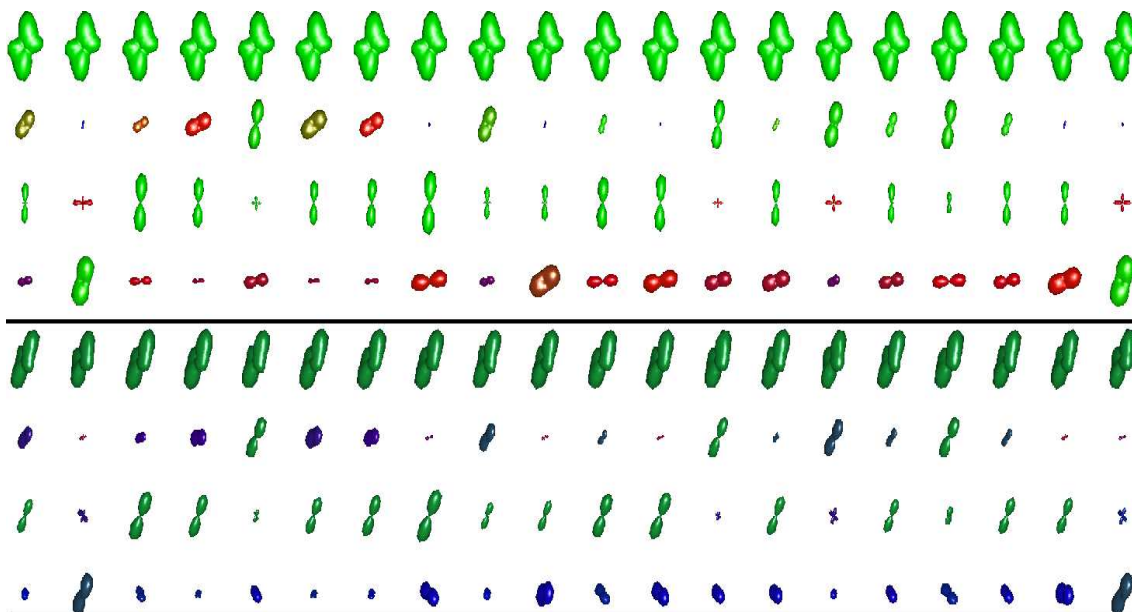


FIGURE 6.6 – La partie haute du schéma correspond à la décomposition de Hilbert avant réorientation d'un même tenseur d'ordre 4. La partie basse correspond à la décomposition après réorientation. La première ligne correspond au tenseur d'ordre 4, et les lignes suivantes représentent les tenseurs d'ordre 2 intervenant dans la décomposition. Chaque colonne correspond à une décomposition de Hilbert calculée avec une nouvelle matrice de rotation \mathbf{R} choisie aléatoirement. On peut observer que les tenseurs d'ordre 4 réorientés ont tous la même forme alors que les décompositions sont très différentes.

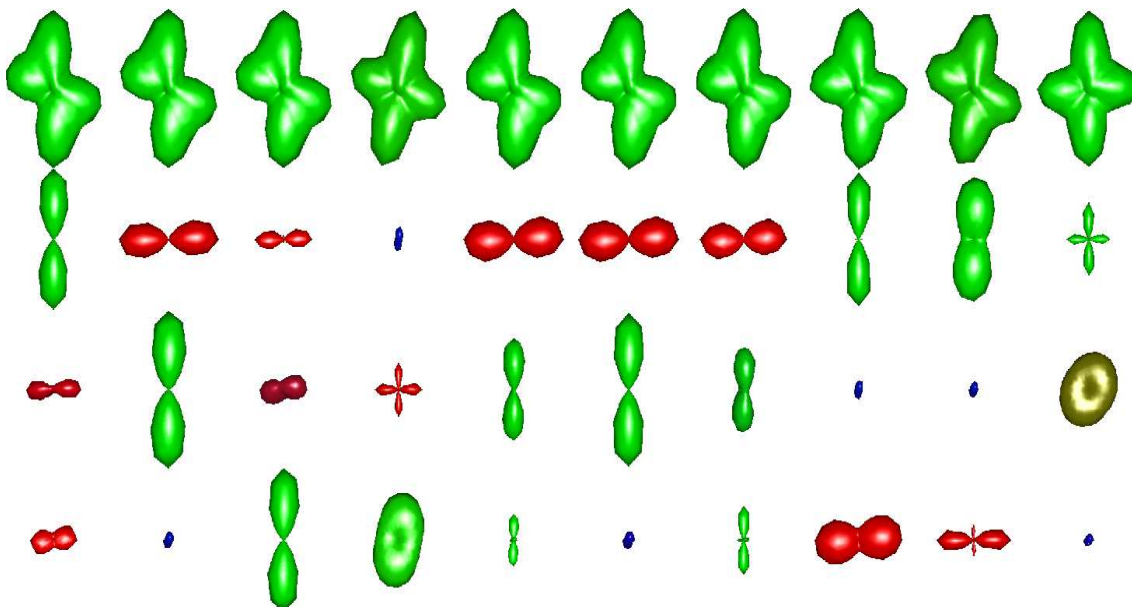


FIGURE 6.7 – On considère un tenseur d'ordre 4 (non représenté), et on observe le résultat de la réorientation avec les différentes composantes de la décomposition de Hilbert. Nous visualisons le tenseur d'ordre 4 réorienté sur la première ligne. Chaque colonne correspond à la décomposition du tenseur d'ordre 4 de la première ligne. Nous observons nettement que la présence d'un croisillon (voir ligne 1, colonnes 4 et colonne 10) dans la décomposition modifie fortement le résultat de la réorientation. En l'absence de croisillon, les résultats de la réorientation ne dépendent pas (ou dépendent très peu) de la décomposition.

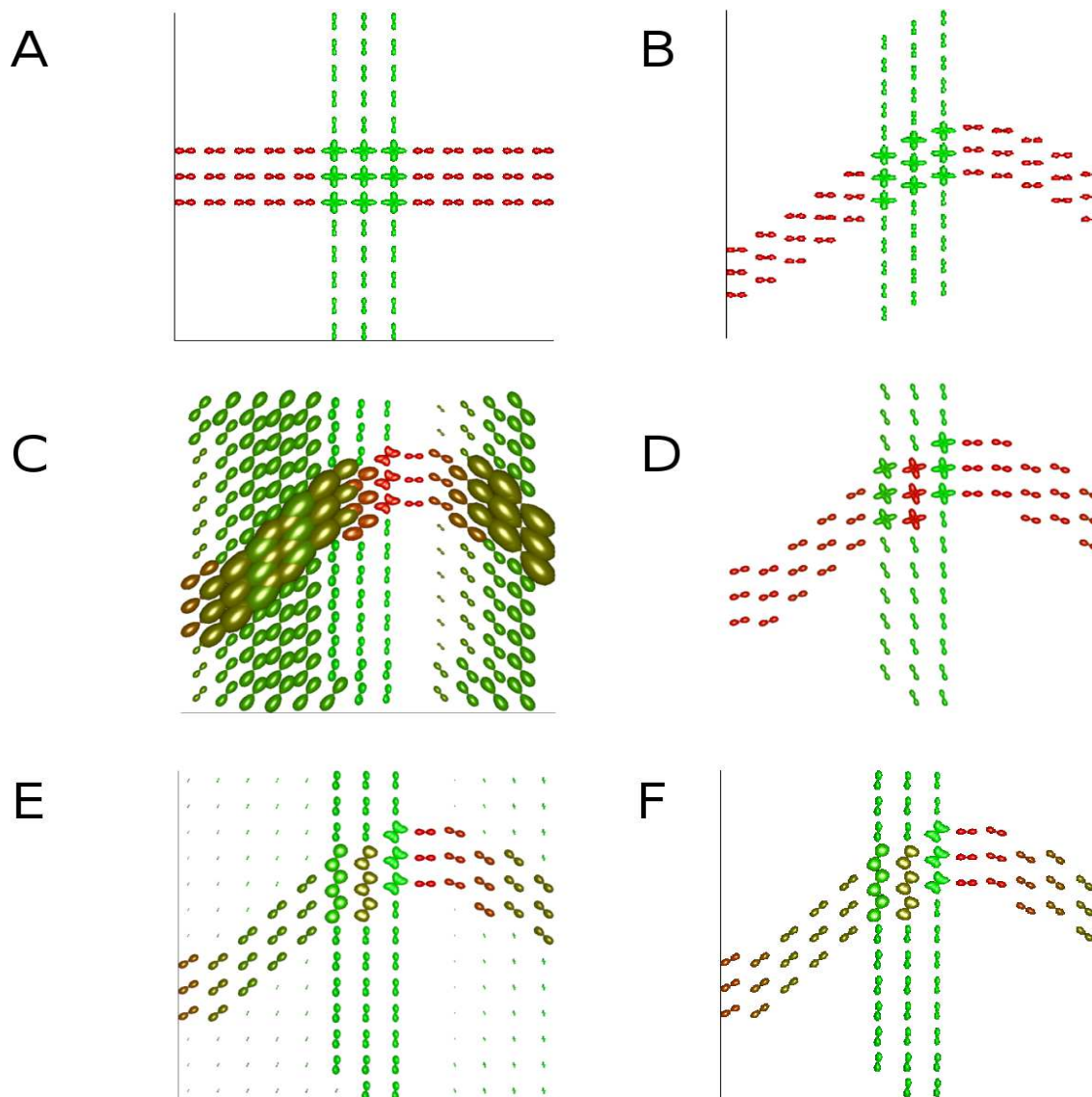


FIGURE 6.8 – A : simulation d'un croisement de fibres. B : application de la transformation sans réorientation. C : application de la SA. D : application de la stratégie FS. E : application de la SS. F : application de la SH. On remarque que les stratégies existantes permettent soit de réorienter indépendamment les paquets de fibres présents dans un croisement (SA), soit de conserver les propriétés locales du milieu (FS). Cependant, ces stratégies ne permettent pas d'avoir ces deux caractéristiques simultanément. On observe que les méthodologies élaborées dans le cadre de cette thèse (SS et SH) permettent de réorienter indépendamment les paquets de fibres présents dans un croisement, tout en conservant les propriétés locales de diffusion.

Chapitre 7

Mise en correspondance des modèles semi-globaux et globaux

1.	Mise en correspondance pour les tractographies	89
2.	Mise en correspondance pour les connectomes	90
2.1	Etat de l'art	90
2.2	Stratégie proposée dans le cadre de cette thèse	91
3.	Conclusion partielle	91

Dans ce chapitre, nous nous intéresserons brièvement à la mise en correspondance des tractographies. Dans un deuxième temps, nous nous intéresserons à celle nécessaire pour les connectomes.

1. Mise en correspondance pour les tractographies

Dans le cadre de la tractographie, nous pouvons obtenir deux types d'objets à mettre en correspondance. Le premier type correspond à des cartes de probabilités de présence de fibres. A notre connaissance, aucun article ne s'intéresse spécifiquement au recalage de ces cartes scalaires.

Bien que la mise en correspondance de cartes scalaires semble aisée, il est nécessaire de prendre des précautions. En effet, ces cartes sont fortement dépendantes des paramètres d'initialisation, tels que les ROIs d'initialisation ou le nombre de graines utilisées. Dans [YO et coll. 2009], les auteurs proposent de comparer différentes méthodes de tractographies déterministes et probabilistes. Ils proposent de normaliser ces cartes par la taille de la ROI ainsi que le nombre de graines utilisées.

Le deuxième type d'objet correspond à des structures tubulaires, obtenues généralement avec des tractographies déterministes. La mise en correspondance de ces objets comprend deux sous-cas. En effet, il existe une mise en correspondance intra-sujet et inter-sujet.

Dans le cadre intra-sujet, le but est de mettre en correspondance un ensemble de trajectoires possédant les mêmes propriétés géométriques. Ceci revient à effectuer un regroupement de fibres. Différents articles traitent du sujet [GERIG et coll. 2004, O'DONNELL & WESTIN 2007, WASSERMANN et coll. 2010]. L'article de O'Donnell et Westin [O'DONNELL & WESTIN 2007] est particulièrement intéressant. Les auteurs proposent de projeter les données dans un espace réduit, puis d'effectuer la classification dans cet espace réduit. L'avantage de ces méthodes est de ne plus utiliser la description complexe de séquences de points correspondant à une fibre.

Dans le cadre inter-sujet, différentes méthodologies ont été proposées. Une stratégie relativement simple est d'estimer la transformation sur une carte scalaire telle la FA [O'DONNELL et coll. 2009] ou directement sur les tenseurs [WASSERMANN et coll. 2010] [ZHU et coll. 2010], puis d'appliquer cette transformation à l'ensemble des trajectoires de fibres. Une deuxième méthode consiste à obtenir une carte scalaire reflétant la géométrie des paquets de fibres [GOODLETT et coll. 2009]. Dans [LEEMANS et coll. 2006], les auteurs estiment la transformation sur les cartes scalaires de courbure et de torsion des paquets de fibres.

Une autre méthode consiste à projeter les données dans un espace de grande dimension, et de considérer chaque trajectoire de fibre comme un point dans cet espace. Le problème consiste maintenant à aligner deux ensembles de points [ZVITIA et coll. 2008, ZVITIA et coll. 2010, MAYER et coll. 2011].

En raison de la complexité de la modélisation d'une trajectoire, peu d'articles proposent de recalculer directement les trajectoires entre elles. En effet, il faut, avant de pouvoir estimer la transformation optimale, savoir quelles fibres sont appariées. L'article [XU et coll. 2009] propose une méthode probabiliste pour à la fois estimer l'appariement et la transformation.

2. Mise en correspondance pour les connectomes

2.1 Etat de l'art

Il existe plusieurs stratégies pour permettre de comparer les connectomes de différents individus entre eux. Les deux points essentiels vont être le choix de la parcellisation et celui du critère de connectivité.

Le choix de la parcellisation dépend fortement de l'étude menée. Par exemple, dans [HAGMANN et coll. 2004], l'objet étudié n'est pas le connectome en soi mais des coefficients liés à celui-ci tel que le critère petit monde. Lorsqu'on veut uniquement comparer des critères extraits des connectomes, la parcellisation ne doit pas être forcément identique pour chaque individu mais on doit être à la même échelle de parcellisation [ZALESKY et coll. 2010]. On peut, par exemple, s'intéresser à l'article [HAGMANN et coll. 2004]. La première étape consiste à effectuer une parcellisation grossière en recalant un atlas segmenté sur l'image de la personne à étudier. La deuxième étape consiste à effectuer une nouvelle parcellisation, plus fine, à l'intérieur de ces régions segmentées. Cela permet aux auteurs de contrôler le niveau de parcellisation, et de s'adapter à l'anatomie de chaque individu.

Si on veut étudier l'ensemble du connectome, la parcellisation devra alors être identique pour chaque individu. Pour cela, on recalcule un atlas segmenté en N ROIs sur l'individu à étudier. Un atlas particulièrement utilisé est l'atlas AAL. On remarquera que l'on préfère toujours recalculer un atlas sur l'individu plutôt que d'amener l'ensemble des images IRM dans l'espace de l'atlas. En effet, on a vu précédemment la difficulté de transporter des trajectoires de fibres ou même des acquisitions d'IRM.

Sachant que la construction du connectome se fait dans l'espace individu et que les méthodes de tractographies sont fortement dépendantes des conditions initiales, on doit veiller à normaliser les critères de connectivité.

Dans [ITURRIA-MEDINA et coll. 2008, ROBINSON et coll. 2010], les critères de connectivité correspondent à la moyenne d'un coefficient d'anisotropie. Ce coefficient est estimé le long des trajectoires de fibres reliant deux ROIs.

Si on considère comme critère de connectivité le nombre de fibres entre deux ROIs, il faut penser à normaliser le résultat par le nombre de voxels de la ROI initiale dans le cadre déter-

ministe [YO et coll. 2009]. Il faut une normalisation pour les tractographies probabilistes en divisant par le nombre de graines utilisées.

2.2 Stratégie proposée dans le cadre de cette thèse

Nous allons présenter la méthodologie employée dans le cadre de cette thèse pour construire des connectomes et les mettre en correspondance.

Tout d’abord, nous utilisons l’atlas AAL pour parcelliser les cerveaux en 90 ROIs. Nous estimons ensuite la transformation de l’atlas AAL vers l’individu à étudier grâce à un recalage non linéaire basé sur les cartes de FA.

Dans l’espace individu, nous effectuons une tractographie déterministe grâce au logiciel MedInria (<http://www-sop.inria.fr/asclepios/software/MedINRIA/>). Nous comptons alors le nombre de fibres passant d’une ROI à l’autre. Nous normalisons le nombre obtenu par la taille de la ROI initiale [YO et coll. 2009]. Ce coefficient correspond à la mesure de connectivité utilisée dans le reste du mémoire. Toutes les valeurs obtenues sont alors stockées dans une matrice 90×90 . Remarquons que cette matrice n’est pas symétrique car pour chaque paire de ROIs i et j , nous normalisons dans un cas par la ROI i et dans l’autre cas par la ROI j .

3. Conclusion partielle

La mise en correspondance des modèles de tractographies peut s’avérer complexe, en raison de la complexité des trajectoires. Dès lors, plusieurs méthodes ont proposé de travailler dans des espaces intermédiaires, où les trajectoires sont exprimées de manière plus simple. Pour la mise en correspondance des connectomes, la méthode est relativement simple car on se ramène le plus souvent à un recalage d’atlas scalaires.

Conclusion

La mise en correspondance est une étape essentielle lors de la construction d'atlas. Plus qu'une simple étape de normalisation, elle doit tenir compte de l'information contenue dans les modèles. Par exemple, dans le cadre des modèles locaux, la mise en correspondance a mis en évidence la notion d'orientation intrinsèque aux modèles. De plus, elle a souligné que les modèles locaux n'étaient pas simplement un vecteur de paramètres, mais que ces modèles représentaient des formes particulières avec des directions privilégiées.

La mise en correspondance a permis aussi de se rendre compte que les différents modèles locaux, semi-globaux et globaux étaient complexes, et que l'information contenue dans ces modèles était riche et difficilement accessible de manière brute. De nombreux articles proposent alors de passer par des coefficients intermédiaires pour effectuer la mise en correspondance.

Dans le cadre des études statistiques, nous devons prendre en compte ces notions, privilégier les paramétrisations ayant un sens interprétable.

Troisième partie

Création d'atlas en IRM de diffusion

Introduction

Dans cette troisième partie, nous nous intéresserons spécifiquement à la conception et à l'utilisation d'un atlas. Nous présenterons d'abord les atlas existants en IRMd. Nous mettrons en évidence les difficultés rencontrées lors de leurs conceptions. Plus précisément, l'IRMd est une modalité délivrant des signaux complexes de grande dimension, nécessitant des précautions particulières notamment lors de l'analyse statistique.

Dans un deuxième temps, nous présenterons des stratégies appelées modèles génératifs pour la modélisation d'une population. Ces modèles se basent sur des représentations réduites de l'espace initial. Ces représentations réduites sont supposées capturer la variabilité inter-individuelle présente dans une population de manière compacte, c'est-à-dire que les principales variations seront décrites par un faible nombre de paramètres. On espère de cette manière là avoir un compromis entre une bonne généralisation du modèle (*i.e.* avoir un modèle concis) et une bonne description de la variabilité inter-individuelle.

Dans un troisième temps, nous proposerons une méthode originale basée sur ces représentations compactes pour décrire le problème spécifique de la comparaison d'un individu par rapport à un groupe de référence. Nous porterons une attention particulière au réglage des paramètres de la méthode. Dans un dernier chapitre, nous présenterons des résultats applicatifs dans le cadre de l'étude des patients présentant des troubles de la conscience.

Chapitre 8

Atlas en imagerie par résonance magnétique

1.	Définition et création d'un atlas	99
2.	Méthodologies proposées pour les modèles locaux	100
3.	Méthodologies proposées pour les modèles semi-globaux	103
4.	Méthodologies proposées pour les modèles globaux	105
5.	Conclusion partielle	106

La création d'atlas correspond à une chaîne de traitements complexes. Dans un premier temps, nous expliquerons quelles sont les différentes étapes de la construction d'un atlas ainsi que les techniques pour utiliser cet atlas. Dans un deuxième temps, nous décrirons les différentes méthodologies mises en œuvre en IRM.

1. Définition et création d'un atlas

Un atlas correspond à un modèle statistique plus ou moins complexe d'une population de référence. Plus précisément, un atlas permet de décrire la variabilité inter-individuelle. Une fois l'atlas créé, on peut l'utiliser pour différentes applications : la segmentation spécifique de ROI, l'aide au diagnostic, l'analyse de pathologies, la régression de population en fonction de covariables. . .

Il existe deux étapes dans les traitements utilisant un atlas : une étape de conception de l'atlas et une étape d'utilisation de l'atlas. La conception de l'atlas dépend fortement de l'application considérée. Par exemple, on construira un atlas basé sur des modèles locaux si on souhaite observer localement l'effet d'une pathologie ou si l'on désire effectuer une segmentation. On préférera l'utilisation de modèles semi-globaux et globaux si l'on désire observer l'impact d'une pathologie sur des mécanismes plus complexes mettant en jeu des paquets de fibres de plusieurs ROIs par exemple.

Une fois la classe d'objet sélectionnée, la conception d'un atlas comporte essentiellement deux parties :

- la mise en correspondance des objets à étudier dans un même référentiel,
- l'apprentissage d'un modèle statistique (par exemple l'estimation de la moyenne et de la variance).

La mise en correspondance a été étudiée en détail dans la partie précédente. Cependant le choix de l'espace de référence, *i.e.* le *template*, reste une question intrinsèque à la conception de l'atlas.

L'apprentissage du modèle est lui fortement dépendant de l'utilisation de l'atlas.

L'utilisation d'atlas peut être considérée comme un problème de reconnaissance de formes (*pattern recognition*). Le but est de trouver un modèle statistique approprié, capable de rendre compte des propriétés pertinentes des données. Plus généralement, le modèle statistique est paramétrique, et les paramètres sont estimés sur un ensemble d'apprentissage (en général, les différentes données initiales). A partir de ce modèle, une règle de décision (*decision rule*) est déterminée, et utilisée comme classifieur. Cela revient le plus généralement à effectuer un test statistique.

Enfin, lorsqu'un nouvel individu doit être comparé à l'atlas, il subit les mêmes pré-traitements que les données ayant servi à la construction de l'atlas, puis il est classifié à partir de la règle de décision.

Dans les prochaines sections, nous essaierons de soulever les différents points clés de ces étapes en IRMd.

2. Méthodologies proposées pour les modèles locaux

La construction d'atlas en IRMd pour les modèles locaux est apparue relativement vite [JONES et coll. 2002] après l'apparition des premiers modèles d'IRMd [BASSER et coll. 1994]. En effet, l'IRMd est une modalité complexe, et la construction d'atlas est un outil puissant permettant une interprétation plus aisée des cartes obtenues. Néanmoins, de par la complexité des signaux en IRMd, la création d'atlas ainsi que leur utilisation nécessitent des outils spécifiques. Dans cette section, nous attacherons à montrer les spécificités de la conception et de l'utilisation d'atlas en IRMd. Nous présenterons principalement des résultats liés à l'imagerie du tenseur de diffusion. En effet, l'utilisation d'atlas pour des modèles locaux plus complexes reste encore relativement rare.

Tout d'abord, intéressons-nous au choix du paramètre étudié pour les modèles locaux. Les premières études ont porté sur des mesures scalaires (FA, MD) [KUBICKI et coll. 2002]. Ce choix a été fait pour deux raisons principales. Premièrement les algorithmes de mise en correspondance étaient peu développés dans le cadre tensoriel. De plus, les coefficients scalaires ont une signification médicale concrète.

Néanmoins, l'interprétation de ces coefficients reste délicate. Prenons l'exemple de la FA. L'idée consensuelle est de penser qu'une pathologie entraîne systématiquement une baisse significative de la FA. Cependant une étude a montré que les altérations dues à une pathologie peuvent entraîner une augmentation de FA [HOEFT et coll. 2007]. De plus, l'interprétation de l'altération de l'anisotropie reste délicate. Seule une analyse conjointe de plusieurs coefficients, telles que la FA et la MD, permettent une analyse médicale pertinente [ZHANG et coll. 2010].

De nombreuses études se sont basées sur des vecteurs d'attributs [CHAPPELL et coll. 2008], sur des paramètres de formes [ZHANG et coll. 2009], sur les directions principales du tenseur [SCHWARTZMAN et coll. 2005], et sur le tenseur directement (voir [BANSAL et coll. 2007] [WHITCHER et coll. 2007, SCHWARTZMAN 2010]). Le fait de rajouter de l'information permet, en plus d'une meilleure interprétation, une meilleure détection. Prenons par exemple un ensemble de tenseurs ayant la même forme mais des directions principales différentes. On ne pourra pas observer de variation dans la FA, mais on pourra détecter des modifications au niveau de l'orientation [BANSAL et coll. 2007]. Cependant, trop d'information peut rendre les modifications moins visibles. Dans l'article [SCHWARTZMAN 2010], les auteurs détectent mieux les variations avec la direction principale du tenseur d'ordre 2 qu'avec l'ensemble du tenseur.

De plus, bien que comportant plus d'information que les paramètres scalaires, le tenseur sous sa forme canonique, *i.e.* représenté avec les six coefficients D_{ij} , reste complexe à interpréter. Dans [BANSAL et coll. 2007, WHITCHER et coll. 2007], les auteurs suggèrent que la paramétrisation du tenseur sous la forme $(\lambda_1, \lambda_2, \lambda_3, \alpha_1, \alpha_2, \alpha_3)$, avec les λ_i correspondant aux valeurs propres et les α_i aux angles d'Euler, permettrait une interprétation plus immédiate.

Une fois le modèle étudié choisi, il faut choisir comment effectuer la mise en correspondance, ainsi que le choix de l'espace de référence commun à l'ensemble des individus de l'étude. Nous avons vu dans la deuxième partie du mémoire que la mise en correspondance en IRMd était une étape complexe nécessitant une étape supplémentaire pour tenir compte de l'information d'orientation. De plus, la transformation peut être estimée de différentes manières : à partir des cartes scalaires d'images pondérées en T2 ou de FA ou directement à partir du tenseur. Quelques articles se sont intéressés à l'impact de ces choix dans la construction d'atlas [DÍAZ et coll. 2010, HECKE et coll. 2011]. Il en ressort principalement que l'utilisation d'atlas mis à disposition sur internet [MORI et coll. 2008] n'est pas forcément la plus propice pour une étude standard. Bien que ces atlas aient été construits à partir d'un grand nombre d'individus, ces derniers ne représentent pas forcément la population à étudier (par exemple en raison des différences d'âge) [HECKE et coll. 2011]. Il existe aussi une grande variabilité due aux différentes acquisitions en IRMd (principalement le nombre de gradients et la valeur de b). Il sera donc plus intéressant de créer un *template* spécifique pour une étude donnée [VAN HECKE et coll. 2008].

De plus, pour améliorer les propriétés statistiques de l'atlas, un filtrage spatial est parfois effectué. Cette étape aussi comporte de nombreuses possibilités : taille du filtre, choix du noyau utilisé, méthode d'interpolation. Le lecteur intéressé par le sujet pourra se rapporter aux articles [JONES & PIERPAOLI 2005, JONES & CERCIGNANI 2010].

On observe donc une grande variabilité dans les méthodologies proposées pour la construction d'atlas. On peut légitimement s'interroger sur la conséquence de cette variabilité dans les résultats. Dans l'article de [JONES et coll. 2007], les auteurs proposent de comparer les résultats de neuf méthodes d'étude pour la comparaison de deux groupes. Pour ne pas biaiser la comparaison, deux jeux de données correspondant à deux groupes distincts sont distribués à neuf laboratoires en leur proposant d'utiliser leurs traitements statistiques habituels. Il en ressort neuf résultats différents pouvant tous concorder avec la pathologie étudiée. Deux conclusions très importantes peuvent être tirées de cette étude. Premièrement, la méthodologie de la construction d'atlas a un réel impact sur le résultat. Deuxièmement, le fait que l'on n'observe pas de différence n'indique pas qu'il n'existe pas de différence.

Une fois les données mises en correspondance, la méthodologie pour la reconnaissance de formes doit être déterminée. Nous séparerons cette méthodologie en deux catégories : l'estimation des paramètres du modèle telles que la moyenne et la variance, et la détermination de règles de décision.

La modalité d'IRMd étant complexe, il est nécessaire de prendre des précautions lors de l'estimation des paramètres. En effet, l'extraction des coefficients en IRMd n'est généralement pas linéaire. Prenons pour exemple l'article [BANSAL et coll. 2007]. Dans cet article, les auteurs montrent les différences existant dans l'estimation de la FA moyenne pour un ensemble de tenseurs. La première idée est de calculer la FA de chaque tenseur et d'en faire la moyenne. La deuxième idée est de calculer le tenseur moyen, et de calculer la FA de ce dernier. Les auteurs montrent que les deux méthodes donnent des résultats fort différents, et que la seconde méthode a de meilleures propriétés statistiques.

Nous avons vu dans la partie outils mathématiques qu'il existait des métriques différentes pour un même objet mathématique. Le choix de cette métrique va fortement influencer l'estimation des paramètres comme la moyenne et la variance [LENGLET et coll. 2006]. Cette

question a été largement discutée pour l'imagerie du tenseur de diffusion, plus particulièrement pour les métriques euclidienne et log-euclidienne [ARSIGNY et coll. 2007]. Le choix de cette métrique va fortement dépendre de l'étude considérée [PASTERNAK et coll. 2010]. Cependant, d'un point de vue statistique, la question reste largement ouverte [BANSAL et coll. 2007, WHITCHER et coll. 2007, SCHWARTZMAN 2010]. On peut s'interroger sur la métrique ayant les meilleures propriétés statistiques. Plus précisément, plusieurs articles se sont intéressés à la normalité des tenseurs de diffusion d'ordre 2. Dans [SCHWARTZMAN 2010], les auteurs montrent que la normalité n'est vérifiée pour aucune des deux métriques en général. Cependant, grâce au théorème de la limite centrale, les moyennes des tenseurs suivent une distribution normale en présence d'un grand nombre d'échantillons [BANSAL et coll. 2007, SCHWARTZMAN 2010]. Dans [WHITCHER et coll. 2007, BANSAL et coll. 2007], les auteurs n'observent pas de différence entre les deux métriques. De plus, Bansal et coll. soulignent que l'estimation de la moyenne est une opération convexe, donc que celle-ci effectuée avec une métrique euclidienne et une métrique log-euclidienne donnera dans les deux cas un tenseur moyen défini positif. Il n'y a donc pas de réel consensus sur l'estimation des paramètres.

Une fois ces paramètres estimés, des tests statistiques paramétriques [BANSAL et coll. 2007, SCHWARTZMAN 2010] et non paramétriques [WHITCHER et coll. 2007] ont été proposés.

Dans [VERMA et coll. 2007], les auteurs proposent d'apprendre la variété où résident les tenseurs, et de travailler dans un espace de dimension réduite où les nouveaux vecteurs correspondant aux tenseurs possèdent les mêmes distances géodésiques que dans l'espace initial (voir Fig 8.1).

D'autres études ont utilisé des algorithmes utilisés en reconnaissance de formes basés sur l'analyse en composantes principales (ACP). L'ACP est un changement de repère, privilégiant les axes de variances maximales par rapport à un ensemble de données. On ne considère alors que les axes dont la variance est élevée. De cette manière, on effectue une réduction de dimension, avec une perte minimale d'information. Néanmoins, cette réduction de dimension est linéaire (la transformation entre les deux espaces est linéaire). Dans le cadre de l'IRMd, deux ACPs modifiées ont été proposées. La première est effectuée dans l'espace des tenseurs avec une métrique riemannienne et est appelée analyse en géodésiques principales [FLETCHER & JOSHI 2004]. La deuxième utilise des méthodes à noyaux pour effectuer l'ACP dans un nouvel espace dit de caractéristiques, et est appelée ACP à noyaux [KHURD et coll. 2007].

Dans [KHURD et coll. 2007], les auteurs proposent de comparer les techniques de réduction de dimension et l'ACP à noyaux. Ces différentes techniques proposent une alternative intéressante aux études plus conventionnelles de tests d'hypothèses. Elles permettent de rendre compte de la variabilité des données. Ces méthodes ne sont plus paramétriques, dans le sens où l'on ne suppose plus que les tenseurs appartiennent à un espace spécifique (espace euclidien ou variété riemannienne au sens de [FLETCHER & JOSHI 2004] ou de [ARSIGNY et coll. 2007]). Plus précisément, ces méthodes supposent que les tenseurs résident dans des sous-variétés plus complexes pouvant notamment dépendre des pathologies rencontrées [KHURD et coll. 2007]. Cette idée se retrouve dans l'article [GERBER et coll. 2010] où les auteurs différencient la variété de l'espace et la sous-variété induite par les données.

La création d'atlas pour d'autres modalités que l'imagerie du tenseur de diffusion reste peu développée [BARMPOUTIS et coll. 2009, CAAN et coll. 2009, BOUX et coll. 2010]. Les études portent sur la création d'un individu moyen avec des images pondérées en diffusion [BOUX et coll. 2010], des modèles utilisant deux tenseurs d'ordre 2 [CAAN et coll. 2009] et des tenseurs d'ordre 4 [BARMPOUTIS et coll. 2009]. On peut aussi remarquer que de nombreuses autres études statistiques ont été proposées dans le cadre du tenseur d'ordre 2 ainsi que pour des modèles plus complexes pour la classification des tissus. Bien que ces études traitent de problèmes liés à la reconnaissance de formes, elles ne seront pas abordées dans cette thèse.



FIGURE 8.1 – (a) : schéma représentant les tenseurs appartenant à des variétés plus complexes que la variété des matrices symétriques définies positives. (b) : schéma représentant deux chemins reliant deux tenseurs : le trait plein en rouge correspond au chemin respectant la géométrie de la variété sous-jacente, et celui en vert en pointillé correspond au chemin si les tenseurs appartenaient à un espace euclidien. Figure extraite de [VERMA et coll. 2007].

On peut donc observer que la construction d'atlas pour les modèles locaux est complexe, et nécessite beaucoup de précautions et de rigueur. On peut se rendre compte que plutôt que d'utiliser un modèle rigide, que ce soit pour la sélection du *template* ou pour l'étude des tenseurs, il est préférable d'utiliser des modèles plus adaptés aux données étudiées.

3. Méthodologies proposées pour les modèles semi-globaux

Dans le cadre de l'étude des paquets de fibres de substance blanche, l'atlas est utilisé principalement de deux manières : la classification automatique des différentes trajectoires en paquets de fibres cohérents (d'un point de vue anatomique ou géométrique), et l'étude de pathologie pour des paquets de fibres spécifiques.

La classification automatique des différentes trajectoires de fibres est intéressante pour de multiples raisons. D'une part, la géométrie complexe des trajectoires, ainsi que le nombre élevé de ces dernières (pour un cerveau, nous pouvons estimer plusieurs dizaines de milliers de fibres) empêchent une classification manuelle exhaustive. D'autre part, une fois les paquets obtenus, nous pouvons comparer ces derniers entre plusieurs individus.

Certains algorithmes de classification sont intéressants car ils représentent les différentes trajectoires de manière compacte en utilisant des techniques de réduction de dimension et d'apprentissage de variétés (*manifold learning*) (voir par exemple [O'DONNELL & WESTIN 2007] [WASSERMANN et coll. 2008]). L'ensemble des trajectoires de fibres, objets mathématiques résidant dans un espace complexe, est projeté dans un espace de dimension réduite. Les distances entre les trajectoires des fibres sont conservées dans ce nouvel espace réduit. L'étape de classification est alors effectuée dans ce dernier espace. De plus, l'estimation d'un individu moyen ou d'un prototype est plus aisée à réaliser dans cet espace réduit. Pour ce faire, dans [O'DONNELL et coll. 2009], les auteurs sélectionnent le prototype du paquet tel que le point le plus proche de la moyenne traditionnelle des points de l'espace réduit (ils ne sélectionnent pas la moyenne directement pour éviter d'une part de calculer une régression entre les deux espaces, et d'autre part le point sélectionné correspond réellement à une trajectoire ce qui n'est pas assuré par la régression).

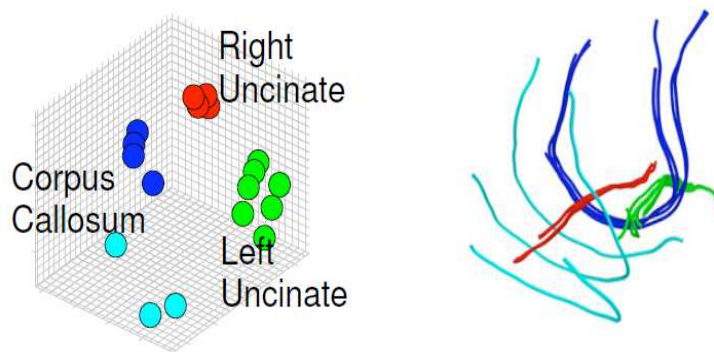


FIGURE 8.2 – Schéma représentant l'espace réduit à gauche et l'espace initial des fibres correspondant à droite. La classification est effectuée dans l'espace réduit. Chaque couleur correspond à un paquet de fibres spécifiques. On observe d'une part que la classification est plus aisée dans l'espace de gauche que dans celui de droite. D'autre part, les résultats de la classification obtenus dans l'espace réduit sont cohérents dans l'espace initial. Figure extraite de [O'DONNELL & WESTIN 2007].

Étudions maintenant l'utilisation d'un atlas pour comparer un individu par rapport à un groupe de référence, ou un groupe de référence par rapport à un autre groupe. L'impact des pathologies sur les paquets de fibres de substance blanche peut être de deux sortes. Premièrement, on peut observer des déformations physiques des trajectoires comme lors de la présence de tumeurs ou de scléroses en plaques. Deuxièmement, on peut observer des altérations des propriétés de diffusion le long des paquets de fibres [CICCARELLI et coll. 2008].

De nombreuses méthodologies ont été proposées pour étudier les propriétés locales de diffusion tels que la FA ou le tenseur de diffusion pour les paquets de fibres, ou les densités locales de fibres. On peut classer ces stratégies en deux catégories : la première considère des formes géométriques et la seconde des images scalaires.

La première catégorie consiste donc à analyser une forme en trois dimensions. Pour cela, on peut projeter l'ensemble des trajectoires d'un même paquet de fibres sur une trajectoire prototype représentant au mieux le paquet, puis étudier localement les coefficients (voir par exemple [HUA et coll. 2008, MADDAH et coll. 2008, GOODLETT et coll. 2009, O'DONNELL et coll. 2009]). Ces techniques nécessitent des paramétrisations souvent complexes rencontrant des problèmes d'appariement entre la trajectoire prototype et les trajectoires du paquet considéré (voir [WASSERMANN et coll. 2010]). Par exemple, on peut rencontrer une forte variabilité à la fin des trajectoires qu'un prototype ne pourra pas décrire. Une autre méthode étudie la densité de probabilité de certains coefficients le long des trajectoires des paquets de fibres [DEMIRALP & LAIDLAW 2010]. Cette technique ne nécessite plus le choix d'un prototype, mais nécessite encore la paramétrisation et l'appariement de chacune des trajectoires. Une dernière stratégie consiste à segmenter l'ensemble des paquets de fibres, afin d'obtenir une forme [OH et coll. 2007, YUSHKEVICH et coll. 2008]. L'analyse statistique considère des voisinages locaux tout le long de cette forme, ou considère le squelette de la forme à la manière du TBSS [SMITH et coll. 2006].

Les algorithmes précédents supposent l'utilisation d'objets géométriques souvent complexes. Une deuxième classe de méthodes est basée sur les cartes de probabilités de présence [HUA et coll. 2008, WASSERMANN et coll. 2010, PANNEK et coll. 2011]. L'utilisation de ces cartes est intéressante car on se ramène à une étude d'une image scalaire conventionnelle. Néanmoins, on perd la structure géométrique des paquets de fibres.

L'étude des formes des paquets de fibres a été moins abordée. Nous pouvons citer [COROUGE et coll. 2006]. Dans cet article, les auteurs proposent une modélisation des formes basées sur une ACP. Cette technique permet d'étudier les différents modes de variabilité correspondant aux premières composantes principales de l'ACP des paquets de fibres. Néanmoins, l'utilisation d'une technique de réduction de dimension non linéaire telle que l'ISOMAP [TENENBAUM et coll. 2000] aurait été plus pertinente pour des objets aussi complexes que les fibres.

4. Méthodologies proposées pour les modèles globaux

L'étude de la connectivité anatomique et fonctionnelle est très récente. Dès lors, il existe moins d'études traitant de la création d'atlas que pour les modèles locaux et semi-globaux.

Les premières études comparant deux populations différentes ont été effectuées sur des coefficients scalaires décrivant des propriétés topologiques des graphes. Plusieurs pathologies ont ainsi été étudiées telles que la maladie d'Alzheimer [STAM et coll. 2007, HE et coll. 2008], la sclérose en plaques [HE et coll. 2009], et la schizophrénie [LIU et coll. 2008, RUBINOV et coll. 2009]. Une étude récente propose d'utiliser un vecteur d'attributs contenant plusieurs de ces coefficients scalaires pour classer deux populations [ITURRIA-MEDINA et coll. 2011]. Les auteurs montrent qu'il existe une redondance de l'information entre ces coefficients, et proposent d'utiliser une technique appelée l'analyse discriminante linéaire (ADL) permettant d'estimer une combinaison des coefficients pour discriminer au mieux les deux populations.

D'autres études ont été réalisées dans le cadre du connectome : les études basées sur l'ensemble du connectome et les études proposant une représentation compacte des connectomes.

Les connectomes résident dans des espaces de grande dimension. Pour rappel, si on considère une parcellisation en 90 ROIs, on obtient un connectome appartenant à un espace de dimension 4005 en considérant qu'il n'existe pas de notion de direction dans la mesure de connectivité entre deux ROIs. Vouloir travailler sur l'ensemble du connectome nécessite de prendre des précautions particulières.

Dans [ZALESKY et coll. 2010], les auteurs proposent une méthode statistique adaptée aux graphes pour prendre en compte le phénomène de comparaisons multiples. Prenons un exemple de ce qui peut se passer si nous effectuons un grand nombre de tests univariés pour chaque arête du graphe. Considérons deux ensembles d'échantillons tirés d'une même distribution normale. Bien que l'on s'attende à obtenir des moyennes identiques, on peut obtenir par pur hasard des moyennes très éloignées de la distribution. Ce phénomène a une probabilité d'apparaître d'autant plus grande que le nombre de tests effectués est important. Ce problème est bien connu en traitement d'images, et plusieurs solutions ont été proposées [SHAFFER 1995, MACHADO 2007]. La contribution de l'article [ZALESKY et coll. 2010] est la prise en compte de la topologie des graphes pour effectuer ces corrections multiples.

Intéressons-nous à une autre étude en connectivité fonctionnelle [VAROQUAUX et coll. 2010], où les auteurs travaillent directement sur les matrices de connectivité en supposant qu'elles appartiennent à l'espace des matrices symétriques définies positives. Ils utilisent alors une métrique riemannienne à la manière des études faites en tenseur de diffusion [LENGLET et coll. 2006]. La grande différence avec ces dernières réside dans le fait que l'espace des matrices de connectivité est de dimension très grande par rapport à l'espace des tenseurs. Les calculs algébriques telle que l'inversion des matrices sont dès lors beaucoup plus complexes.

Comme nous venons de l'observer, les connectomes sont des objets de grande dimension. L'idée d'utiliser des représentations réduites semble pertinente.

Dans [DE BOER et coll. 2011], les auteurs proposent d'effectuer des régressions et des classifications à partir d'un ensemble réduit de données obtenu par exemple par l'ACP. Dans [ROBINSON et coll. 2010], les auteurs proposent aussi d'effectuer une classification des données, *via* un algorithme adapté de l'ADL, dans un espace réduit obtenu par l'ACP. Dans cet article, les auteurs vont plus loin. Une fois la classification obtenue, ils proposent de revenir dans l'espace initial pour visualiser où se situent les modifications dans le connectome.

Dans ces deux derniers articles, les auteurs ont choisi d'effectuer leurs étapes de classification ou de régression dans un espace réduit, où l'information pertinente est suffisante pour discriminer les différents individus.

Certaines études proposent aussi de coupler la connectivité fonctionnelle et anatomique. Citons par exemple l'article [VENKATARAMAN et coll. 2010] où les auteurs effectuent une étude statistique conjointe pour améliorer les détections. Dans l'article [DELIGIANNI et coll. 2010], les auteurs proposent de prédire l'une des connectivités à partir de l'autre. Ils soulignent les problèmes liés à la dimension importante de l'espace. Ils proposent alors d'effectuer la prédiction dans un espace réduit obtenu par l'ACP.

5. Conclusion partielle

Dans ce chapitre, nous avons mis en évidence la complexité de la construction d'atlas en IRM_d, ainsi que la complexité des calculs statistiques correspondants. Les espaces rencontrés sont en général de grande dimension, pas forcément euclidiens. De plus, plusieurs études soulignent le fait que l'information contenue dans ces objets est souvent redondante. Plutôt que travailler dans ces espaces complexes, plusieurs études ont commencé à utiliser des représentations réduites pour effectuer les pré-traitements nécessaires à la mise en place de l'atlas, ainsi que les traitements statistiques.

Chapitre 9

Modèles et représentations compactes en imagerie médicale

1. Méthodes et algorithmes pour les représentations compactes . . .	107
1.1 Réduction de dimension et ISOMAP	108
1.2 Méthodes à noyaux et SVDD	110
2. Modélisation d'une population de référence	113
2.1 Comparaison en imagerie médicale avec des représentations compactes	115
2.2 Comparaison d'une population par rapport à une autre population . .	115
2.3 Détection et analyse de données aberrantes	116
3. Conclusion partielle	117

Nous avons mis en évidence dans le chapitre précédent l'utilité des représentations compactes lors de la modélisation de données complexes en IRMd. Dans le cadre de cette thèse, la pathologie étudiée correspond au coma. L'étiologie des patients peut être fort diverse, et nous oblige à considérer le problème de la comparaison d'un individu par rapport à une population de référence. Ces méthodes peuvent être considérées comme des techniques de détection de données aberrantes. Dans un deuxième temps, il est intéressant de déterminer où se situent les régions modifiées par rapport à la population de référence. Le but est de déterminer si le nouvel individu appartient ou non à une population de référence.

Dans ce chapitre, nous proposons dans un premier temps d'étudier deux méthodes différentes pour représenter des populations de manière compacte. Dans un deuxième temps, nous étudierons les différentes utilisations de ces représentations compactes en IRMd. Dans une dernière partie, nous nous intéresserons spécifiquement aux techniques d'étude pour la détection de données aberrantes.

1. Méthodes et algorithmes pour les représentations compactes

En imagerie médicale, les chercheurs sont souvent confrontés à un problème dénommé le fléau de la dimension (*the curse of dimensionality*), formalisé par Richard Bellman [BELLMAN 1961, LEE & VERLEYSSEN 2007], qui consiste à avoir un faible nombre d'échantillons (une dizaine d'individus) pour des données de très grande dimension (plusieurs milliers de voxels pour une image).

Prenons l'exemple¹, qui consiste à considérer cent observations appartenant à l'intervalle

1. extrait de http://en.wikipedia.org/wiki/Curse_of_dimensionality

unidimensionnel $[0, 1]$ de l'espace des réels. Il est possible, à partir de ces données, d'effectuer un histogramme, et de tirer des inférences statistiques. Cependant, si on considère maintenant que les cent observations appartiennent à l'espace multivarié $[0, 1]^{10}$, ces dernières seront des points isolés dans un espace vaste et vide. Si on voulait avoir une densité de points équivalente à celle des cent observations de $[0, 1]$, il faudrait environ 10^{20} observations, ce qui est généralement complexe informatiquement et impossible en terme d'acquisitions cliniques.

En général, les observations de grandes dimensions résident dans des variétés de plus faible dimension possédant des géométries parfois complexes. Pour résoudre ces problèmes liés à la dimension, des algorithmes de réduction de dimension ont été développés. De plus, des méthodes à noyaux ont aussi été élaborées pour pouvoir réaliser des inférences de façon robuste.

Dans cette section, nous présenterons une méthode de réduction non linéaire dénommée ISOMAP. Nous intéresserons dans un deuxième temps aux méthodes à noyaux, et plus spécifiquement à la méthode du *support vector domain description* (SVDD) permettant l'apprentissage du support de distributions complexes.

1.1 Réduction de dimension et ISOMAP

ISOMAP

Précédemment, nous avons décrit la méthode d'ACP qui vise à effectuer une réduction de dimension des données en recherchant des axes indépendants expliquant au mieux la variabilité des données. L'ACP est une méthode de projection linéaire des données sur les axes représentant une grande variabilité. Cependant, l'ACP n'est pas performante dans le cas non-linéaire (voir Fig 9.1). Une méthode non linéaire dans un espace de dimension réduite a été proposée pour pallier cette limitation : l'ISOMAP [TENENBAUM et coll. 2000].

Le principe de l'ISOMAP consiste à préserver les distances entre les points de l'espace réduit. Plus précisément, des points proches (resp. éloignés) dans l'espace initial, au sens d'un critère de similarité ou de dissimilarité, resteront proches (resp. éloignés) dans l'espace réduit. L'ISOMAP repose sur l'algorithme de mise à l'échelle multidimensionnelle (*multidimensional scaling* (MDS)). Le MDS permet à partir de distances entre points de déterminer un système de coordonnées réduit qui préserve ces distances. Plus précisément, le MDS détermine les coordonnées d'un espace de dimension réduite de dimension d , telles que le critère suivant soit minimisé :

$$\arg \min_{\mathbf{x}_1, \dots, \mathbf{x}_N} \frac{\sum_{i \neq j} (\|\mathbf{x}_i - \mathbf{x}_j\| - d_{ij})^2}{\sum_{i \neq j} d_{ij}^2}, \quad (9.1)$$

où d_{ij} correspond à la distance géodésique \mathcal{D} entre les points \mathbf{y}_i et \mathbf{y}_j dans l'espace initial, et $\|\mathbf{x}_i - \mathbf{x}_j\|$ est la distance euclidienne \mathcal{G} dans l'espace réduit entre les points \mathbf{x}_i et \mathbf{x}_j . Dans le MDS classique, la distance \mathcal{D} utilisée est euclidienne. Dans le cadre de l'ISOMAP, on construit un graphe reliant chaque point \mathbf{x}_i à ses voisins. Les sommets correspondent aux points et les arêtes aux distances entre eux. On définit le voisinage autour de \mathbf{y}_i en considérant les k -plus proches voisins ou une sphère de rayon r centrée sur \mathbf{y}_i .

On estime alors la distance géodésique pour chaque paire de points par la distance la plus courte parcourue sur le graphe (algorithme de Floyd ou de Dijkstra). On applique alors le MDS sur ces nouvelles distances géodésiques pour obtenir un nouveau système de coordonnées.

Il existe un grand nombre de méthodes de réduction de dimension. Le lecteur intéressé pourra se référer au livre [LEE & VERLEYSSEN 2007].

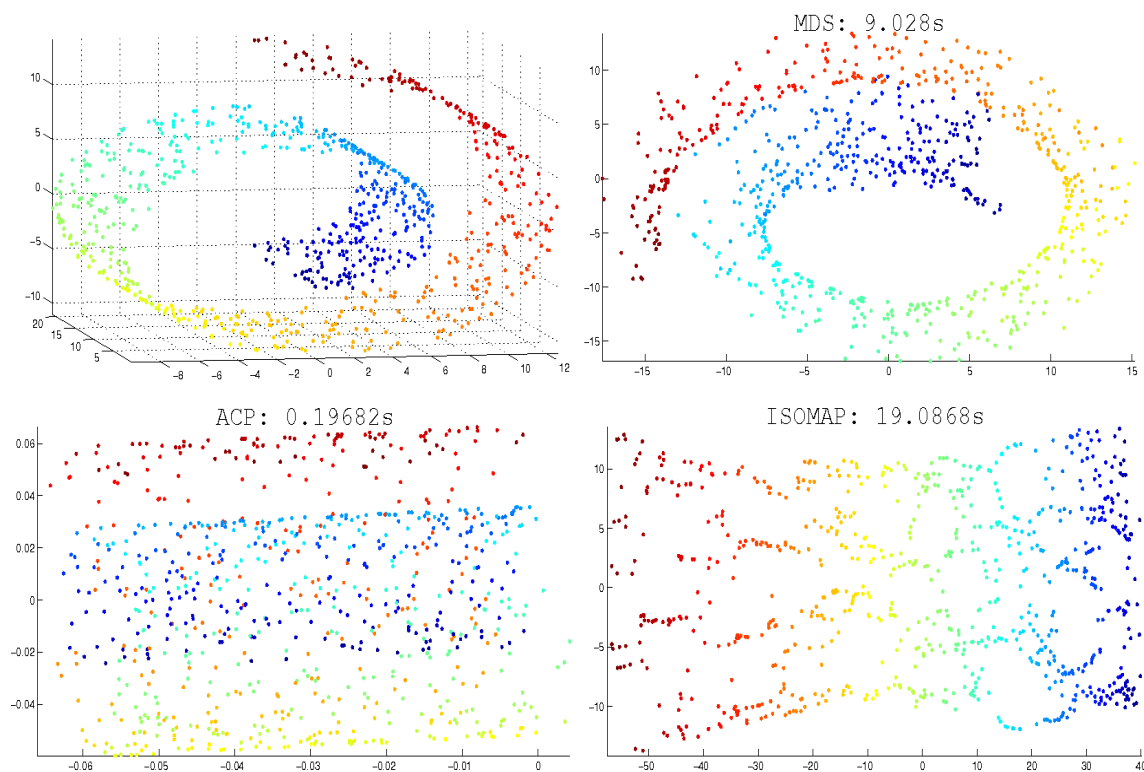


FIGURE 9.1 – Exemple du *swiss roll*. En haut à droite : 800 données initiales appartenant à une variété de dimension deux résidant dans un espace de dimension trois. En haut à gauche : le résultat obtenu par MDS. En bas à gauche : celui obtenu par l'ACP. En bas à droite : celui obtenu par l'ISOMAP. Pour chaque méthode, le temps de calcul figure à côté du nom de la méthode. On observe que seul l'ISOMAP permet de projeter correctement le *swiss roll* qui est une variété non linéaire. On peut aussi remarquer que la méthode de l'ACP est cent fois plus rapide que l'ISOMAP. Plus on aura de données initiales, plus cette proportion entre les deux méthodes sera grande. Figures créées grâce au site web <http://www.math.ucla.edu/~wittman/mani/>.

Expériences sur les tenseurs d'ordre 2

Nous proposons ici une expérience sur les tenseurs d'ordre 2 de dimension deux, paramétrés par des vecteurs appartenant à \mathbb{R}^3 et représentés géométriquement par des ellipses.

Cette expérience a deux buts : premièrement, montrer que l'espace réduit obtenu par l'ISOMAP correspond à ce qu'on attendait intuitivement, et deuxièmement montrer l'impact des métriques lors de la comparaison d'objets. Pour cela, nous appliquons l'ISOMAP sur deux jeux de données différents modélisant des variétés de dimension un et deux plongées dans \mathbb{R}^3 . Le premier jeu contient un ensemble de tenseurs ayant subi différentes rotations paramétrées par un angle. Le deuxième jeu contient des tenseurs paramétrés par deux valeurs propres que l'on fait varier linéairement.

On peut observer que l'ISOMAP permet dans les deux cas d'obtenir la structure attendue de la variété. De plus, les propriétés des métriques de l'espace initial sont conservées dans l'espace de dimension réduite. Plus précisément, dans le cadre des métriques riemannienne et log-euclidienne, les tenseurs négatifs ou nuls sont projetés à l'infini. Cette expérience montre donc qu'il est nécessaire de choisir dans l'espace initial des métriques adaptées.

On peut aussi remarquer que les coordonnées obtenues par le MDS ne sont pas uniques : on peut arbitrairement translater et appliquer une rotation aux vecteurs \mathbf{x}_i de l'espace réduit

sans modifier les distances entre elles.

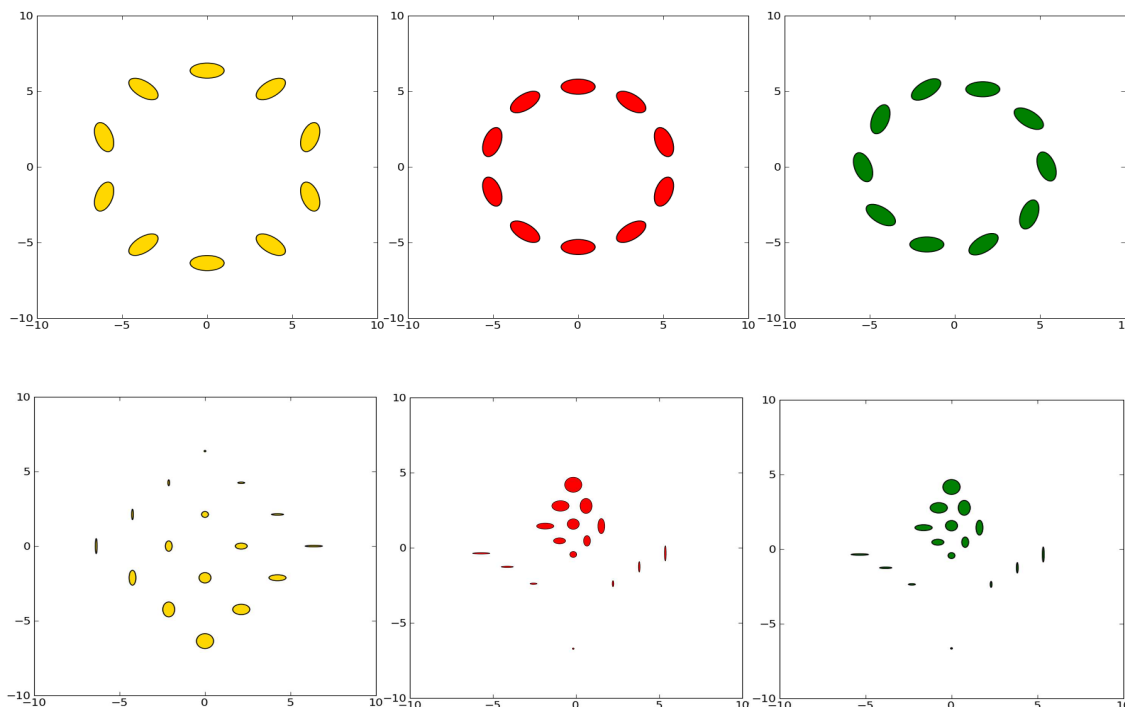


FIGURE 9.2 – Deux ensembles de tenseurs de dimension deux ayant subi différentes transformations, et projetés avec l’ISOMAP avec différentes métriques. De gauche à droite : la métrique utilisée est celle de Frobenius, puis celle riemannienne, et enfin celle log-euclidienne. La ligne du haut correspond à une variété de dimension un paramétrée par un angle de rotation. La ligne du bas correspond à une variété de dimension deux paramétrée par les deux valeurs propres des tenseurs. On observe bien les phénomènes attendus : les métriques riemannienne et log-euclidienne vont projeter les tenseurs non définis ou négatifs à l’infini. On peut remarquer pour la première figure que bien que la variété soit de dimension un, il faut un espace de dimension deux pour pouvoir la projeter correctement.

1.2 Méthodes à noyaux et SVDD

Méthodes à noyaux

Les méthodes à noyaux sont une classe d’algorithmes utilisés en reconnaissance de formes (*pattern recognition*). Le plus connu de ces algorithmes est le *support vector machine* (SVM) [VAPNIK 1999]. Le but général en reconnaissance de formes est d’étudier des relations tels que le classement (*ranking*), la corrélation ou la classification, pour des formes, des images, ou des séquences de points, par exemple.

Les approches à noyaux, contrairement à la réduction de dimension, proposent de plonger les données initiales dans un espace F de caractéristiques (*features space*) de plus grande dimension grâce à une fonction ϕ , où chaque coordonnée correspond à une caractéristique des données initiales. Une fois dans cet espace, de nombreux algorithmes linéaires peuvent être utilisés tels que l’ACP, l’ADL ... Le résultat dans l’espace de départ correspond alors à des traitements non linéaires.

La détermination d’une telle fonction ϕ est complexe et arbitraire. L’idée est donc de ne

jamais décrire cette fonction ϕ . Plus précisément, on se rend compte que la fonction ϕ n'apparaît qu'à travers des produits scalaires $\langle \phi(\mathbf{x}) | \phi(\mathbf{y}) \rangle$, où \mathbf{x} et \mathbf{y} appartiennent à l'espace initial. L'« astuce du noyau » (*kernel trick*) est de considérer une fonction noyau K sans considérer explicitement ϕ . On écrit alors :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i) | \phi(\mathbf{x}_j) \rangle. \quad (9.2)$$

Cette fonction $K(.,.)$ est calculée pour l'ensemble des éléments de l'espace de départ. Les différentes valeurs sont stockées dans une matrice \mathbf{K} . Les différents algorithmes dans l'espace F des caractéristiques utilisent uniquement cette matrice \mathbf{K} .

Les fonctions K à noyaux sont définies positives. Différentes classes de noyaux existent, les plus connues étant le noyau linéaire, les noyaux polynomiaux, et le noyau gaussien. Le noyau gaussien s'écrit :

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), \quad (9.3)$$

avec $\|\cdot\|$ la norme euclidienne, et σ un paramètre à déterminer. Le choix de σ est déterminant pour l'étude [BRERETON & LLOYD 2010]. L'estimation de ce paramètre sera étudiée dans le prochain chapitre. Ce noyau sera utilisé dans la suite du mémoire.

A la manière de l'ISOMAP, on utilise un nouvel espace intermédiaire ayant de bonnes propriétés. L'avantage des méthodes à noyaux est de pouvoir généraliser des algorithmes linéaires, et effectuer des traitements non linéaires. Le principal inconvénient est de ne généralement jamais avoir accès aux éléments $\phi(\mathbf{x})$. Plus précisément, on ne sait pas si l'espace de caractéristiques a un sens réel. On ne peut d'ailleurs que vérifier les résultats dans l'espace de départ. Malgré cette absence de connaissance de l'espace de caractéristiques, les différents algorithmes utilisant ces méthodes à noyaux ont montré par la pratique de bonnes propriétés. Le lecteur intéressé pourra consulter les ouvrages suivants [VAPNIK 1999, SHAW-TAYLOR & CRISTIANINI 2004].

SVDD

Intéressons-nous maintenant à une classe particulière de méthodes à noyaux : le SVDD [TAX & DUIN 1999, TAX & DUIN 2004]. Cet algorithme constitue un classifieur supervisé pour l'apprentissage d'une classe (*one-class supervised classifier*). Une autre catégorie de classifieur en une classe existe appelée ν -SVM [SCHÖLKOPF et coll. 2001]. Dans cette approche, le but est de trouver l'hyperplan optimal qui maximise la distance entre les données dans l'espace F des caractéristiques et son origine. Les méthodes ν -SVM et SVDD sont identiques pour certains noyaux tels que le noyau gaussien. Ces deux approches sont le pendant de l'algorithme du *support vector machine* [VAPNIK 1999] qui est un classifieur supervisé en deux classes. Nous nous intéresserons uniquement à la méthodologie SVDD dans la suite du mémoire.

Le SVDD a pour but de trouver l'hyperboule de rayon minimal R contenant l'ensemble des données \mathbf{x}_i de la base d'apprentissage (voir Fig. 9.3) :

$$\arg \min_{R^2, \mathbf{a}} L_0(R^2, \mathbf{a}, \zeta) = R^2 + C \sum_{i=1}^N \zeta_i \quad (9.4)$$

sous contraintes $\|\mathbf{x}_i - \mathbf{a}\|^2 \leq R^2 + \zeta_i, \zeta_i \geq 0, \forall i = 1, \dots, N.$

Les coefficients ζ_i (*slack variables*) représentent une quantité autorisée pour chaque point \mathbf{x}_i d'être en dehors de l'hyperboule. Le coefficient C est une pondération permettant aux points extérieurs à l'hyperboule d'être plus ou moins loin de la frontière. Si ζ_i est égale à zéro, le point

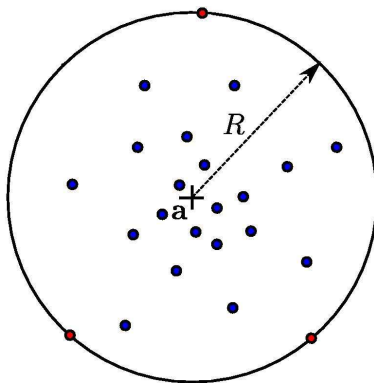


FIGURE 9.3 – Schéma explicitant l'algorithme du SVDD avec un noyau linéaire. \mathbf{a} et R correspondent respectivement au centre et au rayon de l'hyperbole contenant les données. Les données d'apprentissage sont en bleu. Les points en rouge correspondent aux vecteurs-support. Figure adaptée de [SJÖSTRAND et coll. 2007].

\mathbf{x}_i appartient à l'hyperbole. Ces coefficients sont réglés si on a des *a priori* sur des données n'appartenant pas à la classe ou des données connues comme aberrantes. Dans le reste de ce mémoire, nous supposons que tous les éléments de la base d'apprentissage appartiennent à la classe de référence. Par conséquent, ζ_i sera nulle pour tout i . Avec $\zeta_i = 0$, la valeur de C n'a pas d'importance car tous les points appartiennent à l'hyperbole. Sous ces conditions, le problème dual de l'équation 9.4 peut être formulé de la manière suivante :

$$\arg \min_{\boldsymbol{\alpha}} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \langle \mathbf{x}_i | \mathbf{x}_j \rangle - \sum_{i=1}^N \alpha_i \langle \mathbf{x}_i | \mathbf{x}_i \rangle \quad (9.5)$$

sous contraintes $\sum_{i=1}^N \alpha_i = 1, \alpha_i \in [0, 1], \forall i = 1, \dots, N,$

avec $\boldsymbol{\alpha}$ le vecteur contenant les coefficients α_i . Cependant, les hyperboles euclidiennes ne permettent pas de modéliser de manière fine des géométries complexes. En conséquence, on peut utiliser la transformation $\phi : \mathbb{R}^d \rightarrow F$ qui permet de considérer la hyperbole B_F avec un rayon R_F dans l'espace F des caractéristiques. Grâce à l'« astuce du noyau » (voir eq. 9.2) nous permettant de ne pas expliciter ϕ , on peut écrire le précédent problème d'optimisation de la manière suivante :

$$\arg \min_{\boldsymbol{\alpha}} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^N \alpha_i K(\mathbf{x}_i, \mathbf{x}_i) \quad (9.6)$$

sous contraintes $\sum_{i=1}^N \alpha_i = 1, \alpha_i \in [0, 1], \forall i = 1, \dots, N.$

Cette optimisation peut être résolue par une programmation quadratique.

Le centre \mathbf{a}_F de la hyperbole et la distance $d_{\mathbf{x}}$ de l'élément \mathbf{x} au centre \mathbf{a}_F dans l'espace F des caractéristiques peuvent s'exprimer uniquement à partir des α_i :

$$\mathbf{a}_F = \sum_{i=1}^N \alpha_i \phi(\mathbf{x}_i), \quad (9.7)$$

$$d_{\mathbf{x}}^2 = K(\mathbf{x}, \mathbf{x}) - 2 \sum_{i=1}^N \alpha_i K(\mathbf{x}_i, \mathbf{x}) + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j).$$

Le rayon R_F de la hyperboule B_F correspond à la distance d_i de n'importe quel point \mathbf{x}_i se situant sur la frontière de la hyperboule B_F .

Le SVDD est une technique permettant à la fois d'estimer le support d'une distribution complexe [SCHÖLKOPF et coll. 2001] et à la fois de donner un critère d'appartenance d'un individu à une population donnée. Plus précisément, si le nouvel individu est en dehors de la population, on peut quantifier la distance de cet individu à cette population.

Une autre propriété importante de cet algorithme est sa capacité à modéliser localement la frontière de la distribution. Typiquement, si on déplace ou on enlève un point se situant sur la frontière de la hyperboule ou proche de celle-ci alors on altère fortement la frontière. Par contre, si on enlève un individu à l'intérieur de la distribution, la frontière ne sera pas modifiée. Ce phénomène s'explique grâce à l'utilisation des noyaux qui modélisent localement l'interaction entre les individus.

2. Modélisation d'une population de référence

Dans cette section, nous discuterons de la modélisation d'une population en imagerie médicale. Nous nous intéresserons particulièrement aux modèles génératifs \mathbf{f} basés sur des représentations compactes de la forme :

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) + \varepsilon, \quad (9.8)$$

avec \mathbf{y} appartenant à \mathbb{R}^D correspondant à une observation, \mathbf{x} un vecteur appartenant à \mathbb{R}^d correspondant à une paramétrisation compacte de \mathbf{y} , et ε considéré comme du bruit. En général, nous avons D très grand devant d . Les vecteurs \mathbf{y} sont considérés comme des points de l'espace \mathbb{R}^D . Ils peuvent décrire une grande variété d'objets : images d'IRM ou de PET-SCAN vectorisé, sillons corticaux, tenseurs de diffusion ... On entend par modèle génératif un modèle capable de générer des données \mathbf{y} réalistes. Ce type de modèle est à opposer à un modèle discriminatif qui cherche à décrire les relations entre les données \mathbf{y} avec un minimum d'hypothèses sur la structure de ces données \mathbf{y} [BOUCHARD 2005]. Un modèle compact, quant à lui, permet de paramétrer la population avec peu de composantes, ces dernières rendant compte des variations principales à l'intérieur de cette population.

Les premiers modèles (voir par exemple [TALAIRACH & TOURNOUX 1988]) ont considéré un individu unique. Cependant, un individu ne permet pas de représenter la variabilité existant dans une population donnée. Dans [TALAIRACH & TOURNOUX 1988], le modèle présenté est le cerveau *post mortem* d'une femme âgée de soixante ans, peu représentatif de la population moyenne.

Les atlas ont ensuite pris la forme suivante :

$$\mathbf{y} = \boldsymbol{\mu}_y + \varepsilon, \quad (9.9)$$

où $\boldsymbol{\mu}_y$ est la moyenne des individus. Cela correspond à une fonction $\mathbf{f}(\mathbf{x})$ constante dans l'équation 9.8. Ces atlas sont constitués d'un grand nombre d'individus (152 individus dans le cadre de l'atlas proposé dans [MAZZIOTTA et coll. 1995]). Ils ont l'inconvénient d'omettre les particularités des individus. La variabilité inter-individuelle est contenue dans ε .

Une autre idée a été de déterminer des modèles de référence pour des sous-populations. On peut alors modéliser \mathbf{y} de la manière suivante :

$$\mathbf{y} = \boldsymbol{\mu}_k + \varepsilon, \quad (9.10)$$

avec $\boldsymbol{\mu}_k$ correspondant à la moyenne de la sous-population k contenant l'individu \mathbf{y} . Cela correspond à une fonction $\mathbf{f}(\mathbf{x})$ constante par morceaux dans l'équation 9.8. Plusieurs techniques

probabilistes ont permis de déterminer les différents sous-groupes ainsi que leurs prototypes μ_k [BLEZEK & MILLER 2007, SABUNCU et coll. 2009]. Pour des objets de grande dimension, élaborer des algorithmes de classification peut être délicat à cause de problèmes de robustesse. Dans [SUN et coll. 2009], les auteurs proposent d'effectuer l'étape de classification dans un espace réduit obtenu par ISOMAP. Un inconvénient possible de ces méthodes est la détermination du nombre de sous-groupes à considérer.

Des modèles génératifs plus complexes ont vu le jour. Dans un premier temps des modèles linéaires ont été conçus :

$$\mathbf{y} = \boldsymbol{\mu}_y + \mathbf{W}\mathbf{x} + \boldsymbol{\varepsilon}, \quad (9.11)$$

avec \mathbf{W} une matrice $D \times d$. Les colonnes de la matrice \mathbf{W} sont formées par les premières composantes principales obtenues lors de l'ACP [ROMANIUK et coll. 2003, VIK et coll. 2005]. L'avantage des modèles génératifs linéaires obtenus par des méthodes d'ACP est l'estimation simultanée des éléments de l'espace réduit et des fonctions de régression pour passer d'un espace à l'autre. De plus, ces techniques permettent d'extraire les modes principaux des données (*i.e.* les directions où les données varient le plus). L'inconvénient de ces modèles est de ne traiter qu'une faible classe de problèmes (ceux linéaires). Plus précisément, les modèles linéaires ne permettent pas de réduire suffisamment les données, c'est-à-dire les points dans l'espace réduit ne remplissent pas tout l'espace. Il peut donc exister des zones non-admissibles dans l'espace réduit pouvant être atteintes avec les modèles linéaires. Dans [VIK 2004], Torbjørn Vik propose d'ajouter des lois de probabilités non gaussiennes aux modèles linéaires pour empêcher ces zones non-admissibles d'être atteintes. Les modèles non linéaires permettent de réduire les données de manière satisfaisante, c'est-à-dire d'avoir un espace réduit dense.

Pour traiter des modélisations de populations plus complexes, deux catégories de stratégies ont été proposées : celles basées sur des techniques d'ACP plus complexes prenant en compte la non linéarité des populations, et celles basées sur des réductions de dimensions plus complexes telles que l'ISOMAP.

Pour les méthodologies basées sur l'ACP, on peut conserver l'algorithme d'ACP tel quel, mais on ajoute des estimations probabilistes plus complexes prenant en compte la non linéarité de l'espace réduit [VIK et coll. 2007]. On peut aussi effectuer l'ACP dans des espaces non euclidiens grâce à l'utilisation de méthodes à noyaux (ACP à noyaux) [KHURD et coll. 2007] ou grâce à l'utilisation d'une métrique riemannienne (*principal geodesic component*) [FLETCHER & JOSHI 2004].

Considérons maintenant les techniques utilisant des réductions de dimension non linéaires. Dans [GERBER et coll. 2010], les auteurs proposent le même modèle que celui de l'équation 9.8.

Cette méthode comporte trois étapes : l'estimation des coordonnées des points \mathbf{x} dans l'espace réduit, puis l'estimation d'une fonction de régression \mathbf{g} telle que $\mathbf{g}(\mathbf{y}) = \mathbf{x}$, et enfin l'estimation de la fonction de régression \mathbf{f} de l'équation 9.8. Contrairement à l'ACP, les relations de passage entre l'espace initial et l'espace réduit ne sont pas obtenues durant la phase de réduction de dimension.

Le choix de la méthode de régression est donc crucial. Dans [GERBER et coll. 2010], les auteurs proposent d'utiliser des estimateurs à noyau de type Nadaraya-Watson pour les fonctions de régression :

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^N \frac{K_\sigma(d(\mathbf{x}, \mathbf{x}_i))}{\sum_{j=1}^N K_\sigma(d(\mathbf{x}, \mathbf{x}_j))} \mathbf{y}_i, \quad (9.12)$$

avec $K_\sigma(r) = \exp(-r^2/\sigma^2)$ le noyau gaussien, et $d(.,.)$ la distance dans l'espace réduit. Cette régression peut être considérée comme une espérance conditionnelle $\mathbf{f}(\mathbf{x}) = \mathbf{E}[\mathbf{Y}|\mathbf{g}(\mathbf{y}) = \mathbf{x}]$.

Les auteurs en déduisent que cette fonction de régression permet de bien généraliser la variété dans l'espace réduit.

Dans [ALJABAR et coll. 2010], les auteurs introduisent un nouveau formalisme basé sur la réduction de dimension :

$$\mathbf{y} = \mathbf{f}^{qual}(\mathbf{x}^{qual}) + \mathbf{f}^{quant}(\mathbf{x}^{quant}) + \varepsilon, \quad (9.13)$$

avec \mathbf{x}^{qual} et \mathbf{x}^{quant} respectivement des données quantitatives telles que l'intensité des voxels et qualitatives tels que l'âge ou le diamètre du crâne, et \mathbf{f}^{qual} et \mathbf{f}^{quant} des modèles génératifs respectivement pour les données quantitatives et qualitatives. Les différentes catégories de données sont traitées séparément avec des métriques spécifiques pour chacune d'entre elles. Cette étude montre la possibilité d'utiliser les modèles génératifs pour un grand nombre de types de données différentes.

2.1 Comparaison en imagerie médicale avec des représentations compactes

Les études traditionnelles effectuent des tests multivariés sur l'ensemble des coordonnées du vecteur d'attributs (par exemple un vecteur image où les attributs sont les voxels, ou un vecteur de coordonnées d'amers (*landmarks*) pour l'analyse de formes). Ces tests multivariés supposent généralement l'indépendance statistique des différentes coordonnées du vecteur. Les modèles génératifs et les représentations compactes permettent de prendre en compte des relations possibles entre ces différentes coordonnées, et de les analyser [GOLLAND et coll. 2005].

Dans une première partie, nous décrirons brièvement les principes de comparaison de groupes entre eux. Dans une deuxième partie, nous nous intéresserons spécifiquement aux problèmes liés à comparaison d'individus par rapport à un groupe contrôle.

2.2 Comparaison d'une population par rapport à une autre population

Lors de la comparaison entre deux groupes, il existe principalement deux étapes. La première étape consiste à déterminer les caractéristiques sur lesquelles s'appuie le classifieur. La deuxième étape consiste à trouver un classifieur séparant de manière optimale les deux groupes.

Dans [THOMAZ et coll. 2007, ROBINSON et coll. 2010], l'ADL est utilisée comme un classifieur linéaire sur des coordonnées réduites obtenues par ACP. L'ADL permet de projeter sur un axe les deux groupes. Cet axe correspond à la caractéristique la plus déterminante. Il est alors facile de reconstruire l'objet dans l'espace initial en effectuant la projection inverse de l'ADL puis de l'ACP pour observer les caractéristiques les plus discriminantes dans l'espace initial. Dans [KHURD et coll. 2007], les auteurs proposent de généraliser l'ADL grâce à l'utilisation de méthodes à noyaux.

Une technique bien connue pour séparer deux groupes de manière non linéaire correspond aux SVMs [VAPNIK 1999]. Ces méthodes utilisent les méthodes à noyaux. L'hyperplan séparant les deux groupes de manière linéaire est décrit dans l'espace des caractéristiques. Intuitivement, si on se déplace de manière orthogonale à cet hyperplan, on aura les caractéristiques qui diffèrent le plus entre les deux groupes. Le seul inconvénient est que cette direction orthogonale à l'hyperplan est décrite dans l'espace des caractéristiques. Dans [GOLLAND et coll. 2005], les auteurs décrivent une manière d'obtenir la direction la plus proche correspondant à cette direction orthogonale dans l'espace des caractéristiques. On ne peut pas obtenir la direction exacte car l'espace des caractéristiques est de dimension plus grande que l'espace initial, chaque élément de l'espace initial n'a pas forcément de correspondant (*i.e.* ϕ permettant de passer de l'espace

initial à l'espace des caractéristiques est injective).

La comparaison entre deux groupes est basée l'estimation d'un hyperplan séparant ces deux groupes. On peut donc s'interroger sur la manière de comparer un individu par rapport à un groupe.

2.3 Détection et analyse de données aberrantes

Détecter des données aberrantes (*outliers*) est complexe. En effet, on doit discriminer des données qui n'ont jamais été apprises par un modèle. Plus précisément, ces données peuvent avoir une très grande variabilité, beaucoup plus importante que celle de la population de référence, et sont en général très peu nombreuses. Le principe consiste donc dans l'apprentissage de la population de référence, et la comparaison du nouvel individu par rapport à celle-ci. De nombreuses méthodes ont été proposées dans la littérature. Le lecteur intéressé pourra lire les articles [MARKOU & SINGH 2003, MARKOU & SINGH 2003].

Le SVDD est une méthode souvent utilisée dans la détection de données aberrantes. Dans [SJÖSTRAND et coll. 2007], les auteurs proposent d'utiliser une méthode à noyaux, proche de l'algorithme du SVDD, permettant de décrire pour un ensemble d'objets l'objet le plus similaire à l'objet le moins similaire (voir Fig. 9.4). Cet article permet de donner une notion de rang à des objets multivariés complexes. Pour rappel, la notion de rang pour des espaces de dimension de taille supérieure à 1 est mal définie. La notion de rang est une autre manière de déterminer si un individu appartient à une population ou non.

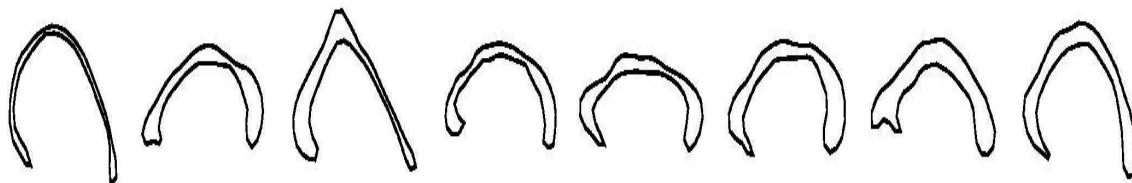


FIGURE 9.4 – De gauche à droite : des profils de corps calleux ordonnés du moins similaire au plus similaire grâce à un algorithme proche du SVDD [SJÖSTRAND et coll. 2007]. On peut remarquer que la détermination d'un ordre pour des données de ce type est complexe. Image extraite de l'article [SJÖSTRAND et coll. 2007].

Une fois les données aberrantes détectées, une deuxième étape consiste à chercher les zones considérées comme anormales. Nous proposons d'utiliser ici les modèles génératifs.

Dans [VIK et coll. 2005], les auteurs considèrent que la variable ε de l'équation 9.8 correspond à un bruit gaussien multivarié, indépendant pour chaque composante, de moyenne nulle et de variance Ψ , avec Ψ une matrice diagonale.

Supposons que les $\varepsilon_i = \mathbf{y}_i - \mathbf{f}(\mathbf{x}_i)$ correspondent aux erreurs résiduelles de l'image \mathbf{y}_i par rapport au modèle $\mathbf{f}(\mathbf{x}_i)$. Les coefficients diagonaux Ψ_i de la matrice de variance Ψ sont alors

estimés de la manière suivante :

$$\Psi = \text{diag} \left(\frac{1}{N-1} \sum_{i=1}^N \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T \right), \quad (9.14)$$

avec N le nombre d'observations et $\text{diag}(\cdot)$ l'opérateur qui met les coefficients hors-diagonale à zéro.

Les anomalies dans une nouvelle observation \mathbf{y} sont alors détectées en effectuant une comparaison de chaque composante du vecteur z -score $\mathbf{t} = \Psi^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x}))$ par rapport à un seuil. Lorsque la composante j de l'observation \mathbf{y} correspond au modèle, alors la composante j du vecteur z -score doit suivre une distribution de Student de $N - 1$ degrés de liberté. Dans le cas contraire, la composante j est supposée ne pas suivre le modèle.

Dans [GERBER et coll. 2010], les auteurs décrivent la composition \mathbf{gof} des fonctions \mathbf{f} et \mathbf{g} , permettant de passer respectivement de l'espace réduit à l'espace initial et inversement, comme un opérateur projetant sur la variété. En présence d'un nouvel individu n'appartenant pas à la population, cet individu va être projeté sur la variété dans l'espace réduit. On peut alors utiliser la modélisation proposée dans [VIK et coll. 2005] pour détecter les zones ne correspondant pas au modèle.

Dans l'article [GERBER et coll. 2010], les auteurs utilisent uniquement une population dense. Les fonctions de régression \mathbf{f} et \mathbf{g} sont estimées correctement sur un support contenant cette population. On peut s'interroger sur la capacité de généralisation de ces fonctions de régression pour des points n'appartenant pas à ce support. Plus précisément, plus un point est loin de la population de référence, plus il a de chances d'être projeté incorrectement sur la population de référence.

3. Conclusion partielle

Dans cette partie, nous avons décrit différents outils permettant de modéliser une population de manière plus ou moins complexe. Ces modélisations prennent en compte la variabilité existant dans la population. Dans un deuxième temps, nous avons montré qu'il était possible de détecter et d'analyser des données aberrantes à partir de ces modèles.

Chapitre 10

Atlas proposé

1.	Création de l’atlas	121
1.1	Utilisation d’un espace réduit	121
1.2	Estimation du support par SVDD	121
1.3	Estimation de la régression f	128
2.	Utilisation de l’atlas	129
2.1	Réduction de dimension d’un nouvel individu	130
2.2	Étape de détection	130
2.3	Étape d’analyse	131
3.	Applications	133
3.1	Application en détection	133
3.2	Application en analyse	135
4.	Conclusion partielle	135

Dans ce chapitre, nous présentons le modèle original d’atlas élaboré dans le cadre de cette thèse. Nous le mettrons en œuvre pour la comparaison d’un individu par rapport à un groupe, dans le but de fournir une aide au diagnostic. Plus précisément, nous voulons créer un modèle statistique compact permettant de rendre compte des données de référence et de détecter les données aberrantes. Nous utilisons des représentations compactes pour ne pas nous exposer à des problèmes liés à la malédiction de la dimension [BELLMAN 1961].

Dans un premier temps, ce modèle statistique doit être capable de déterminer si un nouveau jeu de données est aberrant ou non par rapport aux données de référence, et dans un deuxième temps, de localiser les composantes anormales des données considérées comme aberrantes.

Pour cela, il existe deux grandes phases : une phase de conception de l’atlas et une phase d’utilisation de l’atlas.

Pour la phase de conception, on procède en trois étapes principales : la réduction de dimension pour obtenir un modèle compact, l’apprentissage du support de la distribution dans l’espace réduit, et l’estimation de la régression f pour passer de l’espace réduit à l’espace initial.

Nous utiliserons l’algorithme ISOMAP [TENENBAUM et coll. 2000] pour la réduction de dimension. Cette réduction de dimension nous permet d’obtenir une représentation compacte des données.

Pour l’apprentissage du support de la distribution, nous allons utiliser la méthode du SVDD [TAX & DUIN 1999, TAX & DUIN 2004] avec des noyaux gaussiens (voir la section 1.2). Le SVDD dépend fortement de la paramétrisation du noyau, et plus particulièrement du coefficient σ (voir éq. 9.3). Nous proposons dans un premier temps de nouvelles méthodes pour la sélection

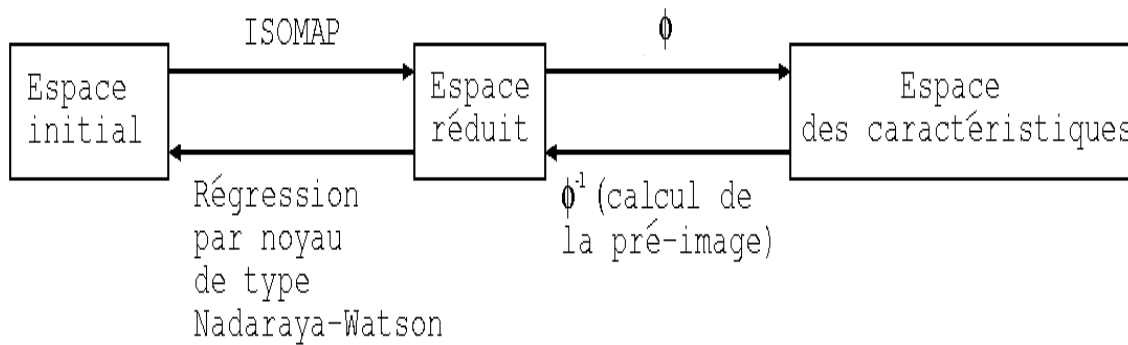


FIGURE 10.1 – La création de l’atlas implique plusieurs étapes et deux espaces complémentaires : l’espace réduit pour une modélisation compacte des données de référence et l’espace des caractéristiques pour la détermination du support des données de référence.

automatique de ce paramètre σ . Dans un deuxième temps, nous montrerons que l’utilisation de l’espace réduit améliore l’estimation du support de la distribution.

Pour estimer la fonction f de régression permettant de passer de l’espace réduit à l’espace initial, nous utiliserons des fonctions à noyaux de type Nadaraya-Watson.

Dans une deuxième phase, nous décrirons l’utilisation de l’atlas pour déterminer si un nouvel individu est anormal par rapport à la population de référence, et pour identifier les zones correspondant à cette anomalie.

Plus précisément, pour détecter si un individu est anormal, nous proposons un modèle discriminatif basé sur le SVDD. Ce modèle nous permet d’obtenir une règle de décision simple.

Pour identifier les zones anormales, nous proposons un modèle génératif basé aussi sur le SVDD. Ce modèle statistique peut être formalisé de la manière suivante :

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) + \varepsilon, \quad (10.1)$$

avec \mathbf{y} les données observées dans l’espace initial, \mathbf{x} les variables correspondantes dans l’espace réduit, \mathbf{f} la fonction modélisant le lien entre les espaces réduit et initial, et ε un bruit supposé gaussien. Le principe est que si un nouvel individu \mathbf{y} appartient à la population de référence, alors cet individu \mathbf{y} pourra être expliqué de manière satisfaisante par $\mathbf{f}(\mathbf{x})$, \mathbf{x} appartenant au support estimé précédemment. Par conséquent, l’erreur ε au modèle suivra la même loi gaussienne que les erreurs obtenues pour la population de référence. Par contre, si le nouvel individu \mathbf{y} n’est pas décrit par le modèle, sa représentation dans l’espace réduit \mathbf{x} ne réside pas sur le support de la population de référence ou la différence ε observée entre l’individu \mathbf{y} et le modèle \mathbf{f} est nettement supérieure¹ aux différences observées pour la population témoin.

Cet atlas utilise deux espaces complémentaires successivement : l’espace réduit et l’espace F des caractéristiques. L’espace réduit permet d’obtenir une représentation compacte, tandis que l’espace F des caractéristiques permet d’apprendre des supports d’aspect complexe. La figure 10.1 représente les différents espaces utilisés.

L’intérêt de cet atlas est de pouvoir développer à partir d’un même algorithme, le SVDD, deux modèles, discriminatif et génératif, complémentaires.

1. Il existe des cas où la différence ε observée est faible tout en ayant \mathbf{y} n’appartenant pas au modèle.

1. Création de l'atlas

Dans un premier temps, nous expliquerons l'utilisation de l'espace réduit. Dans un deuxième temps, nous estimerons le support de la distribution des points dans l'espace réduit. Pour cela, nous utilisons l'algorithme SVDD. Nous nous intéresserons plus particulièrement à la paramétrisation du SVDD pour la rendre automatique. Finalement, nous décrirons l'estimation de la fonction de régression f permettant de passer de l'espace réduit à l'espace initial.

1.1 Utilisation d'un espace réduit

Les données utilisées lors de la construction d'atlas appartiennent généralement à des espaces de grande dimension. Dans ces espaces, les données sont isolées, ce qui rend délicates de nombreuses applications tels que la classification, le partitionnement (*clustering*) ou la détection de données aberrantes. Par conséquent, nous effectuons une réduction de dimension pour obtenir des données denses dans l'espace réduit. Le choix de la dimension va dépendre de deux critères : l'erreur de reconstruction et le nombre d'individus lors de l'apprentissage. Plus précisément, l'erreur de reconstruction correspond à l'erreur quadratique $\|\epsilon\|^2$ entre le vecteur observé et le vecteur reconstruit après le passage dans un espace réduit de dimension d [TENENBAUM et coll. 2000]. Plus la dimension d est petite, plus l'erreur de reconstruction est grande.

Cependant, étudier l'erreur de reconstruction nécessite d'estimer une fonction de régression. En pratique, les auteurs de l'article [TENENBAUM et coll. 2000] ont implémenté une autre technique basée sur les distances. Ils étudient le coefficient C :

$$C = 1 - \text{corr}(d, D), \quad (10.2)$$

où $\text{corr}(d, D)$ est la corrélation des distances D inter-individus dans l'espace initial et des distances d inter-individus dans l'espace réduit. Plus les distances dans l'espace réduit sont proches des distances dans l'espace initial, plus le coefficient C est faible.

Le nombre d'individus est en général limité lors de la phase d'apprentissage. Ceci nous contraint à choisir des espaces de petite dimension (trois ou quatre) pour obtenir des données denses. Ce choix concorde avec les travaux de l'article [GERBER et coll. 2010]. Nous devons donc faire un compromis entre une faible erreur de reconstruction et un espace rempli de façon dense.

Par opposition aux travaux [GERBER et coll. 2010], nous décidons d'apprendre conjointement les individus de la base d'apprentissage et ceux à tester. Les auteurs de [GERBER et coll. 2010] cherchent à modéliser une population de référence d'images de cerveaux en utilisant un espace réduit. Contrairement à nous, il n'existe pas d'individu pathologique dans leur étude.

1.2 Estimation du support par SVDD

L'estimation du support de la distribution des points dans l'espace réduit est une phase cruciale de l'élaboration de l'atlas proposé dans cette thèse. Nous choisissons d'effectuer cette estimation par SVDD. Le but de cette partie est de proposer une paramétrisation automatique du SVDD. La paramétrisation automatique du SVDD est décrite à la fin de l'article [SCHÖLKOPF et coll. 2001] comme une étape essentielle du SVDD. Cette étude de la paramétrisation automatique s'inscrit dans un cadre beaucoup plus général que celui de la création d'atlas.

L'hypothèse principale est que tous les individus appartenant à la population de référence doivent résider à l'intérieur du support. Nous présentons Fig. 10.2 un exemple montrant l'im-

portance de l'estimation du support de la distribution de la population de référence. On peut

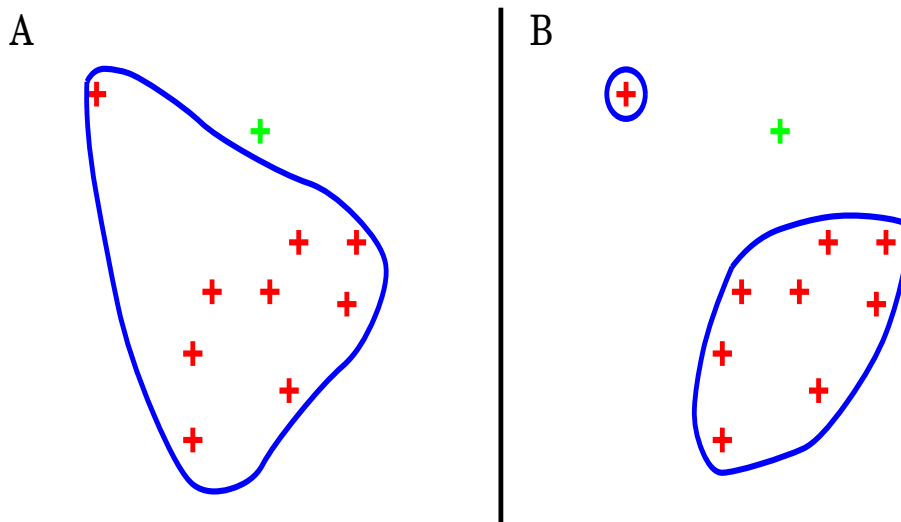


FIGURE 10.2 – Schéma représentant l'estimation de deux supports différents (visualisés en bleu) pour le même jeu de données d'apprentissage (individus en rouge). On observe un individu de la base d'apprentissage très différent des autres individus (l'individu en haut à gauche). Selon la manière d'estimer le support, un nouvel individu, représenté ici en vert, sera proche (voir le schéma A) ou éloigné (voir schéma B) du support de la distribution de la population de référence.

remarquer que dans le cadre des SVMs [VAPNIK 1999], l'estimation automatique des paramètres peut être effectuée par validation croisée ou par maximisation de la marge entre la frontière et les deux groupes [CHAPELLE et coll. 2002]. Dans le cas du SVDD sans inclusion de données aberrantes, ces stratégies ne peuvent pas être mises en œuvre. Plus précisément, la validation croisée est une méthode séparant les données en un échantillon d'apprentissage et un échantillon de test. Généralement, les deux échantillons sont composés conjointement de données appartenant et n'appartenant pas à la population de référence. Cependant, dans le cas de la détection de données aberrantes, les données aberrantes sont rares. De plus, la variabilité des données aberrantes est beaucoup plus grande que celle de la population de référence. On ne peut donc pas utiliser les données aberrantes. Considérons maintenant uniquement la population de référence pour effectuer une validation croisée. La validation croisée estime un paramètre donnant un support très grand comportant toutes les données de l'échantillon test. Ce support ne correspondra pas à l'allure de la distribution de la population de référence. De plus, la frontière dépend uniquement des vecteurs supports, donc ôter des individus qui ne sont pas sur la frontière n'affectera pas celle-ci [BRERETON 2011].

Rappelons ici la formule du noyau gaussien²

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right).$$

Soit \mathbf{K} la matrice ayant pour coefficient $K(\mathbf{x}_i, \mathbf{x}_j)$ pour la i^{eme} ligne et la j^{eme} colonne. Le coefficient σ est crucial dans l'estimation du support par SVDD [BRERETON & LLOYD 2010]. Il va fortement influencer la géométrie du support de la distribution ainsi que le nombre d'éléments

2. Cette formule est extraite des articles [TAX & DUIN 1999]. On peut aussi trouver le noyau gaussien formulé $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$ [EVANGELISTA et coll. 2007] ou $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma}\right)$ [SCHÖLKOPF et coll. 2001].

sur la frontière de l'hyperboule. Ces éléments sont dénommés vecteurs-support (*support vector*).

Prenons deux exemples extrêmes pour bien comprendre l'importance de ce coefficient σ . Supposons σ très proche de zéro. Tous les coefficients hors diagonale de \mathbf{K} vont être égaux à zéro. Dans [TAX & DUIN 1999, TAX & DUIN 2004], les auteurs montrent que tous les individus sont à une même distance $1 - \frac{1}{N}$ du centre de l'hyperboule B_F . Plus précisément, tous les éléments seront situés sur la frontière de l'hyperboule B_F .

Le deuxième cas correspond à choisir σ très grand par rapport au maximum des distances entre les éléments. Dans ce cas-là, la matrice \mathbf{K} contient uniquement des uns. Tous les individus sont à une distance nulle les uns des autres dans l'espace F des caractéristiques. Dans [TAX & DUIN 1999, TAX & DUIN 2004], les auteurs montrent que cela correspond à l'utilisation d'un noyau linéaire (ils effectuent un développement limité de l'exponentielle au voisinage de zéro). On a alors dans ce cas-ci une hyperboule euclidienne englobant les données dans l'espace F des caractéristiques. Seuls quelques éléments se situent sur la frontière de l'hyperboule B_F .

Ces problèmes sont visualisés pour un cas réel Fig. 10.3. Dans cette figure, on fait varier le coefficient σ entre zéro et un nombre très supérieur au maximum des distances inter-éléments. On observe la variation du support de la distribution, qui est d'abord centré sur chaque point et qui tend à devenir une boule dans cet exemple en deux dimensions.

Nous proposons trois nouvelles méthodes d'estimation automatique de σ et les comparons à la méthode de référence des articles [TAX & DUIN 1999, TAX & DUIN 2004].

Estimation de σ basée sur le nombre de vecteurs-support

Dans [TAX & DUIN 1999, TAX & DUIN 2004], les auteurs définissent σ de manière à ce que le nombre de vecteurs-support soit fixe. Cette stratégie sera notée σ_{Nbs} . L'avantage de cette méthode est sa simplicité. L'inconvénient est qu'elle ne dépend pas de la géométrie de la distribution. Plus précisément, si le support de la distribution est simple, cela nécessitera peu de vecteurs-support. Par contre, si la distribution a un support complexe, cela nécessitera un plus grand nombre de vecteurs. Dans [TAX & DUIN 1999, TAX & DUIN 2004], les auteurs utilisent un seuil fixe de 10% de vecteurs-support.

Estimation de σ basée sur les distances des k -plus proches voisins

Dans [TAX & DUIN 1999, TAX & DUIN 2004], les auteurs décrivent l'importance de la relation entre le coefficient σ et les distances inter-individus pour les individus appartenant à la base d'apprentissage. Plus précisément les auteurs indiquent que σ doit être compris entre la plus petite et la plus grande distance inter-individus.

Les distances des k -plus proches voisins sont représentatifs de la densité locale. Dans la section 5 de l'article [CHANDOLA et coll. 2009], les auteurs considèrent que les données supposées normales apparaissent dans des voisinages locaux denses, tandis que les anomalies apparaissent loin de leurs plus proches voisins. Cette densité locale est si importante qu'un algorithme modifié du SVDD la prend en compte dans l'estimation du support [LEE & VERLEYSSEN 2007].

Le but est d'estimer le coefficient σ comme un seuil séparant les données appartenant à des voisinages denses des données isolées. Pour cela, nous calculons l'histogramme des distances pour les k -plus proches voisins. Nous modélisons ensuite cet histogramme par une loi Gamma, et nous fixons σ pour ne conserver que $N\%$ de l'aire totale. Dans le reste du mémoire, cette valeur N sera fixée à 99.

Une variable aléatoire X suit une loi Gamma de paramètres k et θ strictement positifs,

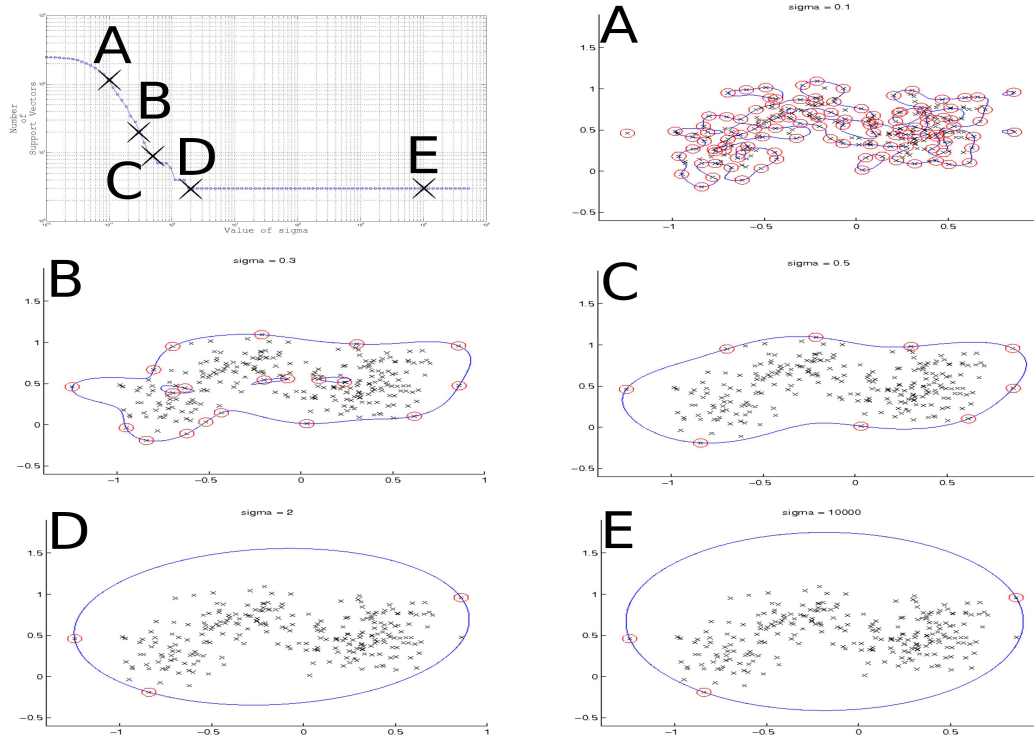


FIGURE 10.3 – En haut à gauche : courbe représentant le nombre de vecteurs-support en fonction du coefficient σ . A, B, C, D, E : les différentes représentations du support pour des coefficients σ croissants. Les individus sont modélisés par des croix. Les vecteurs-support sont entourés en rouge. La pré-image de l'hyperboule est modélisée en bleu. On observe que pour les σ petits, beaucoup de points sont sur la frontière, le support de la distribution correspond à l'union des voisinages locaux de chaque individu. Un nouvel individu a peu de chance d'appartenir au support. On a donc une faible capacité de généralisation du modèle. Pour les σ grands, peu d'individus sont sur la frontière. Le support de la distribution est très grand par rapport à la distribution des individus. Un nouvel individu, même éloigné, a des chances d'appartenir au support. On a ici une trop grande capacité de généralisation.

notée aussi $X \sim \Gamma(k, \theta)$, si sa fonction de densité de probabilité se met sous la forme :

$$f(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k) \theta^k} \mathbb{I}_{(x \geq 0)}, \quad (10.3)$$

où $\mathbb{I}_{(\cdot)}$ est la fonction indicatrice. Les paramètres k et θ de la loi Gamma sont reliés à la moyenne m et à la variance σ^2 par les relations suivantes :

$$m = k\theta, \quad (10.4)$$

$$\sigma^2 = k\theta^2. \quad (10.5)$$

Etant donné que nous supposons la présence possible de données aberrantes, nous effectuons une estimation robuste des paramètres de la loi Gamma. Pour cela, nous estimons moyenne m et variance σ^2 respectivement par la médiane et la *median absolute deviation* (MAD) au lieu des estimations empiriques classiques de la moyenne et de la variance. Plus précisément, le MAD peut être utilisé comme un estimateur consistant de l'écart type avec la relation suivante : $\hat{\sigma} = 1.4826 \text{ MAD}$.

Le coefficient σ obtenu sera noté σ_{GL} . L'avantage de cette méthode est la prise en compte du voisinage local. L'inconvénient est de devoir choisir un seuil N . Nous pouvons observer l'estimation de σ_{GL} en Fig. 10.4.

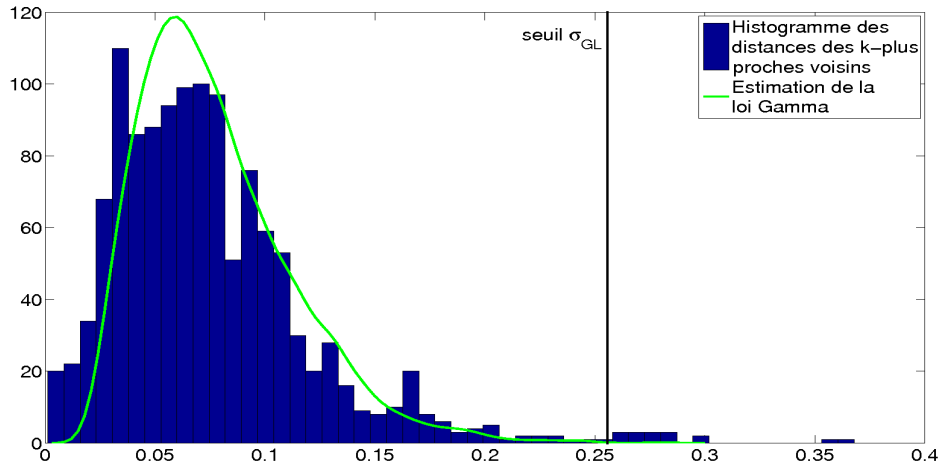


FIGURE 10.4 – En bleu : l’histogramme des distances des k -plus proches voisins. En vert : l’estimée de la loi Gamma. Le seuil σ_{GL} est symbolisé par une barre verticale.

La deuxième idée pour séparer l’histogramme en deux parties de manière automatique est l’utilisation d’un algorithme traditionnel de seuillage [OTSU 1979]. L’algorithme original est utilisé pour effectuer un seuillage automatique d’une image à niveaux de gris en une image binaire à partir de l’histogramme des intensités des pixels. Le but est de séparer l’image en deux classes (le premier plan et l’arrière-plan) Pour cela, on calcule le seuil optimal qui sépare ces deux classes afin que leur variance intra-classe soit minimale. On dénommera σ_{OTSU} le coefficient obtenu par cette méthode. L’avantage de cette méthode est d’être automatique.

L’inconvénient est que les deux classes de l’histogramme n’ont pas le même poids, ce qui peut entraîner des estimations incorrectes. Plus précisément, il y a beaucoup plus de données normales denses que de données aberrantes par définition. On peut visualiser ce phénomène en Fig. 10.5.

Estimation de σ basée sur la distribution des distances

On observe que lorsque σ est petit, tous les individus sont sur la frontière du support, ce qui constitue un cas critique. Idéalement on souhaite avoir peu d’individus sur la frontière. Les autres individus sont situés à l’intérieur du support de manière plus ou moins éloignée de la frontière. On suppose qu’un bon critère serait que les distances d_i des individus au centre de l’hyperboule (calculées selon l’éq. 9.7) soient distribuées selon une loi uniforme. On estime donc la distribution des distances. Pour cela, on utilise une estimation par noyau (méthode de Parzen-Rosenblatt). L’estimateur non-paramétrique de la densité par la méthode du noyau peut s’écrire :

$$\hat{f}_h(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right), \quad (10.6)$$

où K est un noyau et h un paramètre nommé largeur de bande (*bandwidth*). Le noyau K choisi est gaussien : $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.

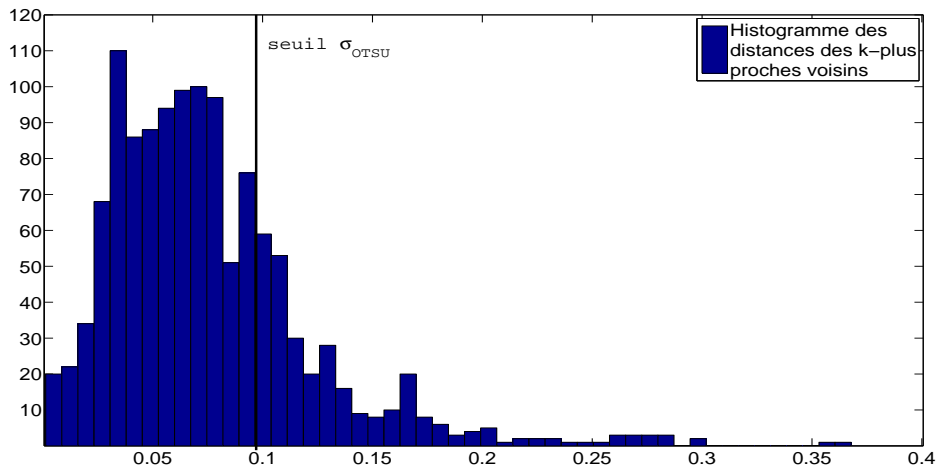


FIGURE 10.5 – En bleu : l’histogramme des distances des k -plus proches voisins. Le seuil σ_{OTSU} est symbolisé par une barre verticale. On observe deux modes dans l’histogramme. Le premier mode correspond aux données proches les unes des autres, le deuxième mode correspond aux données éloignées. On remarque que le deuxième mode est moins peuplé que le premier. Le seuil σ_{OTSU} sous-estime le seuil réel.

On calcule ensuite la distance de Kullback-Leibler D_{KL} entre cette distribution et la distribution uniforme. La distance de Kullback-Leibler D_{KL} est une mesure de dissimilarité entre deux distributions P et Q , pouvant s’écrire sous la forme suivante :

$$D_{KL}(P\|Q) = \int P(x) \log \frac{P(x)}{Q(x)} dx. \quad (10.7)$$

En considérant Q uniforme, on peut reformuler l’équation précédente :

$$D_{KL}(P\|Q) = \int P(x) \log(P(x)) dx + C, \quad (10.8)$$

avec C une constante. On remarque cette dernière équation correspond à l’opposé de l’entropie de P .

On choisit σ pour que cette distance soit minimale. On notera σ_{KL} le résultat obtenu par cette méthode. L’avantage de cette méthode est qu’elle est totalement automatique. L’inconvénient est que la distribution des distances inter-individus ne suit pas une loi uniforme en pratique. Cependant, ce critère pénalise fortement le cas critique où tous les individus sont sur la frontière.

Simulations

Nous proposons d’évaluer ces quatre stratégies sur trois différents exemples simulés.

La première simulation consiste à étudier un support en forme d’anneau pour montrer que les différentes estimations sont correctes en présence d’une distribution complexe. Les quatre supports obtenus peuvent être visualisés Fig. 10.6. Ils correspondent bien au support réel.

En imagerie médicale, et dans les cas pratiques réels, la population supposée de référence peut contenir des données aberrantes. Nous allons étudier les propriétés des quatre estimations en présence de données aberrantes. Nous considérons un support en forme de banane (voir

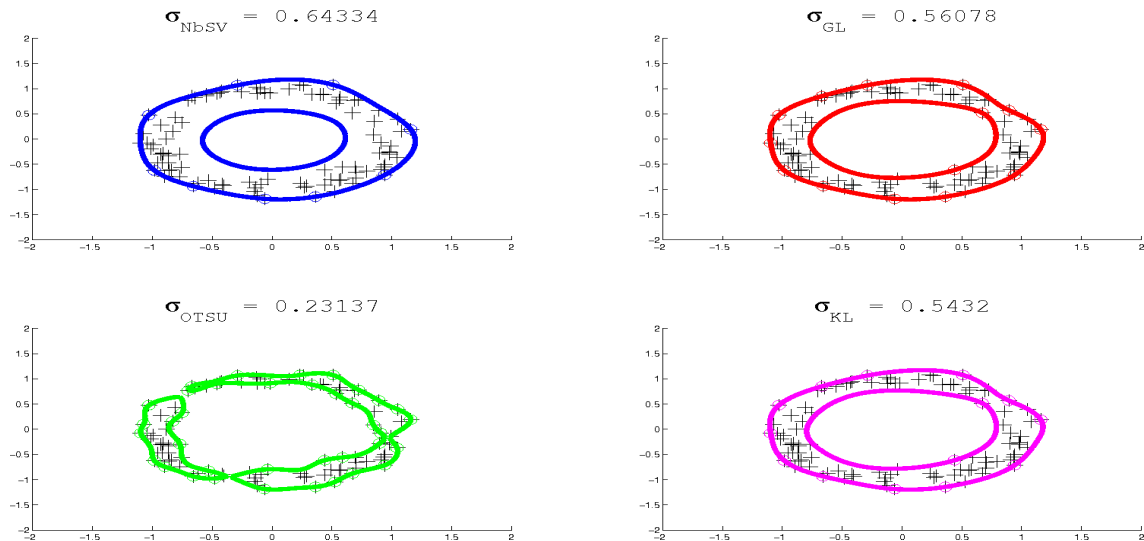


FIGURE 10.6 – Quatre supports obtenus avec quatre estimations de σ sur un support en forme d'anneau. On observe que les quatre méthodes correspondent bien à la distribution réelle. On peut remarquer que σ_{OTSU} donne le support le plus restreint.

Fig. 10.7). On peut observer que les quatre supports correspondent bien au support réel de la distribution. Nous ajoutons 50% de données aberrantes, ce qui correspond à un cas vraiment extrême (voir Fig. 10.8). Nous observons que les stratégies basées sur les distances des k -plus proches voisins (σ_{GL} et σ_{OTSU}) donnent les meilleurs résultats. Nous remarquons que la méthode proposée dans [TAX & DUIN 1999, TAX & DUIN 2004] est très sensible aux données aberrantes.

Un dernier test a été la comparaison de l'évolution des coefficients σ_{Nbs} et σ_{GL} lorsqu'on rajoute progressivement des données aberrantes (voir Fig. 10.9). Nous observons que σ_{Nbs} varie très rapidement alors que σ_{GL} est peu affecté par ces données aberrantes.

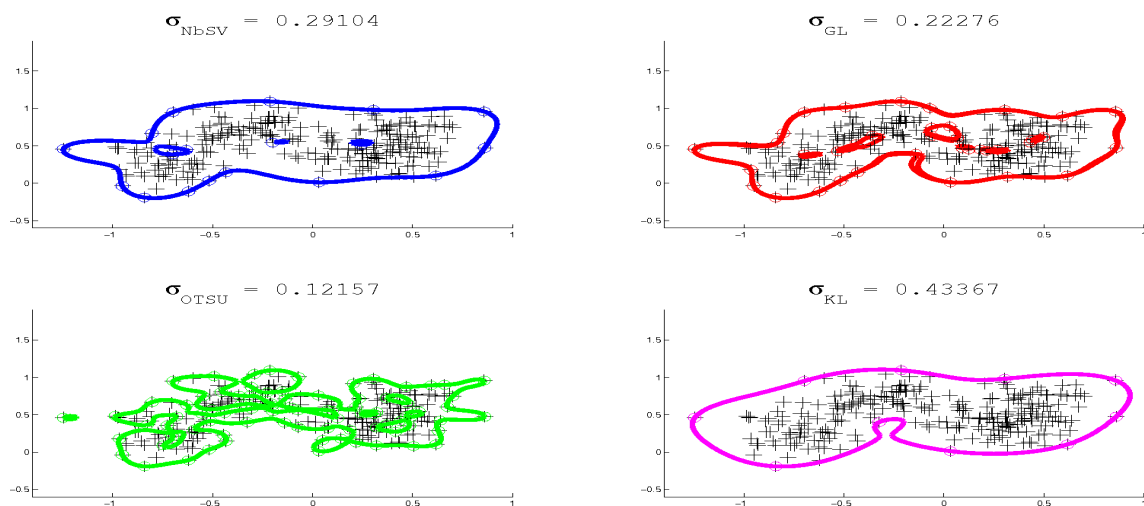


FIGURE 10.7 – Quatre supports obtenus avec quatre estimations de σ sur un support en forme de banane. On observe que les quatre stratégies conduisent à des supports plus ou moins grands.

En conclusion, nous constatons que l'estimation du coefficient σ est très dépendante des données aberrantes présentes dans le jeu de données d'apprentissage. Dans la suite du mémoire,

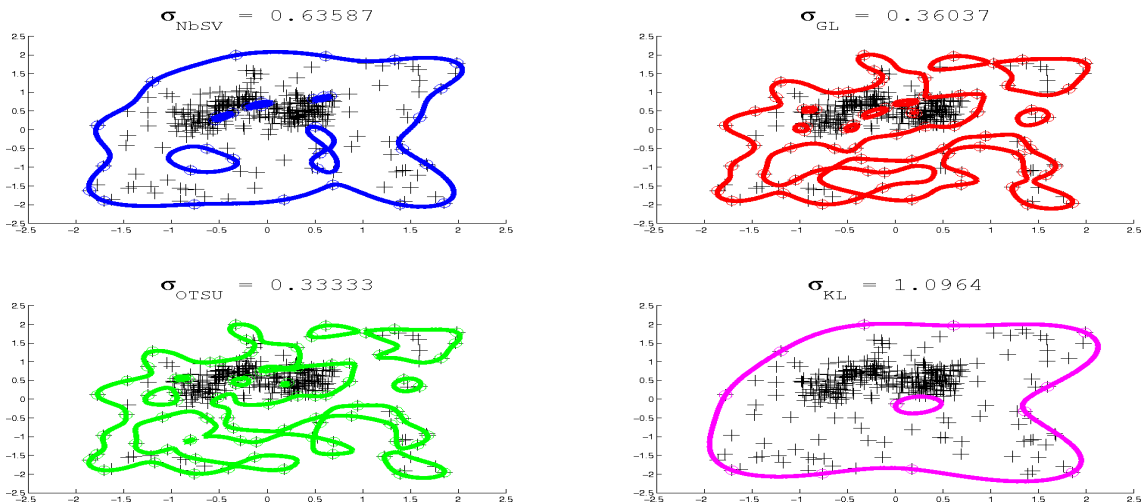


FIGURE 10.8 – Quatre supports obtenus avec quatre estimations de σ sur un support en forme de banane auquel on a rajouté 50% de données aberrantes. On observe que les stratégies GL et OTSU arrivent bien à séparer la distribution banane des données aberrantes.

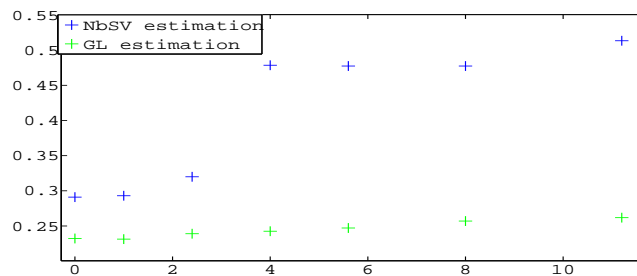


FIGURE 10.9 – Graphique représentant la valeur estimée du coefficient σ pour la stratégie NbS (en bleu, courbe du haut) et celle GL (en vert, courbe du bas) en fonction du pourcentage de données aberrantes introduites. La population de référence contient 250 individus. On peut observer que la stratégie NbS est fortement impactée par le nombre de données aberrantes, contrairement à la stratégie GL.

nous paramètrons le SVDD avec le coefficient σ_{GL} . D'une part, ce coefficient est peu sensible à la présence de données aberrantes, et d'autre part, avec ce coefficient, on obtient des supports plus grands qu'avec la méthodologie Otsu. En effet, avoir des supports de petite taille peut être un inconvénient. Plus précisément, si le support est centré sur les données d'apprentissage (le support n'est plus connexe et il existe une composante pour chaque donnée d'apprentissage), le modèle construit sera peu généralisable. Un nouvel individu non pathologique devra obligatoirement correspondre à un individu déjà présent dans la base d'apprentissage. Si le support est plus grand (le support est connexe), un nouvel individu correspondra à une sous-partie de la population.

1.3 Estimation de la régression f

Nous utilisons une technique de régression à noyaux de type Nadaraya-Watson (NNW) pour passer de l'espace réduit à l'espace initial. L'utilisation de la régression NNW est intéressante car elle ne laisse pas de trous alors que les régressions de type *radial basis functions* (RBFs) en

laissent (voir Fig. 10.10). Cette propriété est intéressante, car lorsque le nombre d'individus de la base d'apprentissage est insuffisant pour obtenir un espace réduit rempli de manière dense, des trous peuvent apparaître dans le support si on utilise la régression par RBFs.

Rappelons ici l'équation correspondant à la régression NNW :

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^N \frac{K_\sigma(d(\mathbf{x}, \mathbf{x}_i))}{\sum_{j=1}^N K_\sigma(d(\mathbf{x}, \mathbf{x}_j))} \mathbf{y}_i, \quad (10.9)$$

avec $K_\sigma(r) = \exp(-r^2/\sigma^2)$ le noyau gaussien, et $d(.,.)$ la distance dans l'espace réduit. Ecrivons maintenant l'équation correspondant à la régression de type RBFs :

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^N w_i K_\sigma(\|\mathbf{x} - \mathbf{x}_i\|), \quad (10.10)$$

avec w_i un poids associé à \mathbf{x}_i , donnant plus ou moins d'importance au noyau K_σ associé. On estime ces poids w_i en minimisant la fonction \mathcal{C} de coût suivante :

$$\mathcal{C}(w_1, \dots, w_N) = \sum_{i=1}^N \|\mathbf{y}_i - \mathbf{f}(\mathbf{x}_i)\|^2. \quad (10.11)$$

Un avantage de la régression NNW est qu'elle ne nécessite qu'un coefficient σ (voir éq. 10.9) pouvant être estimé par une technique de validation croisée. Pour la régression de type RBFs, il faudrait faire dépendre le coefficient σ de l'individu \mathbf{x}_i de la base d'apprentissage afin d'éviter d'avoir des trous.

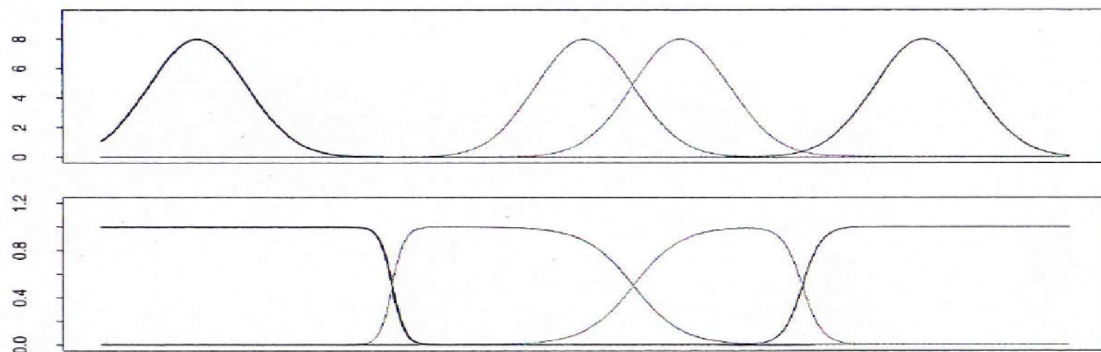


FIGURE 10.10 – Certaines régressions, telles que celles utilisant les RBFs, peuvent laisser des supports avec des trous. La régression à noyaux de type Nadaraya-Watson permet d'éviter ces problèmes. En haut : RBFs gaussiennes dans \mathbb{R} avec des largeurs de bande constantes. En bas : fonctions de base de type NNW. Figure extraite du livre [HASTIE et coll. 2009].

2. Utilisation de l'atlas

Dans la section précédente, nous avons proposé une méthodologie pour la conception d'un atlas. Nous proposons dans cette section deux applications. La première application consiste à déterminer si un individu appartient à la population de référence ou non (étape de détection). La deuxième application consiste à déterminer les régions dans lesquelles les différences sont observées (étape d'analyse). Ces deux applications utilisent la réduction de dimension. Dans un premier temps, nous décrirons la manière d'effectuer la réduction de dimension. Dans un deuxième temps, nous présenterons les deux applications. L'ensemble de la méthodologie est schématisé Fig. 10.11.

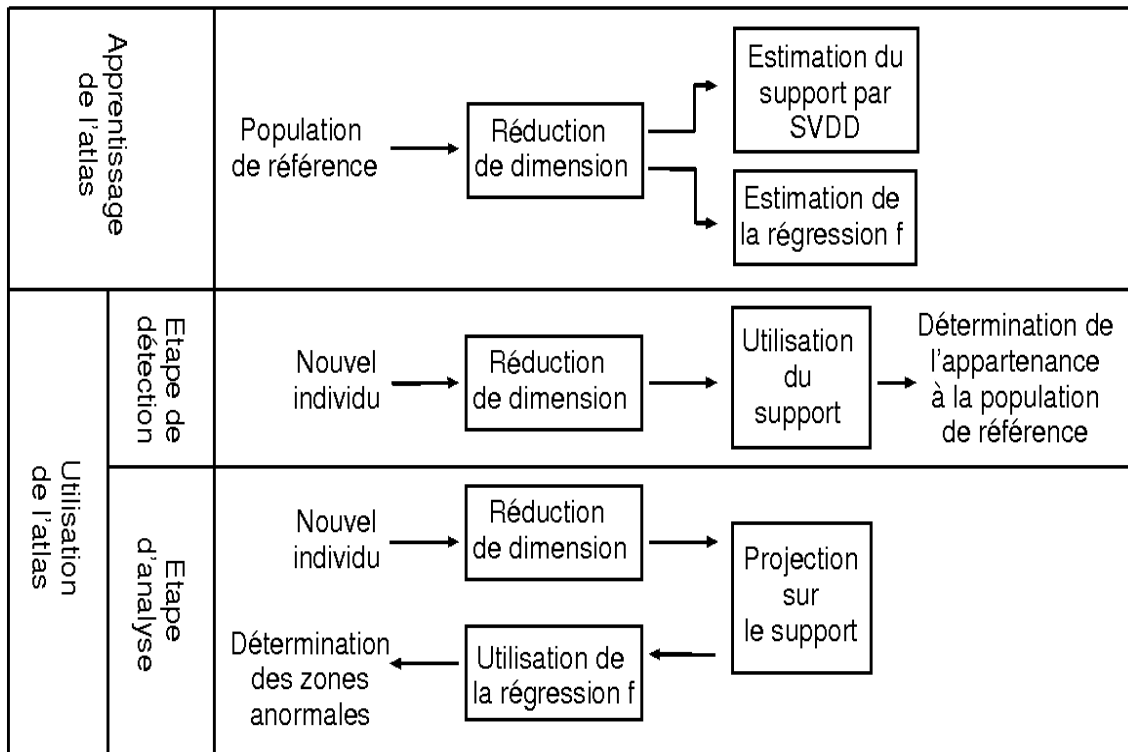


FIGURE 10.11 – Schéma représentant l'élaboration de l'atlas ainsi que son utilisation. Il existe principalement deux phases distinctes. La première phase consiste en l'apprentissage de l'atlas, et la seconde en l'utilisation de l'atlas. Dans chaque cas, une réduction de dimension est effectuée pour obtenir une représentation compacte. Le support estimé par le SVDD dans la phase d'apprentissage sera utilisé dans les deux étapes de détection et d'analyse. On remarquera aussi que l'étape de détection s'effectue dans l'espace réduit, tandis que l'étape d'analyse s'effectue dans l'espace initial, nécessitant l'utilisation d'une fonction f de régression estimée préalablement durant la phase d'apprentissage.

2.1 Réduction de dimension d'un nouvel individu

Si on effectuait la réduction de dimension ISOMAP pour chaque nouvel individu avec la population de référence, la position \mathbf{x}_i des individus de référence serait différente à chaque réduction de dimension. Le support par SVDD et la fonction de régression f appris sur les anciennes valeurs de \mathbf{x}_i ne correspondraient plus aux nouvelles valeurs.

Pour éviter ce problème, nous effectuons l'algorithme ISOMAP en contraignant les individus \mathbf{x}_i de la population de référence à rester à une position identique à celle de la phase d'apprentissage. Les paramètres à optimiser durant l'ISOMAP correspondent uniquement à la position du nouvel individu.

2.2 Étape de détection

L'étape de détection est directe grâce à l'algorithme SVDD. Cet algorithme peut être considéré comme un modèle discriminatif, au sens où il permet de discriminer un individu par rapport à un groupe de référence. Plus précisément, la distance d_{obs} d'un nouvel individu \mathbf{y}_{obs} au centre \mathbf{a}_F de l'hyperboule B_F est calculée (voir éq. 9.7). Si cette distance est plus grande que le rayon R_F de l'hyperboule B_F , alors le nouvel individu est supposé ne pas appartenir à la population de référence. On peut aussi considérer que plus cette distance est grande, plus le nouvel individu

est différent de l'ensemble d'apprentissage. Cependant, il est difficile d'interpréter directement cette distance d_{obs} car, pour chaque ensemble d'apprentissage, le rayon R_F varie. On considère alors une distance normalisée, notée d_{obs}^{norm} , correspondant au rapport de d_{obs} par R_F . De cette manière, la règle de décision s'écrit de la manière suivante :

$$\text{individu} = \begin{cases} \text{normal} & \text{si } d_{obs}^{norm} \leq 1, \\ \text{anormal} & \text{si } d_{obs}^{norm} > 1. \end{cases} \quad (10.12)$$

En pratique, considérer la distance d_{obs} ou la distance d_{obs}^{norm} ne modifie pas le résultat de la détection. Une fois l'étape de détection effectuée, il est intéressant de déterminer quelles sont les zones considérées comme anormales.

2.3 Étape d'analyse

L'étape d'analyse est basée sur un modèle génératif. Tout d'abord, nous estimons la représentation \mathbf{x}_{obs} dans l'espace réduit de l'individu \mathbf{y}_{obs} à tester grâce à l'ISOMAP. L'algorithme ISOMAP est effectué simultanément avec les individus de la population de référence. Si le nouvel individu \mathbf{x}_{obs} est détecté comme anormal, nous le projetons sur la frontière du support appris grâce au SVDD. A partir du nouvel individu projeté \mathbf{x}_{proj} , nous générons un individu $\mathbf{f}(\mathbf{x}_{proj})$, \mathbf{f} étant préalablement appris à partir des données de la population de référence. Cet individu $\mathbf{f}(\mathbf{x}_{proj})$ généré (noté aussi \mathbf{y}_{proj}) représente la partie de \mathbf{y}_{obs} expliquée par le modèle. Ensuite nous calculons l'écart ε_{obs} entre le modèle \mathbf{y}_{proj} et le nouvel individu \mathbf{y}_{obs} . Si cet écart est comparable aux écarts obtenus avec la population de référence, alors on considérera que l'individu appartient à cette population. Si cet écart est supérieur en valeur absolue aux écarts de référence, alors l'individu n'a pu être expliqué par le modèle \mathbf{y}_{proj} . Plus précisément, on effectuera un test d'hypothèse pour déterminer si l'écart correspond ou non aux écarts de référence.

L'étape d'analyse se compose de quatre parties :

- étape 1 : si l'individu observé $\phi(\mathbf{x}_{obs})$ est en dehors de l'hyperboule B_F , projeter l'individu sur la frontière de l'hyperboule B_F en $\phi(\mathbf{x}_{proj})$ (cette action est effectuée dans l'espace F des caractéristiques) ;
- étape 2 : déterminer la pré-image (*i.e.* l'image dans l'espace initial de la projection) de \mathbf{x}_{proj} ;
- étape 3 : générer l'individu \mathbf{y}_{proj} à partir de \mathbf{x}_{proj} grâce à \mathbf{f} préalablement appris ;
- étape 4 : tester l'erreur $\varepsilon = \mathbf{y}_{obs} - \mathbf{y}_{proj}$ par rapport à l'erreur estimée sur la population d'apprentissage (voir [VIK et coll. 2007] et section 2.3).

Décrivons maintenant en détail ces quatre étapes. La première étape correspond à l'étape de projection. Cette étape de projection a tout d'abord été conçue pour débruiter [PARK et coll. 2007]. Nous pouvons concevoir que la zone anormale correspond à un bruit plus fort que le bruit de la population de référence. Détecter la zone anormale revient alors à détecter ce surplus de bruit.

Dans l'espace initial³, le support de la distribution peut être complexe. Projeter les données sur ce support peut être délicat. L'idée est de se placer dans l'espace F des caractéristiques, où le support est décrit par une hyperboule B_F plus simple que le support dans l'espace initial. La fonction ϕ , permettant de travailler dans l'espace F des caractéristiques, n'est connue qu'au

3. L'espace initial peut être l'espace où les individus sont observés ou l'espace réduit. La grande différence entre ces deux espaces est la robustesse dans l'estimation du support par le SVDD. Si l'espace initial correspond à l'espace où les individus sont observés alors l'étape 3 de régression n'est pas nécessaire.

travers de la fonction K . C'est pourquoi, lorsque nous effectuons des calculs dans l'espace F des caractéristiques, nous ne considérons que les produits scalaires des vecteurs $\phi(\mathbf{x}_i)$ entre eux.

La première étape consiste à projeter le point $\phi(\mathbf{x}_{obs})$ sur l'hyperbole B_F en un point $\phi(\mathbf{x}_{proj})$, si $\phi(\mathbf{x}_{obs})$ n'appartient pas au support de la distribution. Cependant cette projection ne se fait pas de manière euclidienne dans l'espace F . Les éléments de l'espace appartiennent à la sphère unité de l'espace F (avec un noyau gaussien, $K(\mathbf{x}, \mathbf{x}) = \|\phi(\mathbf{x})\|^2 = 1$, quel que soit \mathbf{x} appartenant à l'espace initial). Il faut donc contraindre la projection à rester sur la sphère unité de l'espace F (voir Fig. 10.12). Les détails de calcul sont décrits dans l'article [PARK et coll. 2007].

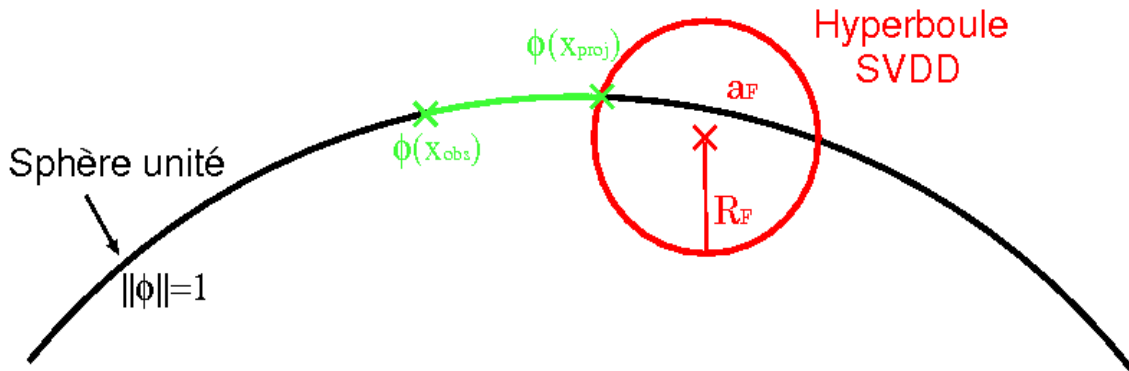


FIGURE 10.12 – Schéma représentant l'espace F des caractéristiques obtenu avec un noyau gaussien. L'hyperbole SVDD est visualisée en vert. Toutes les données de l'espace initial appartiennent à la sphère unité représentée en noir. Nous observons que la trajectoire de la projection visualisée en rouge réside sur cette sphère unité.

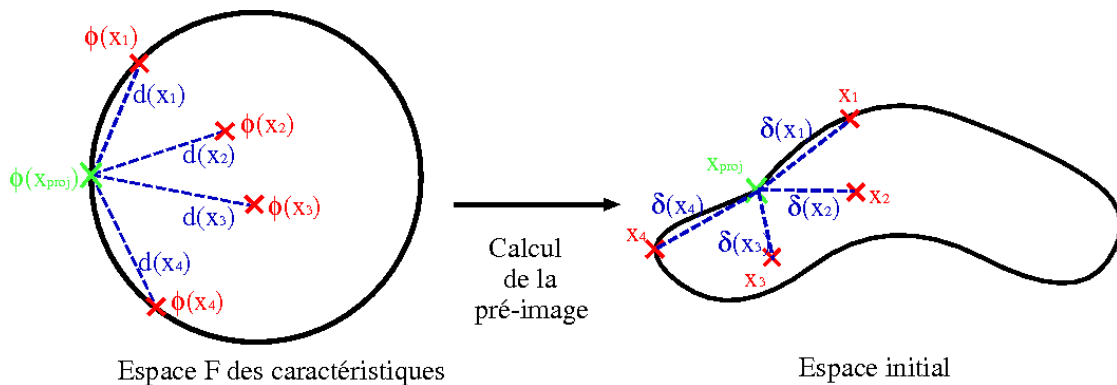


FIGURE 10.13 – Schéma représentant le calcul de la pré-image \mathbf{x}_{proj} du point projeté $\phi(\mathbf{x}_{proj})$ à partir de quatre de ses plus proches voisins. Le principe est d'estimer \mathbf{x}_{proj} tel que les distances $\delta(\mathbf{x}_i)$ dans l'espace initial soient les mêmes que les distances $d(\mathbf{x}_i)$ dans l'espace F des caractéristiques.

Si l'individu appartient au support, nous n'effectuons aucune opération de projection.

Nous pouvons considérer cette projection sur un atlas comme une extension des travaux décrits dans [GOLLAND et coll. 2005] pour une étude un individu par rapport à un groupe. Dans [GOLLAND et coll. 2005], les auteurs ne considèrent pas le projeté, mais s'intéressent à la direction de projection orthogonale à l'hyperplan séparant les deux populations.

Une fois la projection effectuée, il faut estimer la pré-image \mathbf{x}_{proj} de $\phi(\mathbf{x}_{proj})$. Cette estimation est complexe car nous n'avons pas la connaissance de la fonction $\phi(\cdot)$ [MIKA et coll. 1999]. Nous proposons d'utiliser la méthode décrite dans [PARK et coll. 2007], basée sur le MDS. Plus précisément, cette stratégie considère les k -plus proches voisins $\phi(\mathbf{x}_i)$ de $\phi(\mathbf{x}_{proj})$, ainsi que les k distances $d(\mathbf{x}_i)$ entre les voisins $\phi(\mathbf{x}_i)$ et la projection $\phi(\mathbf{x}_{proj})$. Ensuite, on estime la pré-image \mathbf{y}_{proj} telle que les distances $\delta(\mathbf{x}_i)$ dans l'espace initial entre la pré-image et les k voisins \mathbf{y}_i soient les plus similaires possibles aux distances $d(\mathbf{x}_i)$ dans l'espace F des caractéristiques. Ce calcul est effectué même si l'individu appartient au support.

Cette stratégie est particulièrement intéressante dans le cas de la création du modèle. En effet, nous utilisons uniquement une sous-partie de la population pour reconstruire l'individu considéré. Cette idée de propagation d'atlas à partir d'une sous-partie de la population se retrouve dans [WOLZ et coll. 2010].

Une fois la représentation \mathbf{x}_{proj} estimée, il faut générer l'individu \mathbf{y}_{proj} qui est la partie de \mathbf{y}_{obs} expliquée par le modèle. On génère \mathbf{y}_{proj} grâce à la fonction \mathbf{f} appliquée à \mathbf{x}_{proj} .

Nous calculons ensuite pour l'ensemble de la base d'apprentissage l'erreur au modèle telle que $\boldsymbol{\varepsilon}_i = \mathbf{y}_{obs}^i - \mathbf{y}_{modele}^i$, avec \mathbf{y}_{modele}^i correspondant à la pré-image d'un individu de la base d'apprentissage, projetée ou non sur l'hyperboule B_F , et i variant pour l'ensemble de la base d'apprentissage. Nous supposons que la variable aléatoire $\boldsymbol{\varepsilon}_i$ suit une loi gaussienne de moyenne nulle et de matrice de variance-covariance Ψ diagonale. Nous estimons cette matrice selon l'équation 9.14. Les anomalies dans le nouvel individu \mathbf{y}_{obs} à tester sont détectées en étudiant chaque composante du z-score $\mathbf{t} = \Psi^{-\frac{1}{2}}\boldsymbol{\varepsilon}$, par rapport à un seuil T . Le z-score \mathbf{t} correspond à la probabilité que le résidu $\boldsymbol{\varepsilon}$ suive une loi de Student avec $N - 1$ degrés de libertés, où N est le nombre d'individus de la population de référence [VIK et coll. 2007].

3. Applications

Nous proposons deux exemples pour montrer l'utilisation de l'atlas pour la détection et l'analyse. Nous montrerons l'effet de la réduction de dimension sur chacun des exemples.

3.1 Application en détection

Nous proposons d'étudier l'impact de la réduction de dimension sur l'étape de détection. Nous utilisons pour cette expérience 50 images de FA extraites d'une base disponible sur internet⁴ [BULLITT et coll. 2005]. Nous effectuons une parcellisation de ces cerveaux en 90 ROIs correspondant au *template* AAL [TZOURIO-MAZOYER et coll. 2002]. Ces ROIs sont de tailles et d'intensités très variées (*nombre moyen de voxels* : 1450 ± 770 et *moyenne de FA sur une ROI* : 0.24 ± 0.12).

Pour une des 90 ROIs anatomiques, nous allons inclure pour un des 50 individus une occlusion consistant en une variation de FA à l'intérieur de cette ROI. Nous apprenons le support de la distribution des 49 autres individus. Nous testons si l'individu avec l'occlusion est à l'intérieur ou non du support. Nous effectuons cette opération également pour les 49 autres individus, et pour les 90 ROIs anatomiques. Cette procédure correspond à une méthodologie de validation croisée du type *leave one out*. Cette stratégie est présentée Fig. 10.14.

Dans les occlusions, nous soustrayons successivement différentes valeurs de FA (0.05, 0.1, 0.2, et 0.3). De plus, nous faisons varier la taille de ces occlusions proportionnellement à la taille de la ROI considérée (10%, 20% et 30% du nombre total de voxels de la ROI considérée).

4. Téléchargeable à l'adresse suivante : <http://www.insight-journal.org/midas/community/view/21>.

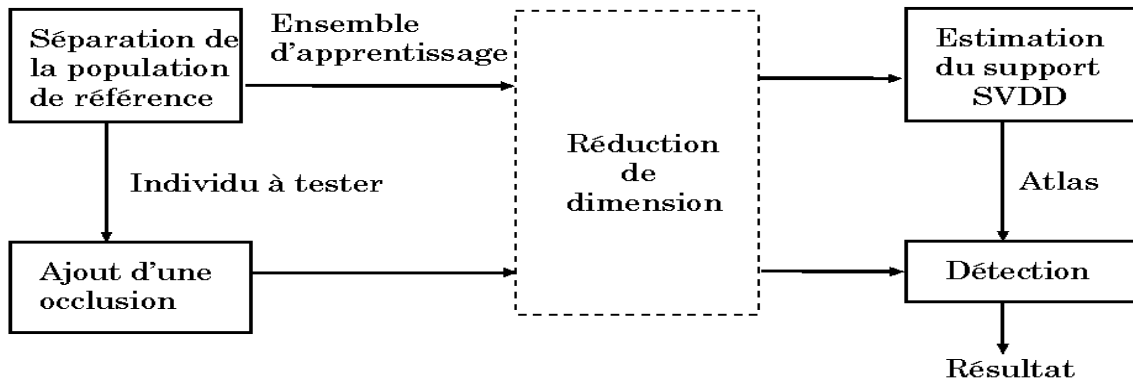


FIGURE 10.14 – Schéma décrivant l'expérience pour une itération. Une image correspond à une des 90 ROIs anatomiques de la parcellisation. Chaque image de la population de référence devient successivement l'image à tester, à laquelle on ajoute une occlusion. On estime le modèle sur les individus restant. Deux choix sont ensuite possibles : utiliser une réduction de dimension ou non. Puis on apprend le support de la distribution. Finalement, on détecte si l'image avec l'occlusion appartient ou non à la population de référence. On réitère cet algorithme pour les 90 ROIs anatomiques. Si on utilise la réduction de dimension, il n'y a pas de régression car l'étape de détection se fait dans l'espace réduit.

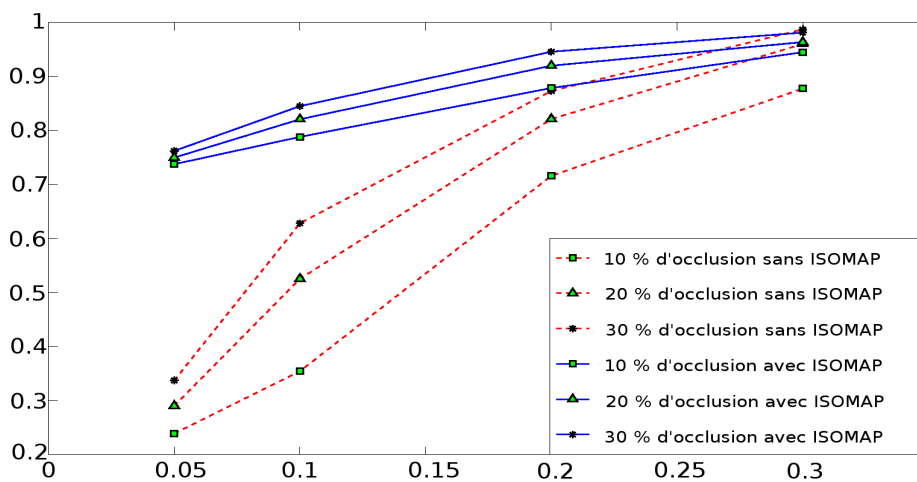


FIGURE 10.15 – Schéma représentant la fraction de bonnes détections en fonction de la variation de FA dans l'occlusion. Les trois courbes rouges en pointillés correspondent aux cas sans réduction de dimension pour des occlusions de 10%, 20% et 30%. Les trois courbes bleues en trait plein correspondent aux cas avec réduction de dimension pour les mêmes variations dans la taille des occlusions que dans le cas sans réduction de dimension. On remarque que plus l'occlusion est grande, plus la ROI est correctement détectée. On observe une meilleure détection lorsqu'on utilise la réduction de dimension.

Finalement, nous effectuons ce test sur les ROIs originales et sur les ROIs représentées dans un espace réduit de dimension égale à trois après utilisation de l'ISOMAP. Nous observons les résultats Fig. 10.15. Nous remarquons que la détection s'améliore nettement avec l'utilisation de l'espace réduit.

3.2 Application en analyse

Nous utilisons ici une base de cent images⁵ de dimension 8×8 représentant des chiffres 9 (voir Fig. 10.16). Nous testons nos atlas avec et sans réduction de dimension. Pour comparaison, nous testerons l'utilisation de la moyenne, ainsi que de l'ACP. Ces différentes techniques ont été décrites au chapitre précédent en section 2.

Nous observons que la moyenne est floue, et ne permet pas de distinguer des modifications complexes. Nous remarquons que les images de chiffres constituent une population trop complexe pour être analysée par des modèles linéaires. Nous observons aussi que les résultats obtenus avec la réduction de dimension permettent d'obtenir des détails plus fins que sans réduction de dimension.

Nous proposons de comparer la méthode proposée dans [GERBER et coll. 2010], nommée dans le reste de ce mémoire méthode de Gerber, à notre atlas sur une base d'images 8×8 représentant des chiffres 6 (voir Fig. 10.17 et 10.18). Tout d'abord, la méthode de Gerber n'est utilisée à la base que pour décrire une population de référence, et n'introduit pas de nouvel individu potentiellement aberrant. Néanmoins cette technique propose un modèle génératif utilisant un espace réduit obtenu grâce à l'ISOMAP, et peut être utilisée pour étudier de nouveaux individus potentiellement aberrants. La différence majeure entre l'atlas proposé dans cette thèse et la méthode de Gerber correspond à la manière de projeter le nouvel individu sur la population de référence dans l'espace réduit. Plus précisément, dans [GERBER et coll. 2010], la projection du nouvel individu est faite directement dans l'espace initial, grâce à une régression par NNW. Cette régression est apprise uniquement sur la population de référence. L'inconvénient est qu'il n'existe pas de moyen de contrôler cette projection. Dans cette thèse, on propose d'estimer le nouvel individu dans l'espace réduit, puis de le projeter sur la frontière du support de la distribution de la population de référence.

Nous proposons un schéma récapitulatif des différentes méthodes étudiées pour l'étape d'analyse en figure 10.19.

4. Conclusion partielle

Dans ce chapitre, nous avons proposé la conception d'un atlas basé sur une méthode à noyaux associée à une réduction de dimension. Les méthodes à noyaux, et plus précisément les algorithmes du SVDD et de projection, permettent une étude locale de la population. La réduction de dimension, quant à elle, permet de représenter les données de façon compacte, ce qui facilite l'estimation du support des distributions, et donne ainsi la possibilité de détecter des phénomènes plus fins. Les outils de détection et d'analyse peuvent être utilisés indépendamment l'un de l'autre. De plus, ces outils peuvent être utilisés pour diverses catégories d'objets, à partir du moment où l'on a une métrique appropriée sur les objets. Cette propriété est particulièrement intéressante dans le cadre de l'IRMd où les modèles sont nombreux.

5. Les images sont celles utilisées dans l'article [PARK et coll. 2007]

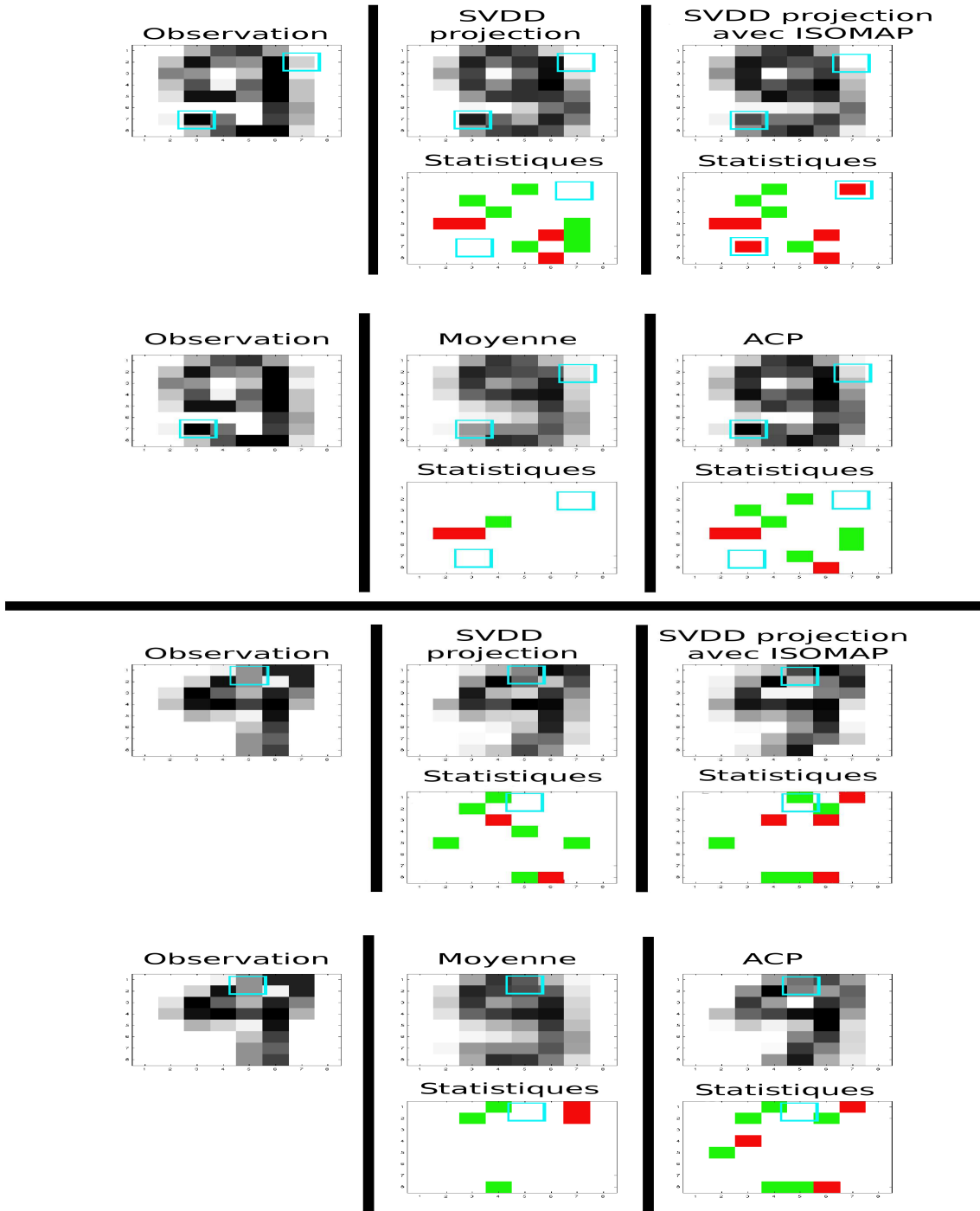


FIGURE 10.16 – Exemple pour deux images du chiffre neuf (« observation »). On compare ici plusieurs stratégies de détections statistiques. Premièrement, le nouvel individu \mathbf{y}_{obs} (« observation ») est représenté par $\mathbf{f}(\mathbf{x}_{obs})$ simulant différents modèles (SVDD projection, SVDD projection avec réduction de dimension, moyenne, ACP). Deuxièmement, on effectue une détection statistique sur les vecteurs $\boldsymbol{\varepsilon}$ des résidus en considérant les z-scores associés. Les images du bas correspondent aux détections statistiques. La couleur verte suppose que l'intensité est statistiquement inférieure à l'intensité prédite par le modèle. La couleur rouge suppose que l'intensité est statistiquement supérieure à l'intensité prédite par le modèle. Le cadre bleu correspond à des pixels d'intérêt supposés anormaux (ce choix est arbitraire). On observe que seul les atlas utilisant le SVDD et la réduction de dimension permettent de détecter ces pixels. On remarque que la moyenne est un modèle flou, éloigné du neuf d'origine.

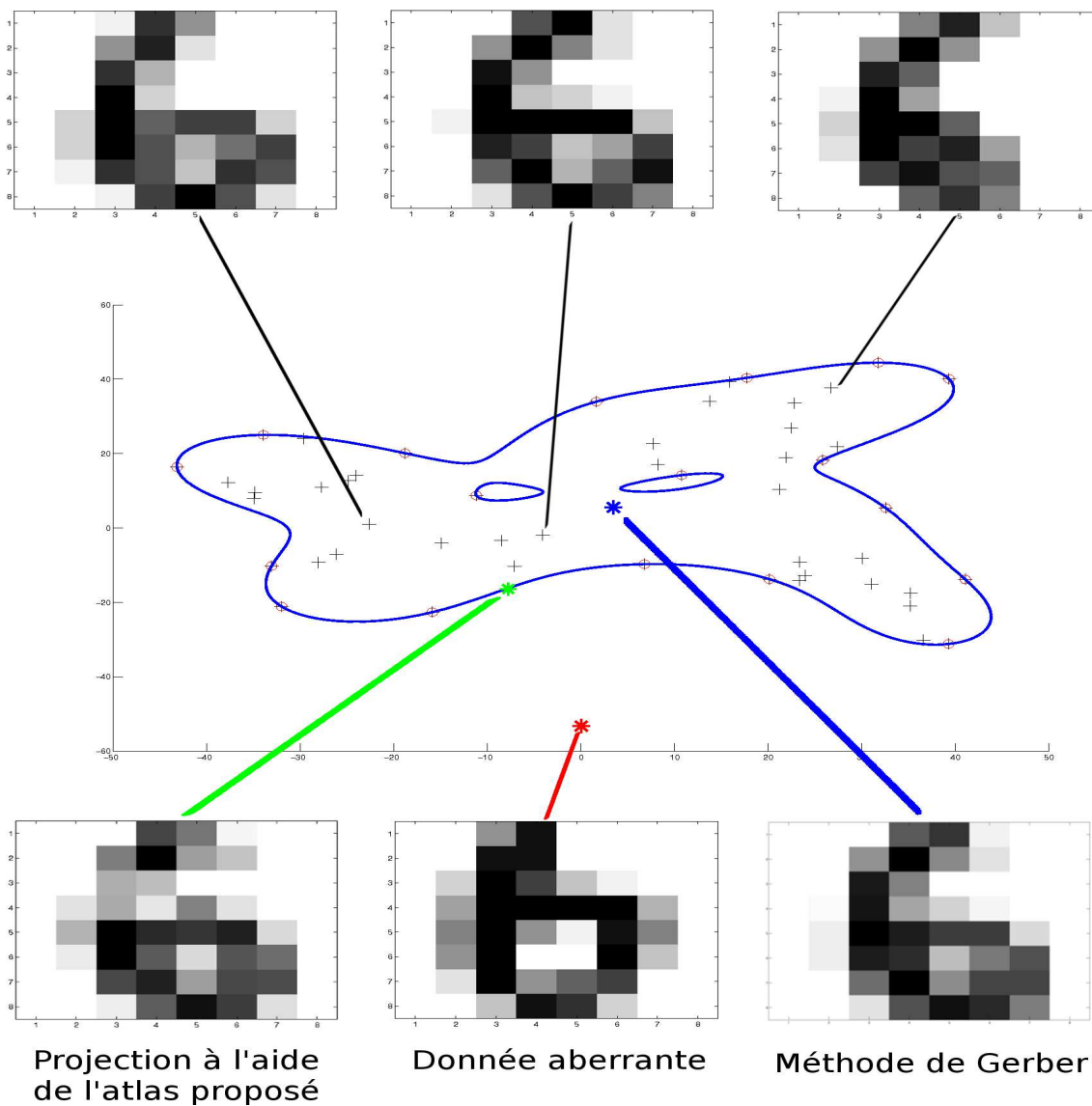


FIGURE 10.17 – Graphe représentant les images de 6 dans un espace réduit de dimension deux. Trois images de 6 de la population de référence sont visualisées sur la ligne supérieure. Le support de la distribution est visualisé en bleu. Un nouvel individu aberrant à analyser est représenté en rouge. En bleu, on visualise la projection obtenue par la méthode de Gerber. En vert, on visualise la projection obtenue par notre méthode d'analyse. La donnée aberrante observée, ainsi que les individus générés par les deux modèles (après l'étape de régression f permettant de passer de l'espace réduit à l'espace initial) sont visualisés dans la ligne inférieure.

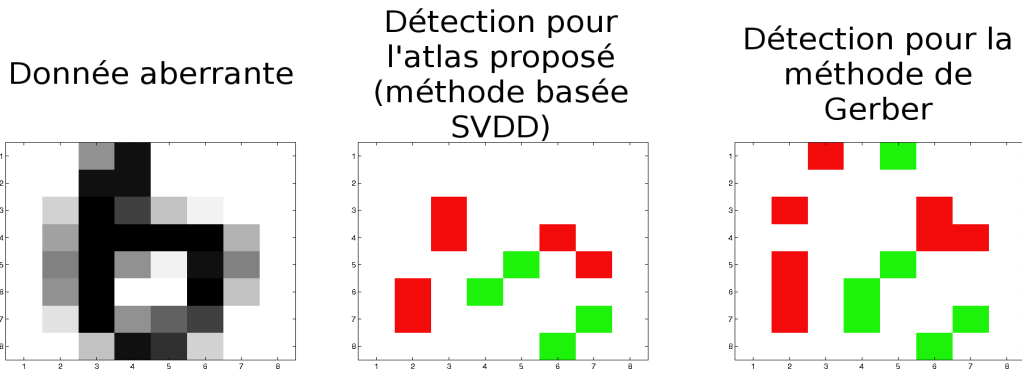


FIGURE 10.18 – On compare ici deux stratégies de détection statistique pour la donnée aberrante de la Fig. 10.17. Nous observons que les détections sont plus nombreuses avec la méthode de Gerber.

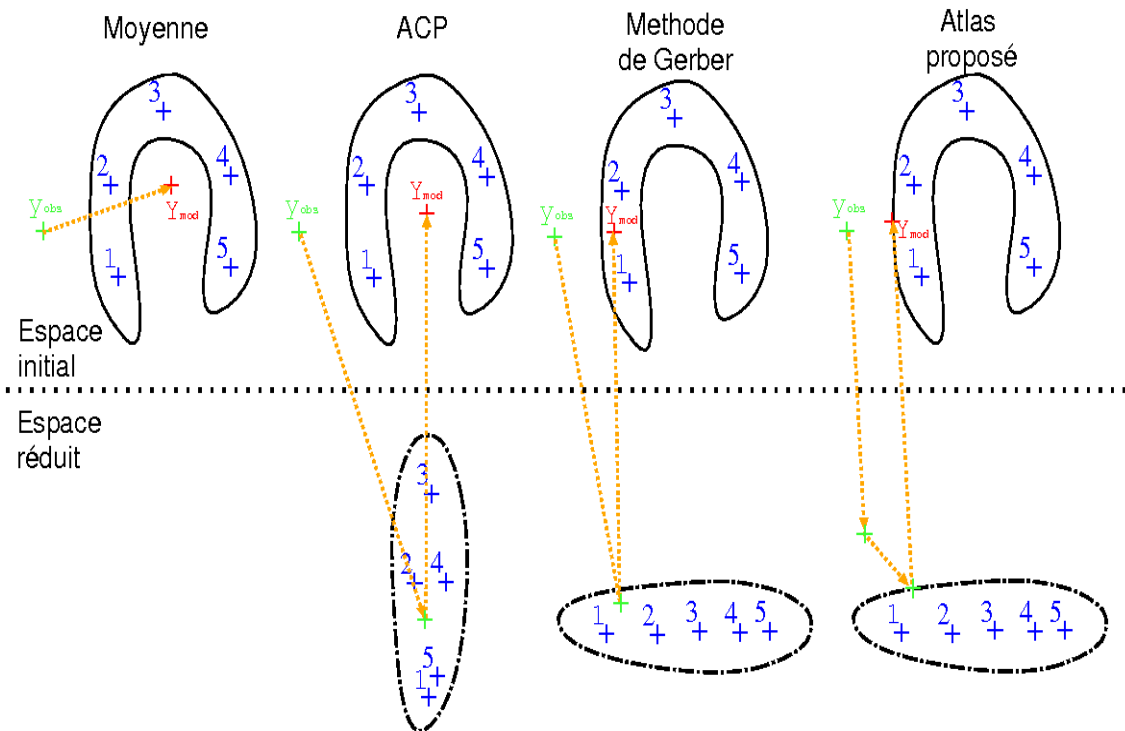


FIGURE 10.19 – On schématise ici quatre méthodologies (moyenne, ACP, méthode de Gerber, méthode proposée dans le cadre de cette thèse) utilisées dans l'étape d'analyse pour estimer \mathbf{y}_{mod} (noté $\mathbf{f}(\mathbf{x}_{obs})$ dans ce mémoire), visualisé en rouge, à partir d'une donnée aberrante observée \mathbf{y}_{obs} , visualisée en vert. Le support de la distribution de la population de référence, modélisé en bleu, est complexe. La moyenne estimée de manière euclidienne n'appartient pas à la distribution. L'ACP estime de manière incorrecte la population dans l'espace réduit (dans ce schéma, les voisinages locaux de l'espace initial ne sont pas conservés dans l'espace réduit). La méthode de Gerber estime correctement l'espace réduit grâce à l'ISOMAP, mais la projection est éloignée de l'individu dans l'espace réduit. La méthode proposée dans cette thèse décompose la projection sur la population de référence dans l'espace réduit en deux étapes. La première étape correspond à la projection dans l'espace réduit de manière non linéaire grâce à l'ISOMAP. La deuxième étape consiste à projeter sur la frontière du support de la distribution estimé par le SVDD.

Chapitre 11

Applications à l'IRM de diffusion

1.	Atlas basé sur la connectivité anatomique et sur les ROIs	139
1.1	Principe	139
1.2	Méthodes	139
1.3	Résultats	141
2.	Atlas basé sur la connectivité anatomique et fonctionnelle	143
2.1	Principe	143
2.2	Méthodes	143
2.3	Résultats	145
3.	Conclusion partielle	147

Dans ce chapitre, nous proposons deux atlas basés sur la connectivité anatomique. Ces atlas sont utilisés dans le cadre de l'étude des patients atteints de troubles de la conscience. La population de patients étant hétérogène (âge, sexe, étiologie du coma, délai entre l'accident et l'IRM), nous considérons la problématique de la comparaison d'un patient par rapport à une population de référence supposée saine.

1. Atlas basé sur la connectivité anatomique et sur les ROIs

1.1 Principe

Le connectome reflète les connexions anatomiques entre les différentes aires cérébrales. Cependant, une connexion entre une ROI \mathbf{A} et une ROI \mathbf{B} peut apparaître lésée dans trois cas : les faisceaux de fibres de substance blanche sont lésés, la ROI \mathbf{A} ou la ROI \mathbf{B} sont pathologiques. Le connectome ne permet pas de distinguer ces trois différents cas. Nous proposons donc de rajouter de l'information pour pouvoir déterminer si une ROI est pathologique ou non. Plus précisément, nous étudierons la FA de chaque voxel de chaque ROI. Nous comparerons ensuite la FA de chaque ROI des individus pathologiques avec celle de la population saine.

A notre connaissance, c'est la première fois qu'on s'intéresse simultanément à l'intégrité des connexions et des ROIs.

1.2 Méthodes

La population de référence comporte 21 individus supposés sains. Les individus pathologiques sont au nombre de 16. Ces individus pathologiques présentent des troubles de la conscience per-

sistants suite à un accident (traumatisme crânien, arrêt cardio-respiratoire ...), sans sédation et en respiration spontanée. Les images pondérées en diffusion ont été acquises dans 30 directions \mathbf{g}_i de gradient (deux acquisitions pour chaque direction) avec une valeur de b égale à 1000 s/mm^2 sur une machine IRM de 1.5 Tesla (Avanto, Siemens, Erlangen, Allemagne). Deux images non pondérées en diffusion ont aussi été acquises.

Le protocole d'acquisition est très important dans le cadre du coma. Plus précisément, les données de chaque patient ont été acquises dans un court délai après l'accident. En effet, il est connu que les dégénérescences observées s'aggravent au cours du temps. Il était donc nécessaire d'étudier les patients après le même délai suite à leurs accidents.

Dans un premier temps, nous effectuons l'étape d'analyse pour les connectomes. L'estimation du connectome et sa mise en correspondance sont décrites à la section 2.2 du chapitre 7. Nous effectuons un seuillage pour ne conserver que les connexions supérieures à 0.1. Ce seuil est arbitraire, et nous permet de ne considérer que 1914 connexions sur les 8010 connexions possibles. Le but du seuillage est de ne conserver que les connexions les plus importantes.

Nous effectuons ensuite une réduction de dimension. La dimension de l'espace réduit est déterminée en fonction de l'erreur de reconstruction [TENENBAUM et coll. 2000]. Nous visualisons dans la Fig. 11.1 l'erreur de reconstruction en fonction de la dimension de l'espace réduit. Nous choisissons une dimension égale à trois. Cette dimension est sélectionnée afin d'obtenir une erreur de reconstruction faible et de considérer un espace réduit rempli densément.

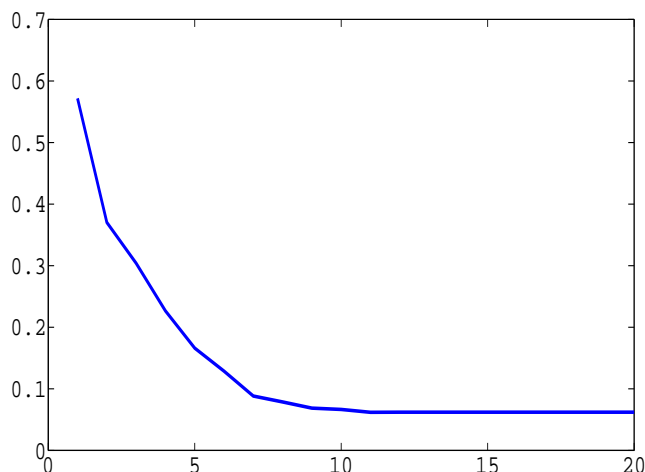


FIGURE 11.1 – Graphe représentant l'erreur de reconstruction en fonction de la dimension de l'espace réduit. Il existe généralement deux parties dans ce type de courbes. Plus précisément, la première partie de la courbe décroît vite. Ceci équivaut à considérer que la dimension de l'espace réduit n'est pas suffisante pour modéliser les données, et qu'une dimension supplémentaire améliore nettement la reconstruction. Dans la deuxième partie de la courbe, la pente est plus faible. Ceci équivaut à dire que l'ajout de dimension supplémentaire n'apporte pas beaucoup d'information. La dimension optimale de l'espace réduit correspond à la limite entre ces deux parties.

Pour l'étape de projection sur le support de la distribution obtenue par SVDD, nous considérons quatre voisins.

Dans un deuxième temps, nous effectuons l'étape de détection pour chacune des 90 ROIs.

Plus précisément, nous considérons les valeurs de FA en chaque voxel de la ROI considérée. Chaque carte de FA a été préalablement recalée avec une transformation non linéaire sur un même référentiel correspondant à l'atlas AAL [TZOURIO-MAZOYER et coll. 2002]. Ces transformations sont les mêmes que celles utilisées lors la mise en correspondance des connectomes.

Ces valeurs de FA sont considérées comme des vecteurs \mathbf{y}_{FA} de grande dimension dans \mathbb{R}^M , où M correspond au nombre de voxels de la ROI étudiée. La distance considérée entre deux vecteurs \mathbf{y}_{FA} est une distance euclidienne. Nous n'effectuons pas d'étape d'analyse. En effet, nous voulons savoir si la ROI considérée est potentiellement lésée ou non (ceci correspond uniquement à l'étape de détection).

Etudions maintenant la réduction de dimension pour chaque ROI. Choisir la dimension de l'espace réduit ROI par ROI serait trop laborieux. Nous choisissons une dimension égale à 3 pour chacune des 90 réductions de dimension. L'idée est toujours d'avoir un compromis entre une erreur de reconstruction faible et un espace dense.

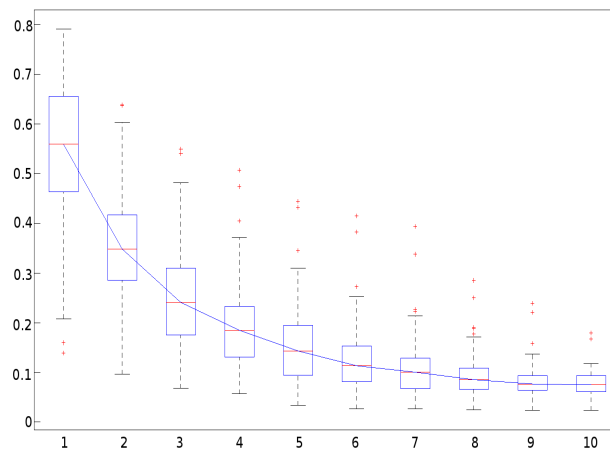


FIGURE 11.2 – Diagramme représentant l'erreur de reconstruction par rapport au nombre de dimensions de l'espace réduit pour les 90 ROIs.

Nous proposons de visualiser la conception de cet atlas dans le schéma Fig.11.3.

1.3 Résultats

Tout d'abord, nous allons nous intéresser à l'erreur ε au modèle estimée lors de l'étape d'analyse grâce au modèle génératif $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \varepsilon$. Dans ce modèle, ε est supposée suivre une loi gaussienne de moyenne nulle. Ceci équivaut à considérer que la fonction \mathbf{f} a réussi à modéliser correctement la population. Pour tester la normalité de ε , nous avons effectué un test de Lilliefors pour chacune des composantes de ε . Nous réalisons donc un grand nombre de tests, et devons effectuer une correction pour compenser les comparaisons multiples [SHAFFER 1995]. L'idée la plus simple est de considérer une correction de Bonferroni proposant un seuil approximativement de 10^{-5} . Cependant, le logiciel utilisé (Matlab) ne permet pas d'avoir des p-valeurs inférieures à 0.001. Nous utiliserons donc ce dernier seuil. Nous obtenons alors 22.6% de composantes qui ne suivent pas une loi gaussienne. Ce pourcentage est relativement élevé. Cependant, la correction due à la comparaison multiple n'a pas été parfaite. De plus, le nombre d'individus est relativement faible (21 individus de référence).

Nous proposons de visualiser deux graphes de normalité pour une composante i dont l'erreur

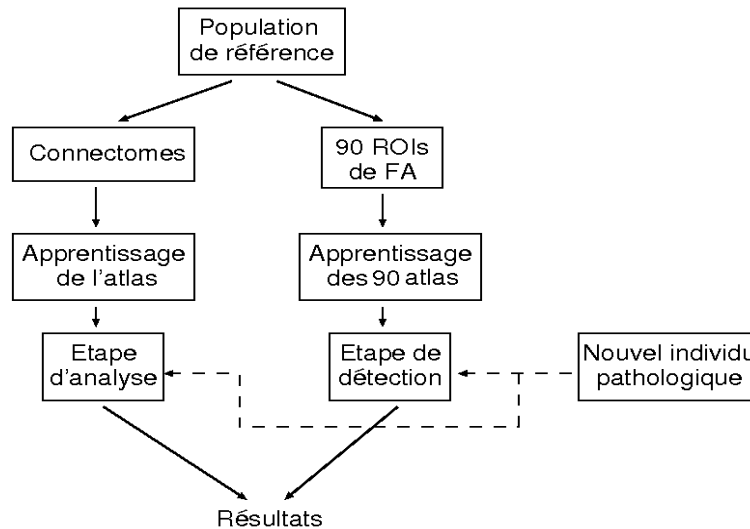


FIGURE 11.3 – Schéma décrivant la création et l'utilisation de l'atlas.

ε_i suit une loi gaussienne et une autre composante j dont l'erreur ε_j ne suit pas une loi gaussienne (voir Fig.11.4).

Les résultats sont visualisés avec le logiciel Gephi¹ en Fig. 11.5. Seules les connexions supposées pathologiques sont visualisées. Ces connexions correspondent aux arêtes des graphes. Les ROIs quant à elles correspondent aux différents nœuds du graphe. Le diamètre du nœud correspond au degré du nœud (*i.e.* le nombre de connexions pathologiques liées à ce nœud). Plus précisément, plus le nœud est grand, plus le nombre de connexions pathologiques attachées à ce nœud est élevé.

Les différences observées aux niveaux des FAs des ROIs sont visualisées par la couleur des nœuds. Plus précisément, le jaune représente une ROI se situant sur la frontière du support des distributions des ROIs de la population témoin, tandis que le rouge représente une ROI très éloignée du support. Une ROI supposé normale est bleue. Cependant, dans le cadre du coma, toutes les ROIs sont en dehors du support de la distribution de la population saine. L'intérêt est de savoir si les ROIs sont plus ou moins loin du support, ce qui correspond médicalement à déterminer si les ROIs sont plus ou moins lésées.

Pour plus de clarté, nous visualisons pour un même patient une vue de profil (à gauche) et une vue de dessus (à droite). On observe deux individus pathologiques sur les seize traités. Premièrement, on remarque que différentes ROIs intervenant dans les phénomènes de conscience apparaissent comme pathologiques à la fois au niveau des connexions anatomiques et au niveau des ROIs en FA. On peut notamment citer les noyaux gris centraux (le thalamus, le putamen, les noyaux caudés), le précunéus, le gyrus fusiforme et le cortex olfactif. Plus précisément, on peut remarquer que les noyaux gris centraux sont à la fois lésés du point de vue des connexions et du point de vue des ROIs. On remarque aussi qu'il existe une altération globale de la FA (tous les nœuds apparaissent jaunes). On peut aussi se rendre compte que les déconnexions importantes d'une ROI ne sont pas forcément reliées aux lésions apparentes dans la ROI en FA.

1. Logiciel disponible à l'adresse web suivante : <http://gephi.org/>

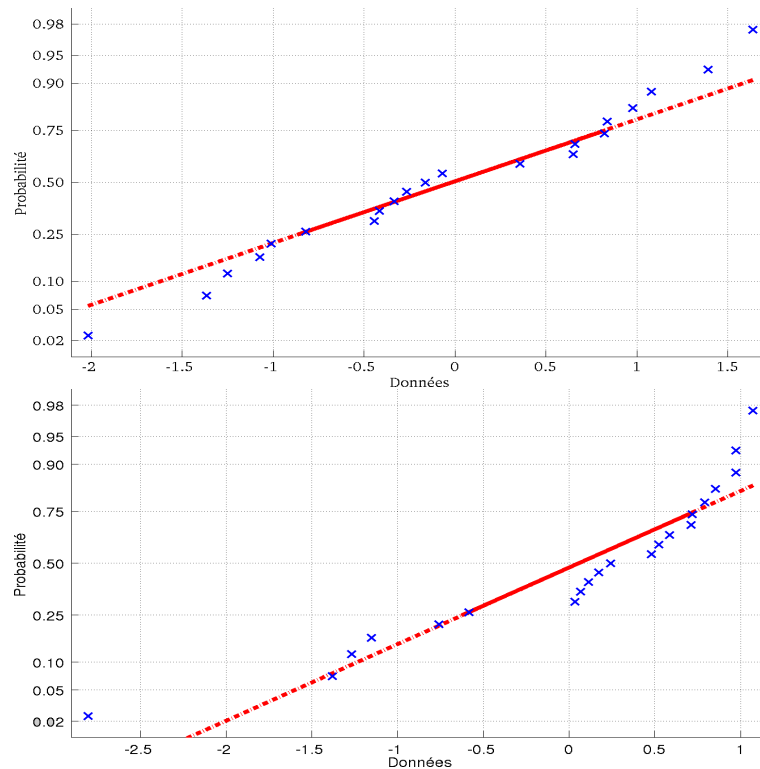


FIGURE 11.4 – Deux graphes de normalité. Le graphe du dessus représente l'erreur suivant une loi gaussienne, tandis que le graphe du dessous montre une erreur ne suivant pas une loi gaussienne. Plus précisément, les deux graphes ont pour abscisse les valeurs des erreurs ε_i pour une composante i . L'axe des ordonnées représente la probabilité des données si elles suivaient une loi normale. Le segment en trait plein de couleur rouge relie le premier quartile au troisième quartile. Ce segment est extrapolé jusqu'au premier et au dernier coefficient pour permettre de mieux visualiser la linéarité. Le but de ce graphe est de visualiser rapidement si les données suivent une loi normale. Plus précisément, lorsque les données suivent une loi normale, ces dernières suivent la droite rouge. Une autre distribution que la loi normale introduirait une courbure.

2. Atlas basé sur la connectivité anatomique et fonctionnelle

2.1 Principe

Certaines études récentes montrent qu'il existe des relations fortes entre les connectivités anatomique et fonctionnelles [DELIGIANNI et coll. 2010, VENKATARAMAN et coll. 2010].

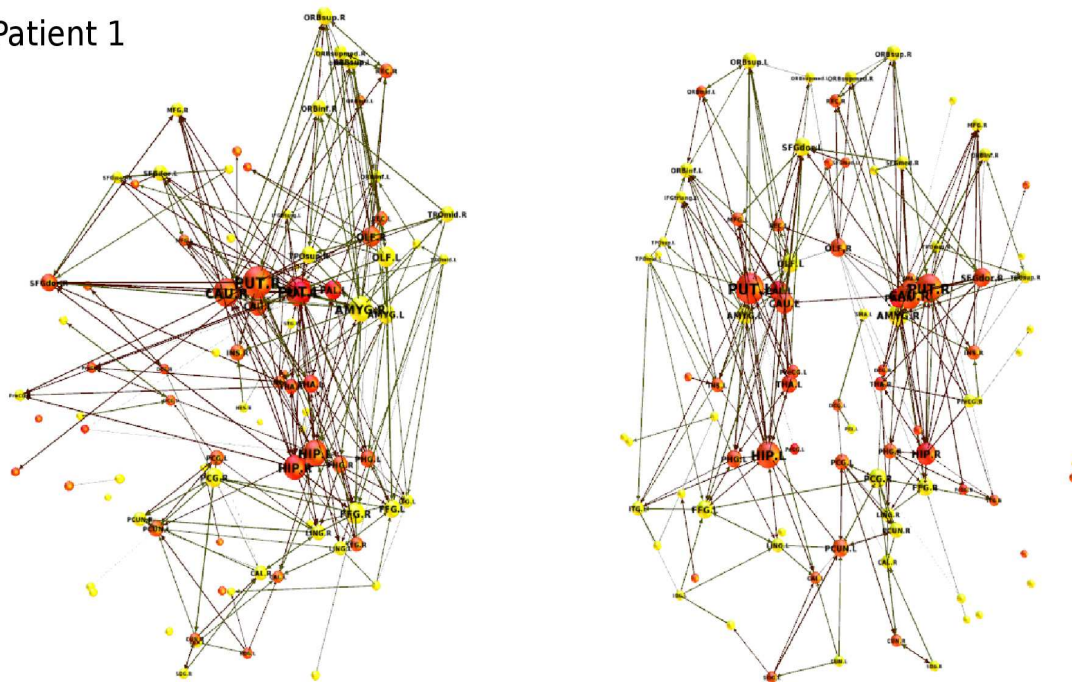
Nous proposons ici d'étudier pour un même patient les connexions anatomique et fonctionnelle pathologiques. La question sous-jacente est de savoir s'il existe une interaction entre connectivité anatomique et connectivité fonctionnelle dans le cadre du coma.

2.2 Méthodes

Pour l'étude anatomique et fonctionnelle, la population de référence comporte 21 et 20 témoins respectivement. Les individus des deux populations sont différents. Nous avons 10 individus pathologiques ayant été acquis avec les deux modalités.

Les acquisitions pour les deux modalités ont été effectuées sur une machine IRM de 1.5 Tesla (Avanto, Siemens, Erlangen, Allemagne). Pour l'IRMd, les images ont été acquises de la

Patient 1



Patient 2

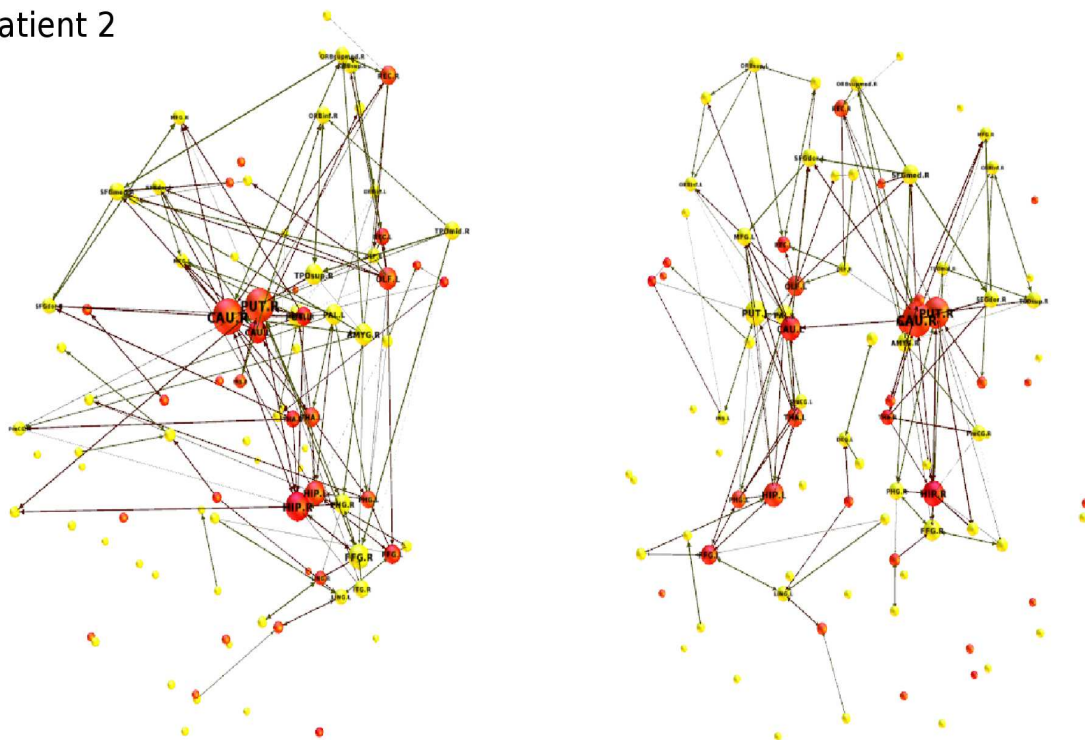


FIGURE 11.5 – Graphes en trois dimensions représentant les connexions pathologiques chez deux patients. À gauche, on observe une vue de profil, et à droite, une vue de dessus. Ces graphes ont été obtenus avec Gephi. L'interprétation de ces graphes est faite dans la section 1.3.

manière décrite dans la section précédente. Pour l'IRMf de repos, les images ont été acquises avec une séquence EPI-BOLD de 20 minutes.

Une précision doit être apportée par rapport au protocole de l'étude de la section précédente. Les patients n'étaient pas sédatisés au cours de l'acquisition. Cette sédation peut provoquer des effets indésirables principalement lors de l'étude en IRMf de repos. L'absence de la sédation permet donc une meilleure comparaison avec le groupe de référence.

Dans cette étude, l'élaboration de l'atlas de connectivité anatomique est le même que celui décrit à la section précédente. Pour la connectivité fonctionnelle, nous considérons les matrices d'association, dont chaque coefficient $a_{i,j}$ correspond à la corrélation des activités temporelles entre la ROI i et la ROI j . L'élaboration de ces matrices est détaillée dans l'article [ACHARD et coll. 2006]. A la différence des matrices représentant la connectivité anatomique, les matrices d'associations de connectivité fonctionnelle sont symétriques.

Pour les deux types de connectivité, nous considérons uniquement les 400 plus fortes connexions en moyenne. Pour l'espace réduit, nous considérons un espace réduit de dimension 3. Pour l'étape de projection sur le support de la distribution appris par le SVDD, le nombre de voisins est égal à 4.

La conception de l'atlas est représentée en Fig. 11.6 .

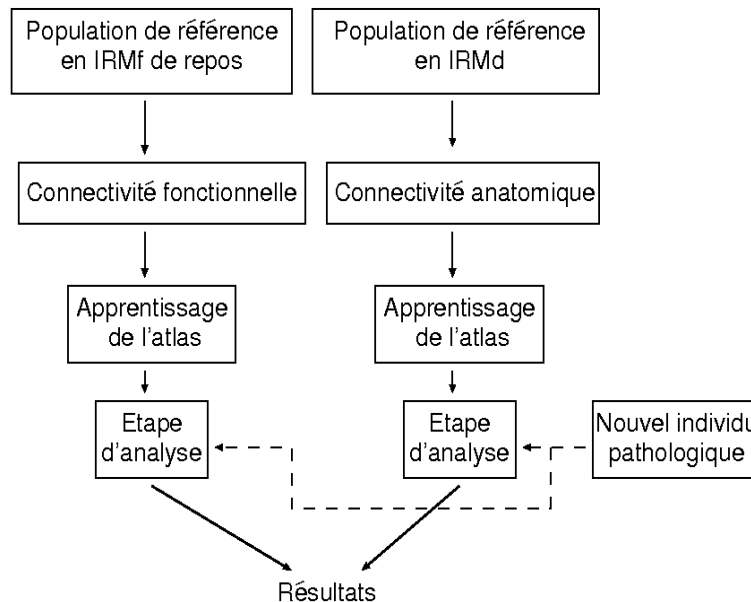


FIGURE 11.6 – Schéma décrivant l'atlas élaboré pour étudier simultanément les connectivités anatomique et fonctionnelle.

2.3 Résultats

Dans un premier temps, nous observons les matrices représentant la connectivité anatomique et la connectivité fonctionnelle (voir Fig. 11.7). Nous remarquons qu'il existe une plus grande variabilité en connectivité fonctionnelle qu'en connectivité anatomique. Ceci s'explique par le fait que, durant l'acquisition de l'IRMf de repos, on observe l'activité cérébrale à un instant t . Cette activité cérébrale peut varier fortement au cours du temps, selon l'activité de l'individu. Pour la connectivité anatomique, les matrices reflètent les structures anatomiques sous-jacentes. Ces structures sont globalement identiques pour chaque individu et ne varient que très lentement dans le temps, ce qui explique la faible variabilité.

Dans la Fig. 11.7, on observe des phénomènes différents entre les individus sains et pathologiques pour les connectivités anatomique et fonctionnelle. Plus précisément, en connectivité anatomique, on observe une diminution globale des coefficients de la matrice. En connectivité fonctionnelle, on observe des variations dans les motifs des matrices plus qu'une diminution de la connectivité fonctionnelle. Cette différence peut s'expliquer par le fait que la connectivité anatomique représente un lien physique entre deux ROIs, tandis la connectivité fonctionnelle correspond à une corrélation entre deux signaux. Les deux signaux peuvent être plus faibles que des signaux de référence mais cependant avoir une forte corrélation entre eux.

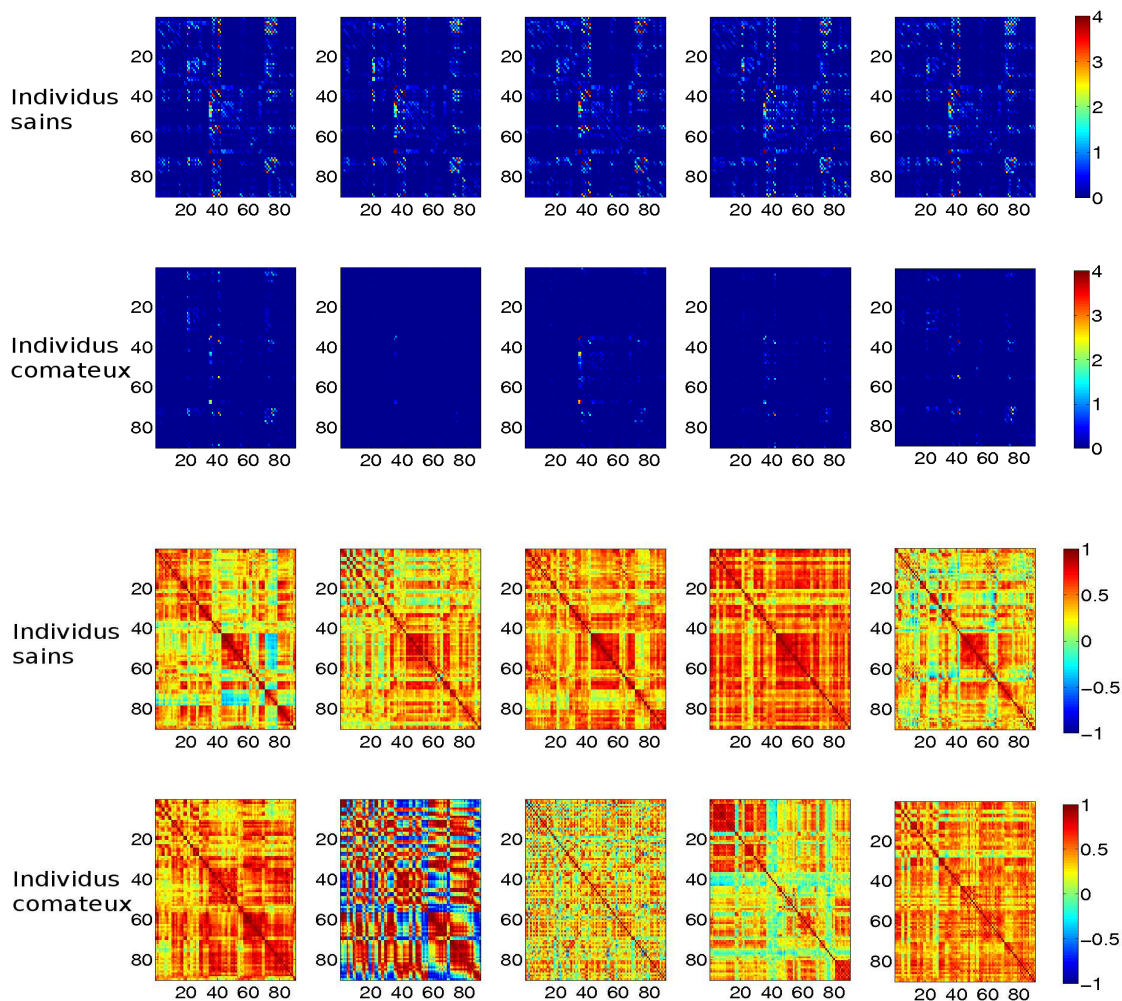


FIGURE 11.7 – Graphique représentant les matrices des connectivités anatomique et fonctionnelle. Les deux premières lignes correspondent à la connectivité anatomique. Les deux dernières lignes correspondent à la connectivité fonctionnelle. On observe que la variabilité existant en connectivité fonctionnelle est plus importante que celle existant en connectivité anatomique.

Nous pouvons observer ces différences de variabilité une fois que nous avons projeté les individus dans l'espace réduit (voir Fig. 11.8). Nous remarquons que les points représentant les patients entourent le nuage de points représentant les individus sains en connectivité fonctionnelles. En connectivité anatomique, les individus sains et pathologiques sont nettement mieux séparés.

Nous visualisons en Fig. 11.9, grâce au logiciel Gephi, les graphes de connectivités anatomique

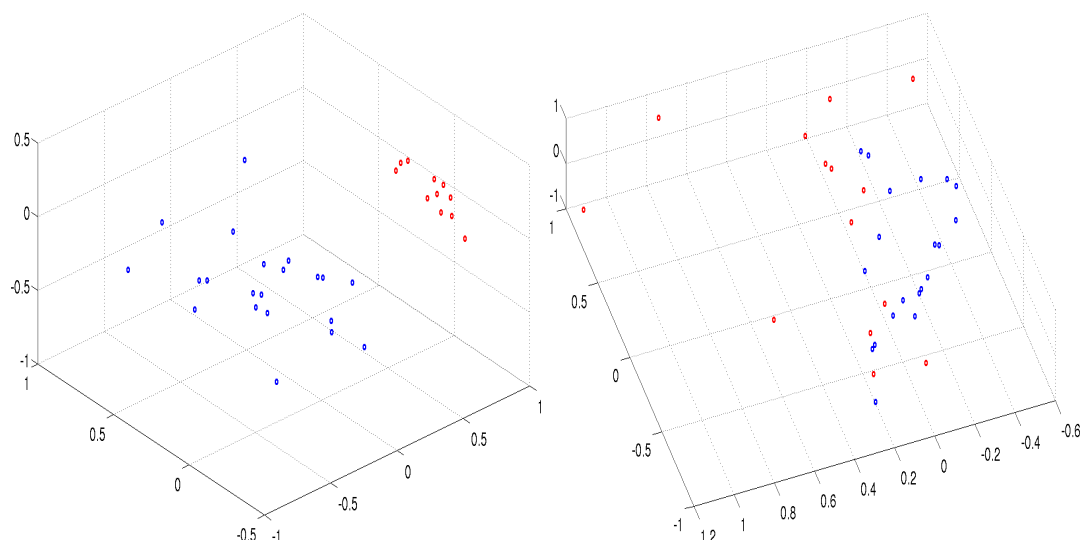


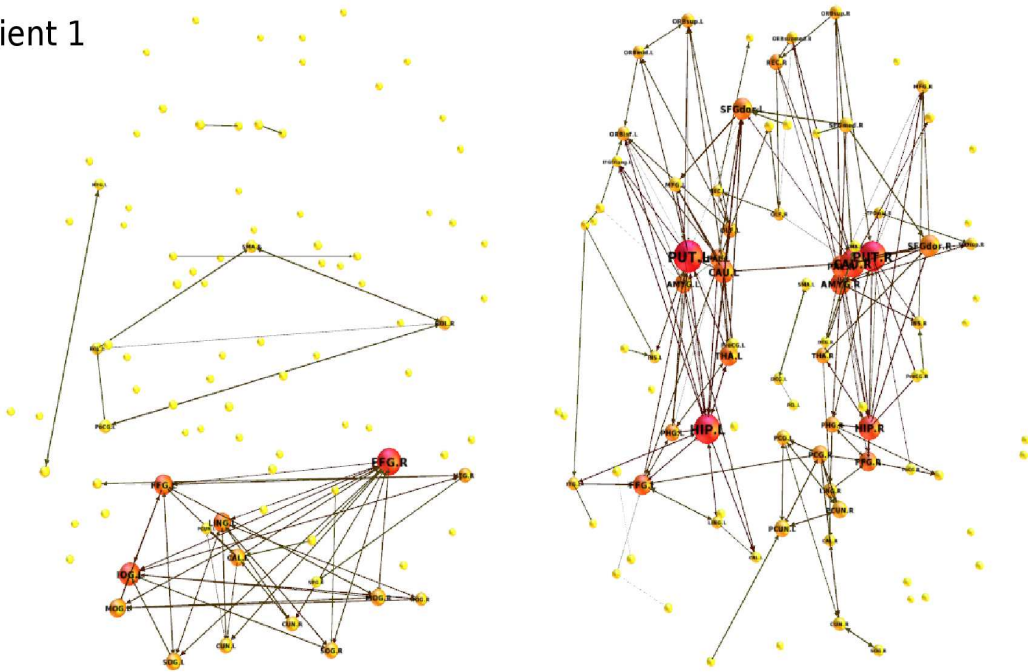
FIGURE 11.8 – Graphique représentant les matrices des connectivités projetées dans un espace réduit de dimension 3. Le graphe de gauche représente les matrices de connectivité anatomique. Celui de droite représente les matrices de connectivité fonctionnelle. La couleur bleue représente les individus sains, tandis que la couleur rouge représente les individus pathologiques.

et fonctionnelle. Nous pouvons observer que les connexions fonctionnelles et anatomiques supposées lésées ne sont pas situées aux mêmes endroits. Dans les deux exemples observés, les déconnexions observées en IRMf de repos sont localisées à l'arrière du cerveau tandis que les déconnexions observées en IRMd se situent au niveau frontal ainsi que dans les noyaux gris centraux. Plusieurs explications peuvent être avancées. Premièrement, nous avons remarqué que les connectivités anatomique et fonctionnelle ne correspondaient pas au même objet. Plus précisément, la connectivité fonctionnelle ne prend pas en compte l'intensité du signal. Deuxièmement, les noyaux gris centraux détectés en IRMd sont complexes à étudier en IRMf de repos. Ces ROIs travaillent simultanément avec de nombreuses autres ROIs. Les signaux temporels sont alors plus complexes à étudier. Une étude [BEHRENS et coll. 2003] propose d'ailleurs d'affiner la parcellisation d'une de ces ROIs (le thalamus). En effet, la parcellisation de ces noyaux gris peut être améliorée, permettant une meilleure étude en IRMf de repos. De plus, on a observé une forte diminution des FAs pour ces ROIs dans l'étude de la section précédente. Cette diminution peut en partie expliquer la forte diminution des connexions. On observe aussi des phénomènes de surconnexion qu'on n'observe pas en IRMd. Cette étude est préliminaire. De par la quantité d'information fournie, ces résultats sont relativement difficiles à analyser. L'approfondissement de l'analyse de ces résultats constitue une perspective de ces travaux de thèse.

3. Conclusion partielle

Dans ce chapitre, nous avons élaboré deux atlas basés sur la connectivité anatomique. Le fait d'utiliser la réduction de dimension nous permet d'étudier des objets mathématiques très différents telles que des ROIs ou des matrices de connectivité. De plus, les atlas proposés permettent soit de détecter des individus pathologiques avec un modèle discriminatif (par exemple les ROIs de FA), soit de localiser les zones pathologiques (les matrices de connectivités anatomique et fonctionnelle).

Patient 1



Patient 2

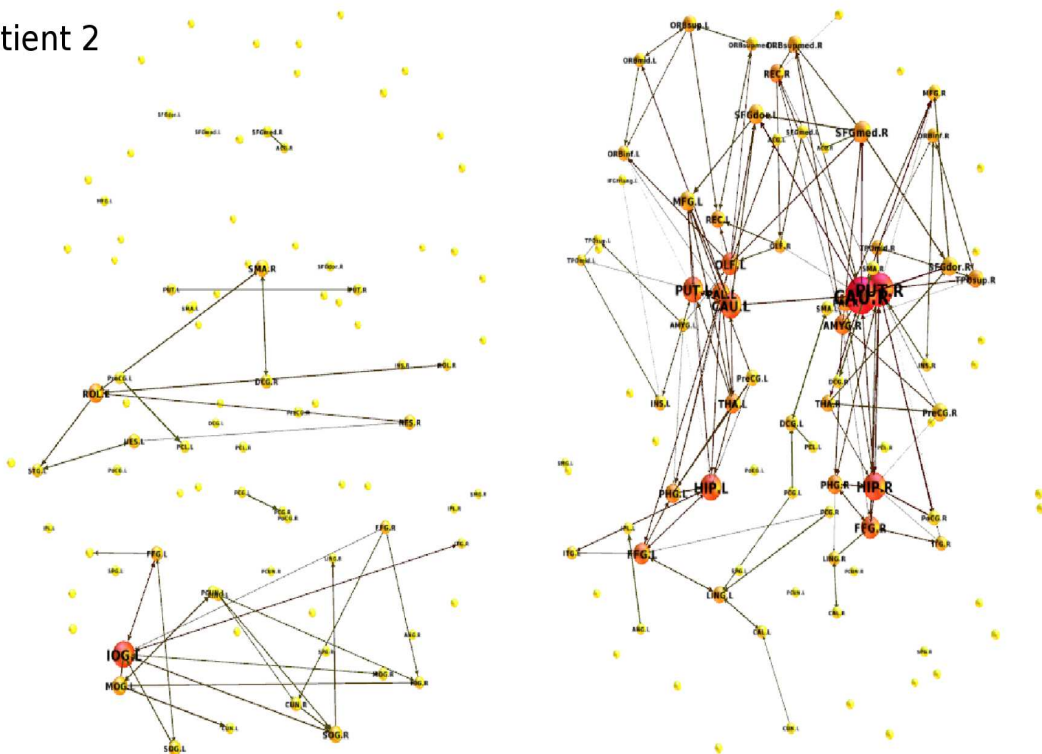


FIGURE 11.9 – Graphes en trois dimensions représentant les connexions fonctionnelles (à gauche) et anatomiques (à droite) supposées pathologiques. Nous pouvons observer que les connexions pathologiques ne se situent pas aux mêmes endroits. Plus précisément, nous observons de nombreuses déconnexions fonctionnelles à l'arrière du cerveau, tandis que les déconnexions anatomiques se situent plus au niveau des noyaux gris et à l'avant du cerveau. La discussion sur les différences observées se situe en section 2.3.

Conclusion

Dans cette troisième partie, nous nous sommes tout d'abord intéressés à la conception d'atlas en IRM. Nous avons remarqué que ces atlas étaient complexes. Cette complexité réside dans le fait que les grandeurs mathématiques considérées appartiennent généralement à des espaces non euclidiens de grande dimension. Différents auteurs ont commencé à introduire des techniques de réduction de dimension pour analyser ces modèles. L'idée sous-jacente est que les données forment une sous-variété dans un espace de plus grande dimension.

Dans un deuxième temps, nous avons effectué un état de l'art approfondi des modèles génératifs existant en imagerie médicale.

Nous avons ensuite proposé l'élaboration d'un nouvel atlas basé sur le SVDD permettant à la fois de détecter des données aberrantes et de localiser les zones pathologiques de ces données.

Finalement, nous avons proposé deux études dans le cadre des troubles de la conscience. Dans le premier atlas, nous avons proposé de coupler l'étude de la connectivité anatomique avec celle de l'intégrité des ROIs en considérant la FA. Dans un deuxième atlas, nous avons comparé les connectivités fonctionnelle et anatomique. La conception de ces deux atlas montre que les méthodologies pour l'élaboration d'atlas proposées dans cette thèse permettent de traiter un grand nombre d'objets mathématiques différents résidant dans des espaces non euclidiens de grande dimension.

Conclusion et perspectives

Nous récapitulons ici les principaux résultats et contributions propres à ce travail de thèse, tout en proposant des directions de recherche futures.

Synthèse des travaux

Dans le cadre de cette thèse, nous avons mis en place une chaîne de traitements pour la conception d'atlas en IRMd.

Dans une première partie, nous avons tout d'abord décrit la modalité d'IRMd. Cette modalité délivre des signaux complexes permettant d'étudier aussi bien des phénomènes à l'échelle du voxel que des mécanismes globaux à l'échelle du cerveau.

Dans un deuxième temps, nous avons étudié les outils existants rentrant en jeu dans la modélisation des différents signaux délivrés par l'IRMd. Nous nous sommes particulièrement intéressés aux tenseurs, ainsi qu'aux différentes métriques pour les comparer. Nous avons montré la richesse de ces modèles permettant d'obtenir un grand nombre de descripteurs.

Ensuite, nous avons présenté les modèles existants pour les phénomènes locaux, semi-globaux et globaux. Lors de cette présentation, nous avons pu observer la diversité des objets mathématiques utilisés. Ces objets peuvent aller d'une séquence de points formant des trajectoires à des graphes, en passant par des fonctions définies sur la sphère. Ces différents objets résident dans des espaces non euclidiens de grande dimension.

Nous avons ensuite étudié les méthodologies pour mettre en correspondance les différents modèles rencontrés. La complexité des modèles contraint d'effectuer des mises en correspondance spécifiques. Par exemple, pour les modèles locaux, des informations d'orientation doivent être prises en compte par rapport à la mise en correspondance d'images scalaires conventionnelles. Dans le cadre de cette thèse, nous avons proposé de nouvelles stratégies de réorientation pour les modèles locaux utilisant des tenseurs d'ordre 4, ainsi que des tenseurs d'ordre supérieur, pour prendre en compte ces informations d'orientation. Conceptuellement, l'originalité de cette méthodologie est l'utilisation de la décomposition d'un tenseur d'ordre 4 modélisant un croisement de paquets de fibres en des tenseurs d'ordre 2 modélisant les paquets de fibres séparément. La deuxième originalité réside dans l'hypothèse d'indépendance et dans le déplacement indépendant de chaque paquet.

Ces méthodologies peuvent être étendues aux modèles utilisant des HSs grâce au lien unissant les deux outils mathématiques. De cette manière, les réorientations proposées peuvent être utilisées pour l'ensemble des modèles basés sur ces outils.

Dans une troisième partie, nous avons tout d'abord effectué un état de l'art des atlas existant en IRMd. Cette étude a permis de se rendre compte de la complexité de la création d'atlas pour des objets de grande dimension résidant dans des espaces non euclidiens.

Ensuite, nous nous sommes intéressés à l'utilisation de modèles génératifs compacts pour la conception d'atlas en imagerie médicale.

Dans un deuxième temps, nous avons développé une nouvelle méthodologie pour la conception d'atlas. Cet atlas est basé sur des modèles compacts non linéaires. Il permet à la fois de discriminer des données aberrantes, ainsi que de déterminer les zones pathologiques. Ces deux types d'études sont complémentaires et sont basés sur des modèles discriminatif et génératif. Ces deux modèles sont générés grâce au même algorithme SVDD. Un des avantages de cet atlas est qu'il peut être utilisé pour un grand nombre de modèles différents. Plus précisément, l'atlas proposé nécessite de connaître une métrique adaptée entre les objets étudiés, mais ne nécessite pas de connaître l'espace où ces objets résident.

Cependant, comme pour la plupart des méthodes à noyaux, la paramétrisation du noyau pour le SVDD reste cruciale. Nous avons proposé des méthodes d'estimation automatique pour limiter l'impact de l'utilisateur dans la création de l'atlas. Ces différentes estimations peuvent être utilisées dans un cadre plus général que celui de la conception de l'atlas proposé.

Finalement, nous avons proposé d'étudier les connectomes. Pour montrer la capacité de nos atlas à prendre en compte des objets très différents, nous avons développé deux atlas différents. Le premier atlas considère simultanément les ROIs anatomiques et la connectivité anatomique, tandis que le deuxième atlas étudie les connectivités anatomique et fonctionnelle. D'un point de vue médical, le fait d'avoir un outil méthodologique permettant d'observer un problème sous différents aspects est très intéressant.

Dans le cadre des troubles de la conscience, nous proposons une description des connexions anatomiques et fonctionnelles pathologiques. Ces informations vont être maintenant étudiées d'un point de vue médical pour essayer de mieux comprendre les phénomènes de troubles de la conscience.

Perspectives

Méthodologiques

Les méthodologies pour la réorientation des tenseurs d'ordre 4 ont été validés sur des cas simulés. Une perspective à court terme serait de valider ces algorithmes de réorientation sur des cas réels.

Dans cette thèse, nous avons proposé des atlas conçus pour traiter un grand nombre de données différentes. Une autre perspective serait d'adapter ces atlas pour les modèles locaux et semi-globaux.

Dans les études sur les connectivités anatomique et fonctionnelle, la population de référence est différente. Si la population de référence avait eu une acquisition en IRMd et en IRMf de repos, on aurait pu concevoir un atlas prenant en compte simultanément les matrices de connectivité anatomique et fonctionnelle. De cette manière, on aurait pu voir l'existence de couplage entre les deux catégories de connectivité.

Théoriques

L'atlas proposé dans ce mémoire a été conçu pour un type spécifique de problème qui est la comparaison d'un individu par rapport à un groupe. Dans le cadre d'une comparaison d'un groupe par rapport à un autre, l'atlas développé privilégierait un groupe au détriment d'un autre. On peut donc chercher à étendre la méthodologie proposée dans cette thèse pour des

problèmes de comparaison de groupes.

Une autre perspective théorique serait d'utiliser une loi de probabilité $p(\mathbf{x})$ au lieu de SVDD pour modéliser les distributions dans l'espace réduit.

Médicales

D'un point de vue médical, une perspective à court terme serait d'étudier de manière approfondie les zones considérées comme anormales dans les connectomes des personnes atteintes de troubles de la conscience.

Les zones anormales sont différentes entre les connectivités anatomique et fonctionnelle. Une autre perspective serait d'étudier les mécanismes entraînant ces différences.

De plus, la compréhension des phénomènes de conscience reste complexe. Pour mieux interpréter ces phénomènes, plusieurs pistes sont possibles.

La première est d'étudier s'il existe des corrélations entre l'état clinique du patient diagnostiqué par un médecin et les zones pathologiques détectées par l'atlas. La deuxième piste serait d'étudier l'évolution des patients au cours du temps (*i.e.* d'effectuer un suivi longitudinal). Un grand intérêt serait porté aux patients étant sortis du coma. On pourrait alors essayer d'extraire des facteurs pronostics pour déterminer l'évolution du patient.

Quatrième partie

Bibliographie

Publications de l'auteur

Distinction

Prix des Meilleures Communications du Congrès de la SRLF 2011 pour la publication [12].

Publications relatives à la thèse

Articles dans des revues internationales à comité de lecture

[1] Kremer S., Renard F., Noblet V., Mialin R., Wolfram-Gabel R., Delon-Martin C., Achard S., Schenck M., Mohr M., Dietemann J.-L., Schneider F., « Diffusion tensor imaging in human global cerebral anoxia : correlation with histology in a case with autopsy », *Journal of Neuroradiology*, pp. 301-303, Vol. 37, Num. 5, 2010.

[2] Achard S., Kremer S., Schenck M., Renard F., Ong-Nicolas C., Namer I.-J., Mutschler V., Schneider F., Delon-Martin C., « Global functional disconnections in post-anoxic coma patient », *The Neuroradiology Journal*, pp. 269-273, Vol. 1, Num. 6, 2011.

[3] Chatelin S., Deck C., Renard F., Kremer S., Heinrich C., Armspach J.-P., Willinger R., « Computation of axonal elongation in head trauma finite element simulation », *Journal of the Mechanical Behavior of Biomedical Materials*, à paraître, 2011.

Articles dans des revues nationales à comité de lecture

[4] Blanc F., Noblet V., Jung B., Rousseau F., Renard F., Bourre B., Longato N., Cremel N., Di Bitonto L., Kleitz C., Collongues N., Foucher J., Kremer S., Armspach J.-P., de Seze J., « Neuromyéélite optique de Devic, troubles cognitifs et imagerie cérébrale par résonance magnétique », *Revue neurologique*, pp. 7-8, Vol. 167, Num. S1, 2011.

Conférences internationales avec actes et comité de lecture

[5] Boisgontier H., Noblet V., Renard F., Heitz F., Rumbach L., Armspach J.-P., « Statistical detection of longitudinal changes between apparent diffusion coefficient images. Application to multiple sclerosis », actes du congrès *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Londres, Royaume Uni, septembre 2009.

[6] Renard F., Noblet V., Heinrich C., Kremer S., « Reorientation strategies for high order tensors », actes du congrès *IEEE International Symposium on Biomedical Imaging (ISBI)*, Rotterdam, Pays-Bas, avril 2010.

[7] Achard S., Kremer S., Coeurjolly J.-F., Schenk M., Renard F., Ong-Nicolas C., Namer I.-J., Mutschler V., Schneider F., Delon-Martin C., « Global functional disconnections and long

memory alterations in consciousness disorder patients », actes du congrès *Human Brain Mapping* (HBM), Barcelone, Espagne, juin 2010.

[8] Renard F., Noblet V., Grigis A., Heinrich C., Kremer S., « Comparison of interpolation methods for angular resampling of diffusion weighted images », actes du congrès *IEEE International Conference on Image Processing Theory, Tools and Applications* (IPTA), Paris, France, juillet 2010.

[9] Grigis A., Noblet V., Renard F., Heitz F., Armspach J.-P., Rumbach L., « Change detection in diffusion MRI using multivariate statistical testing on tensors », actes du congrès *Medical Image Computing and Computer Assisted Intervention* (MICCAI), Pékin, Chine, septembre 2010.

[10] Grigis A., Renard F., Noblet V., Heinrich C., Heitz F., Armspach J.-P., « A new high order tensor decomposition : application to reorientation », actes du congrès *IEEE International Conference on International Symposium on Biomedical Imaging* (ISBI), Chicago, Etats-Unis, avril 2011.

Conférences nationales avec actes et comité de lecture

[11] Renard F., Heinrich C., Kremer S., « L'orthogonalisation simultanée appliquée à l'imagerie du tenseur de diffusion », actes du congrès *ORASIS'09*, Trégastel, France, juin 2009.

[12] Renard F., Delon-Martin C., Achard S., Noblet V., Heinrich C., Schenck M., Harlay M.-L., Diouf M., Namer I.-J., Schneider F., Kremer S., « Apport de l'IRM du tenseur de diffusion à l'exploration des malades cérébrolésés : corrélations anatomo-fonctionnelles », actes du congrès internationale de la Société de Réanimation de Langue Française (SRLF), Paris, France, janvier 2011.

Publications relatives à des travaux effectués avant cette thèse

Conférences internationales avec actes et comité de lecture

[13] Renard F., Yang Y., « Image segmentation for detection of soft plaques in multidetector CT images », *IEEE Southwest Symposium on Image Analysis and Interpretation* (SSIAI), Santa Fe, Etats-Unis, mars 2008.

[14] Renard F., Yang Y., « Image analysis for detection of coronary artery soft plaques in MDCT images », *IEEE International Conference on International Symposium on Biomedical Imaging* (ISBI), Paris, France, avril 2008.

[15] Renard F., Yang Y., « Coronary artery extraction and analysis for detection of soft plaques in MDCT images », *IEEE International Conference on Image Processing* (ICIP), San Diego, Etats-Unis, octobre 2008.

Conférences nationales avec actes et comité de lecture

[16] Renard F., Nguyen M.K., Sushchenko A., Daveau C., « Nouvelle méthode de contrôle non destructif exploitant le rayonnement gamma diffusé », actes du congrès Les Journées Cofrend, Toulouse, France, mai 2008.

Références bibliographiques

- [ACHARD & BULLMORE 2007] S. ACHARD & E. BULLMORE (2007), «Efficiency and cost of economical brain functional networks», *PLoS Computational Biology*, **3**, n° 2, pages 174–183.
- [ACHARD et coll. 2006] S. ACHARD, R. SALVADOR, B. WHITCHER, J. SUCKLING & E. BULLMORE (2006), «A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs», *Journal of Neuroscience*, **26**, n° 1, pages 63–72.
- [ALBERT & BARABÁSI 2002] R. ALBERT & A. BARABÁSI (2002), «Statistical mechanics of complex networks», *Reviews of Modern Physics*, **74**, n° 1, pages 47.
- [ALEXANDER et coll. 1999] D. ALEXANDER, J. GEE & R. BAJCSY (1999), «Similarity measures for matching diffusion tensor images», *Proceedings of the 10th British Machine Vision Conference*, pages 93–102.
- [ALEXANDER & GEE 2000] D. C. ALEXANDER & J. C. GEE (2000), «Elastic matching of diffusion tensor images», *Computer Vision and Image Understanding*, **77**, n° 2, pages 233–250.
- [ALEXANDER et coll. 2001] D. C. ALEXANDER, C. PIERPAOLI, P. J. BASSER & J. C. GEE (2001), «Spatial transformations of diffusion tensor magnetic resonance images», *IEEE Transactions on Medical Imaging*, **20**, n° 11, pages 1131–1139.
- [ALJABAR et coll. 2010] P. ALJABAR, R. WOLZ, L. SRINIVASAN, S. COUNSELL, J. P. BOARDMAN, M. MURGASOVA, V. DORIA, M. A. RUTHERFORD, A. D. EDWARDS, J. V. HAJNAL & D. RUECKERT (2010), «Combining morphological information in a manifold learning framework : application to neonatal MRI», *Proceedings of MICCAI*, **6363**, n° 3, pages 1–8.
- [ANDERSON 2005] A. W. ANDERSON (2005), «Measurement of fiber orientation distributions using high angular resolution diffusion imaging», *Magnetic Resonance in Medicine*, **54**, n° 5, pages 1194–1206.
- [ARSIGNY 2006] V. ARSIGNY (2006), *Processing data in Lie groups : an algebraic approach. Application to non-linear registration and diffusion tensor MRI.*, Thèse de doctorat, Ecole Polytechnique.
- [ARSIGNY et coll. 2006] V. ARSIGNY, P. FILLARD, X. PENNEC & N. AYACHE (2006), «Log-Euclidean metrics for fast and simple calculus on diffusion tensors», *Magnetic Resonance in Medicine*, **56**, n° 2, pages 411–421.
- [ARSIGNY et coll. 2007] V. ARSIGNY, P. FILLARD, X. PENNEC & N. AYACHE (2007), «Geometric means in a novel vector space structure on symmetric positive-definite matrices», *SIAM Journal on Matrix Analysis and Applications*, **29**, n° 1, pages 328–347.
- [ASSEMBLAL 2010] H. ASSEMBLAL (2010), *Traitement et analyse d'images IRM de diffusion pour l'estimation de l'architecture locale des tissus*, Thèse de doctorat, Université de Caen.

- [ASSEMLAL et coll. 2011] H. ASSEMLAL, D. TSCHUMPERLÉ, L. BRUN & K. SIDDIQI (2011), «Recent advances in diffusion MRI modeling : angular and radial reconstruction», *Medical Image Analysis*, **15**, n° 4, pages 369–396.
- [BANERJEE & DIEHL 2006] A. BANERJEE & C. DIEHL (2006), «A support vector method for anomaly detection in hyperspectral imagery», *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pages 2282–2291.
- [BANKMAN 2008] I. BANKMAN (2008), *Handbook of medical image processing and analysis*, Academic Press, deuxième édition.
- [BANSAL et coll. 2007] R. BANSAL, L. H. STAIB, K. J. PLESSSEN, D. XU, J. ROYAL & B. S. PETERSON (2007), «Voxel-wise comparisons of the morphology of diffusion tensors across groups of experimental subjects», *Psychiatry Research : Neuroimaging*, **156**, n° 3, pages 225–245.
- [BARBIERI et coll. 2008] S. BARBIERI, M. WELK & J. WEICKERT (2008), «Variational registration of tensor-valued images», *Proceedings of the IEEE CVPR*, pages 1–6.
- [BARMPOUTIS & VEMURI 2009] A. BARMPOUTIS & B. C. VEMURI (2009), «Groupwise registration and atlas construction of 4th-order tensor fields using the \mathbb{R}^+ Riemannian metric», *Proceedings of MICCAI*, **5761**, pages 640–647.
- [BARMPOUTIS & ZHUO 2011] A. BARMPOUTIS & J. ZHUO (2011), «Diffusion kurtosis imaging : robust estimation from DW-MRI using homogeneous polynomials», *Proceedings of the IEEE ISBI*, pages 262–265.
- [BARMPOUTIS et coll. 2007] A. BARMPOUTIS, B. C. VEMURI & J. R. FORDER (2007), «Registration of high angular resolution diffusion MRI images using 4th order tensors», *Proceedings of MICCAI*, **4791**, n° 1, pages 908–915.
- [BARMPOUTIS et coll. 2009] A. BARMPOUTIS, B. C. VEMURI, D. HOWLAND & J. R. FORDER (2009), «Regularized positive-definite fourth order tensor field estimation from DW-MRI», *NeuroImage*, **45**, n° 1 Suppl, pages S153–S162.
- [BASSER et coll. 1994] P. BASSER, J. MATTIELLO & D. LEBIHAN (1994), «MR diffusion tensor spectroscopy and imaging», *Biophysical Journal*, **66**, n° 1, pages 259–267.
- [BASSER & PAJEVIC 2007] P. J. BASSER & S. PAJEVIC (2007), «Spectral decomposition of a 4th-order covariance tensor : applications to diffusion tensor MRI», *Signal Processing*, **87**, pages 220–236.
- [BASSER et coll. 2000] P. J. BASSER, S. PAJEVIC, C. PIERPAOLI, J. DUDA & A. ALDROUBI (2000), «In vivo fiber tractography using DT-MRI data», *Magnetic Resonance in Medicine*, **44**, n° 4, pages 625–632.
- [BASSETT & BULLMORE 2006] D. S. BASSETT & E. BULLMORE (2006), «Small-world brain networks», *Neuroscientist*, **12**, n° 6, pages 512–523.
- [BASSETT & BULLMORE 2009] D. S. BASSETT & E. T. BULLMORE (2009), «Human brain networks in health and disease», *Current Opinion in Neurology*, **22**, n° 4, pages 340–347.
- [BASSETT et coll. 2008] D. S. BASSETT, E. BULLMORE, B. A. VERCHINSKI, V. S. MATTAY, D. R. WEINBERGER & A. MEYER-LINDENBERG (2008), «Hierarchical organization of human cortical networks in health and schizophrenia», *Journal of Neuroscience*, **28**, n° 37, pages 9239–9248.
- [BATCHELOR et coll. 2005] P. G. BATCHELOR, M. MOAKHER, D. ATKINSON, F. CALAMANTE & A. CONNELLY (2005), «A rigorous framework for diffusion tensor calculus», *Magnetic Resonance in Medicine*, **53**, n° 1, pages 221–225.
- [BEAULIEU 2002] C. BEAULIEU (2002), «The basis of anisotropic water diffusion in the nervous system— a technical review», *NMR in Biomedicine*, **15**, n° 7-8, pages 435–455.

- [BEHRENS et coll. 2003] T. E. BEHRENS, M. W. WOOLRICH, M. JENKINSON, H. JOHANSEN-BERG, R. G. NUNES, S. CLARE, P. M. MATTHEWS, J. M. BRADY & S. M. SMITH (2003), «Characterization and propagation of uncertainty in diffusion weighted MR imaging», *Magnetic Resonance in Medicine*, **50**, n° 5, pages 1077–1088.
- [BEHRENS et coll. 2003] T. E. J. BEHRENS, H. JOHANSEN-BERG, M. W. WOOLRICH, S. M. SMITH, C. A. M. WHEELER-KINGSHOTT, P. A. BOULBY, G. J. BARKER, E. L. SILLERY, K. SHEEHAN, O. CICCARELLI, A. J. THOMPSON, J. M. BRADY & P. M. MATTHEWS (2003), «Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging», *Nature Neuroscience*, **6**, n° 7, pages 750–757.
- [BELLMAN 1961] R. BELLMAN (1961), *Adaptive control processes*, Princeton University Press.
- [BERGMANN et coll. 2007] Ø. BERGMANN, G. KINDLMANN, S. PELED & C. WESTIN (2007), «Two-tensor fiber tractography», *Proceedings of the IEEE ISBI*, pages 796–799.
- [BLEZEK & MILLER 2007] D. J. BLEZEK & J. V. MILLER (2007), «Atlas stratification», *Medical Image Analysis*, **11**, n° 5, pages 443–457.
- [BLOY & VERMA 2010] L. BLOY & R. VERMA (2010), «Demons registration of high angular resolution diffusion images», *Proceedings of the IEEE ISBI*, pages 1013–1016.
- [BOISGONTIER 2010] H. BOISGONTIER (2010), *Détection automatique de changements en IRM de diffusion : application à la sclérose en plaques.*, Thèse de doctorat, Université de Strasbourg.
- [BOUCHARD 2005] G. BOUCHARD (2005), *Les modèles génératifs en classification supervisée et applications à la catégorisation d'images et à la fiabilité industrielle*, Thèse de doctorat, Université Joseph-Fourier - Grenoble I.
- [BOUX et coll. 2010] S. BOUX, Y. RATHI & M. SABUNCU (2010), «Building an average population HARDI atlas», *Workshop MICCAI on computational diffusion MRI*.
- [BRERETON 2011] R. G. BRERETON (2011), «One class classifiers», *Journal of Chemometrics*, **25**, n° 5, pages 225–246.
- [BRERETON & LLOYD 2010] R. G. BRERETON & G. R. LLOYD (2010), «Support vector machines for classification and regression», *Analyst*, **135**, pages 230–267.
- [BRUCHER 2008] M. BRUCHER (2008), *Représentations compactes et apprentissage non supervisé de variétés non linéaires. Application au traitement d'images.*, Thèse de doctorat, Université Louis Pasteur - Strasbourg I.
- [BULLITT et coll. 2005] E. BULLITT, D. ZENG, G. GERIG, S. AYLWARD, S. JOSHI, J. K. SMITH, W. LIN & M. G. EWEND (2005), «Vessel tortuosity and brain tumor malignancy : a blinded study», *Academic Radiology*, **12**, n° 10, pages 1232–1240.
- [BULLMORE & SPORNS 2009] E. BULLMORE & O. SPORNS (2009), «Complex brain networks : graph theoretical analysis of structural and functional systems», *Nature Reviews Neuroscience*, **10**, n° 3, pages 186–198.
- [BULLMORE & BASSETT 2011] E. T. BULLMORE & D. S. BASSETT (2011), «Brain graphs : graphical models of the human brain connectome», *Annual Review of Clinical Psychology*, **7**, n° 1, pages 113–140.
- [CAAN et coll. 2009] M. CAAN, C. SAGE, M. VAN DER GRAAF, C. GRIMBERGEN, S. SUNAERT, L. VAN VLIET & F. VOS (2009), «Dual tensor atlas generation based on a cohort of coregistered non-HARDI datasets», *Proceedings of MICCAI*, **5761**, pages 869–876.
- [CALLAGHAN et coll. 1990] CALLAGHAN, MACGOWAN, PACKER & F. ZELAYA (1990), «High-resolution q-space imaging in porous structures», *Journal of Magnetic Resonance*, **90**, pages 177–182.

- [CAMPBELL et coll. 2005] J. S. CAMPBELL, K. SIDDIQI, V. V. RYMAR, A. F. SADIKOT & G. B. PIKE (2005), «Flow-based fiber tracking with diffusion tensor and q-ball data : validation and comparison to principal diffusion direction techniques», *NeuroImage*, **27**, n° 4, pages 725–736.
- [CAO et coll. 2005] Y. CAO, M. MILLER, R. WINSLOW & L. YOUNES (2005), «Large deformation diffeomorphic metric mapping of vector fields», *IEEE Transactions on Medical Imaging*, **24**, n° 9, pages 1216–1230.
- [CATANI 2006] M. CATANI (2006), «Diffusion tensor magnetic resonance imaging tractography in cognitive disorders», *Current Opinion in Neurology*, **19**, n° 6, pages 599–606.
- [CHANDOLA et coll. 2009] V. CHANDOLA, A. BANERJEE & V. KUMAR (2009), «Anomaly detection : a survey», *ACM Computing Surveys*, **41**, n° 3, pages 15 :1–15 :58.
- [CHAPELLE et coll. 2002] O. CHAPELLE, V. VAPNIK, O. BOUSQUET & S. MUKHERJEE (2002), «Choosing multiple parameters for support vector machines», *Machine Learning*, **46**, pages 131–159.
- [CHAPPELL et coll. 2008] M. H. CHAPPELL, J. A. BROWN, J. C. DALRYMPLE-ALFORD, A. M. ULUĞ & R. WATTS (2008), «Multivariate analysis of diffusion tensor imaging data improves the detection of microstructural damage in young professional boxers», *Magnetic Resonance Imaging*, **26**, n° 10, pages 1398–1405.
- [CHATELIN 2010] S. CHATELIN (2010), *Contribution à la compréhension des mécanismes de lésion cérébrale en situation de choc : inclusion de l'anisotropie et de l'hétérogénéité par techniques d'imagerie médicale.*, Thèse de doctorat, Université de Strasbourg.
- [CHATELIN et coll. 2011] S. CHATELIN, C. DECK, F. RENARD, S. KREMER, C. HEINRICH, J. ARMSPACH & R. WILLINGER (2011), «Computation of axonal elongation in head trauma finite element simulation», *Journal of the Mechanical Behavior of Biomedical Materials*, **Sous presse, à paraître**.
- [CHENG et coll. 2009] J. CHENG, A. GHOSH, T. JIANG & R. DERICHE (2009), «A Riemannian framework for orientation distribution function computing», *Proceedings of MICCAI*, **5761**, pages 911–918.
- [CHEUNG et coll. 2009] M. M. CHEUNG, E. S. HUI, K. C. CHAN, J. A. HELPERN, L. QI & E. X. WU (2009), «Does diffusion kurtosis imaging lead to better neural tissue characterization? A rodent brain maturation study», *NeuroImage*, **45**, n° 2, pages 386–392.
- [CHIANG et coll. 2007] M. CHIANG, A. D. KLUNDER, K. MCMAHON, G. I. DE ZUBICARAY, M. J. WRIGHT, A. W. TOGA & P. M. THOMPSON (2007), «Information-theoretic analysis of brain white matter fiber orientation distribution functions», *Proceedings of the IPMI*, **20**, pages 172–182.
- [CICCARELLI et coll. 2008] O. CICCARELLI, M. CATANI, H. JOHANSEN-BERG, C. CLARK & A. THOMPSON (2008), «Diffusion-based tractography in neurological disorders : concepts, applications, and future developments», *The Lancet Neurology*, **7**, n° 8, pages 715–727.
- [CONTURO et coll. 1999] T. E. CONTURO, N. F. LORI, T. S. CULL, E. AKBUDAK, A. Z. SNYDER, J. S. SHIMONY, R. C. MCKINSTRY, H. BURTON & M. E. RAICHEL (1999), «Tracking neuronal fiber pathways in the living human brain», *Proceedings of the National Academy of Sciences of the United States of America*, **96**, n° 18, pages 10422–10427.
- [COROUGE et coll. 2006] I. COROUGE, P. T. FLETCHER, S. JOSHI, S. GOUTTARD & G. GERIG (2006), «Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis», *Medical Image Analysis*, **10**, n° 5, pages 786–798.
- [CRANK 1980] J. CRANK (1980), *The mathematics of diffusion*, Oxford University Press, USA, deuxième édition.

- [CURRAN & ALEXANDER 2005] K. CURRAN & D. ALEXANDER (2005), «Comparison of similarity measures for driving diffusion tensor registration», *Proceedings of the International Society for Magnetic Resonance in Medicine*, pages 226 – 226.
- [DAVOODI-BOJD & SOLTANIAN-ZADEH 2008] E. DAVOODI-BOJD & H. SOLTANIAN-ZADEH (2008), «Atlas based segmentation of white matter fiber bundles in DT-MRI using fractional anisotropy and principal eigen vectors», *Proceedings of the IEEE ISBI*, pages 879–882.
- [DE BOER et coll. 2011] R. DE BOER, M. SCHAAP, F. VAN DER LIJN, H. A. VROOMAN, M. DE GROOT, A. VAN DER LUGT, M. A. IKRAM, M. W. VERNOOIJ, M. M. BRETELER & W. J. NIESSEN (2011), «Statistical analysis of minimum cost path based structural brain connectivity», *NeuroImage*, **55**, n° 2, pages 557–565.
- [DE LATHAUWER et coll. 2000] L. DE LATHAUWER, B. DE MOOR & J. VANDEWALLE (2000), «Fetal electrocardiogram extraction by blind source subspace separation», *IEEE Transactions on Biomedical Engineering*, **47**, n° 5, pages 567–572.
- [DELIGIANNI et coll. 2010] F. DELIGIANNI, E. C. ROBINSON, C. F. BECKMANN, D. SHARP, A. D. EDWARDS & D. RUECKERT (2010), «Inference of functional connectivity from structural brain connectivity», *Proceedings of the IEEE ISBI*, page 1113–1116.
- [DEMIRALP & LAIDLAW 2010] C. DEMIRALP & D. H. LAIDLAW (2010), «Tract-based probability densities of diffusivity measures in DT-MRI», *Proceedings of MICCAI*, **6361**, n° 1, pages 542–549.
- [DESCOTEAUX 2008] M. DESCOTEAUX (2008), *High angular resolution diffusion MRI : from local estimation to segmentation and tractography*, Thèse de doctorat, Université de Nice - Sophia Antipolis.
- [DESCOTEAUX et coll. 2006] M. DESCOTEAUX, E. ANGELINO, S. FITZGIBBONS & R. DERICHE (2006), «Apparent diffusion coefficients from high angular resolution diffusion imaging : estimation and applications», *Magnetic Resonance in Medicine*, **56**, n° 2, pages 395–410.
- [DESCOTEAUX et coll. 2007] M. DESCOTEAUX, E. ANGELINO, S. FITZGIBBONS & R. DERICHE (2007), «Regularized, fast, and robust analytical q-ball imaging», *Magnetic Resonance in Medicine*, **58**, n° 3, pages 497–510.
- [DÍAZ et coll. 2010] G. DÍAZ, G. PAJARES, E. ROMERO, J. ALVAREZ-LINERA, E. LÓPEZ, J. HERNÁNDEZ-TAMAMES & N. MALPICA (2010), «The effect of the normalization strategy on voxel-based analysis of DTI images : a pattern recognition based assessment», *Proceedings of the 2010 international conference on Brain informatics*, **6334**, pages 78–88.
- [DRYDEN et coll. 2009] I. L. DRYDEN, A. KOLOYDENKO & D. ZHOU (2009), «Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging», *Annals of Applied Statistics*, **3**, n° 3, pages 1102–1123.
- [DUDA et coll. 2003] J. T. DUDA, M. RIVERA, D. C. ALEXANDER & J. C. GEE (2003), «A method for non-rigid registration of diffusion tensor magnetic resonance images», volume 5032, pages 1186–1196, SPIE, 2003.
- [DYRBY et coll. 2007] T. B. DYRBY, L. V. SØGAARD, G. J. PARKER, D. C. ALEXANDER, N. M. LIND, W. F. BAARÉ, A. HAY-SCHMIDT, N. ERIKSEN, B. PAKKENBERG, O. B. PAULSON & J. JELSING (2007), «Validation of in vitro probabilistic tractography», *NeuroImage*, **37**, n° 4, pages 1267–1277.
- [ENGERT & BONHOEFFER 1999] F. ENGERT & T. BONHOEFFER (1999), «Dendritic spine changes associated with hippocampal long-term synaptic plasticity», *Nature*, **399**, n° 6731, pages 66–70.

- [ESSEN et coll. 1992] D. V. ESSEN, C. ANDERSON & D. FELLEMAN (1992), «Information processing in the primate visual system : an integrated systems perspective», *Science*, **255**, n° 5043, pages 419–423.
- [EVANGELISTA et coll. 2007] P. F. EVANGELISTA, M. J. EMBRECHTS & B. K. SZYMANSKI (2007), «Some properties of the Gaussian kernel for one class learning», *Proceedings of the 17th international conference on Artificial neural networks-ICANN'07*, **4668**, n° 1, pages 269–278.
- [FICK 1855] A. FICK (1855), «On liquid diffusion», *Philosophical Magazine*, **10**, pages 30–39.
- [FILLARD 2008] P. FILLARD (2008), *Traitement riemannien des tenseurs pour l'IRM de diffusion et l'anatomie algorithmique du cerveau*, Thèse de doctorat, Université de Nice – Sophia Antipolis.
- [FILLARD et coll. 2007] P. FILLARD, X. PENNEC, V. ARSIGNY & N. AYACHE (2007), «Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics», *IEEE Transactions on Medical Imaging*, **26**, n° 11, pages 1472–1482.
- [FILLARD et coll. 2009] P. FILLARD, C. POUPON & J.-F. MANGIN (2009), «A novel global tractography algorithm based on an adaptive spin glass model», *Proceedings of MICCAI*, **5761**, n° 1, pages 927–934.
- [FILLARD et coll. 2011] P. FILLARD, M. DESCOTEAUX, A. GOH, S. GOUTTARD, B. JEURISSEN, J. MALCOLM, A. RAMIREZ-MANZANARES, M. REISERT, K. SAKAIE, F. TENSAOUTI, T. YO, J.-F. MANGIN & C. POUPON (2011), «Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom», *NeuroImage*, **56**, n° 1, pages 220–234.
- [FLETCHER & JOSHI 2004] P. T. FLETCHER & S. JOSHI (2004), «Principal geodesic analysis on symmetric spaces : statistics of diffusion tensors», *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*, **3117**, pages 87–98.
- [FLETCHER et coll. 2007] P. T. FLETCHER, R. TAO, W. JEONG & R. T. WHITAKER (2007), «A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor MRI», *Proceedings of the IPMI*, **20**, pages 346–358.
- [FORTE & VIANELLO 1996] S. FORTE & M. VIANELLO (1996), «Symmetry classes for elasticity tensors», *Journal of Elasticity*, **43**, n° 2, pages 81–108.
- [GARRIGUES 2007] J. GARRIGUES (2007), *Fondements de la mécanique des milieux continus*, Hermes Science Publications.
- [GEE & ALEXANDER 2006] J. C. GEE & D. C. ALEXANDER (2006), «Diffusion-tensor image registration», dans J. WEICKERT & H. HAGEN, éditeurs, *Visualization and Processing of Tensor Fields*, Mathematics and Visualization, pages 327–342, Springer.
- [GEE et coll. 2002] J. C. GEE, D. C. ALEXANDER, M. RIVERA & J. DUDA (2002), «Non-rigid registration of diffusion tensor MR images», *Proceedings of the IEEE ISBI*, pages 477–480.
- [GENG et coll. 2009] X. GENG, T. J. ROSS, W. ZHAN, H. GU, Y. CHAO, C. LIN, G. E. CHRISTENSEN, N. SCHUFF & Y. YANG (2009), «Diffusion MRI registration using orientation distribution functions», *Proceedings of the IPMI*, **21**, pages 626–637.
- [GERBER et coll. 2010] S. GERBER, T. TASDIZEN, P. THOMAS FLETCHER, S. JOSHI & R. WHITAKER (2010), «Manifold modeling for brain population analysis», *Medical Image Analysis*, **14**, n° 5, pages 643–653.
- [GERIG et coll. 2004] G. GERIG, S. GOUTTARD & I. COROUGE (2004), «Analysis of brain white matter via fiber tract modeling», *Proceedings of IEEE EMBS*, **6**, pages 4421–4424.

- [GHOSH et coll. 2008] A. GHOSH, M. DESCOTEAUX & R. DERICHE (2008), «Riemannian framework for estimating symmetric positive definite 4th order diffusion tensors», *Proceedings of MICCAI*, **5241**, n° 1, pages 858–865.
- [GOLLAND et coll. 2005] P. GOLLAND, W. E. L. GRIMSON, M. E. SHENTON & R. KIKINIS (2005), «Detection and analysis of statistical differences in anatomical shape», *Medical Image Analysis*, **9**, n° 1, pages 69–86.
- [GOODLETT et coll. 2009] C. B. GOODLETT, P. T. FLETCHER, J. H. GILMORE & G. GERIG (2009), «Group analysis of DTI fiber tract statistics with application to neurodevelopment», *NeuroImage*, **45**, n° 1, pages S133–S142.
- [GRIGIS et coll. 2011] A. GRIGIS, F. RENARD, V. NOBLET, C. HEINRICH, F. HEITZ & J.-P. ARMSPACH (2011), «A new high order tensor decomposition : application to reorientation.», dans *Proceedings of the IEEE ISBI*, pages 258–261, mars 2011.
- [GROEMER 1996] H. GROEMER (1996), *Geometric applications of Fourier series and spherical harmonics*, Cambridge University Press.
- [GUIMOND et coll. 2002] A. GUIMOND, C. R. GUTTMANN, S. K. WARFIELD & C. F. WESTIN (2002), «Deformable registration of DT-MRI data based on transformation invariant tensor characteristics», *Proceedings of the IEEE ISBI*, pages 761– 764.
- [HAGMANN 2005] P. HAGMANN (2005), *From diffusion MRI to brain connectomics*, Thèse de doctorat, École polytechnique fédérale de Lausanne.
- [HAGMANN et coll. 2004] P. HAGMANN, T. REESE, W. TSENG, R. MEULI, J. THIRAN & V. WEDEEN (2004), «Diffusion spectrum imaging tractography in complex cerebral white matter : an investigation of the centrum semiovale», *Proceedings of the International Society for Magnetic Resonance in Medicine*.
- [HAGMANN et coll. 2008] P. HAGMANN, L. CAMMOUN, X. GIGANDET, R. MEULI, C. J. HONEY, V. J. WEDEEN & O. SPORNS (2008), «Mapping the structural core of human cerebral cortex», *PLoS Biology*, **6**, n° 7, pages e159.
- [HAGMANN et coll. 2010] P. HAGMANN, L. CAMMOUN, X. GIGANDET, S. GERHARD, P. E. GRANT, V. WEDEEN, R. MEULI, J. THIRAN, C. J. HONEY & O. SPORNS (2010), «MR connectomics : principles and challenges», *Journal of Neuroscience Methods*, **194**, n° 1, pages 34–45.
- [HASTIE et coll. 2009] T. HASTIE, R. TIBSHIRANI & J. H. FRIEDMAN (2009), *The elements of statistical learning*, Springer, deuxième édition.
- [HE et coll. 2007] Y. HE, Z. J. CHEN & A. C. EVANS (2007), «Small-World anatomical networks in the human brain revealed by cortical thickness from MRI», *Cerebral Cortex*, **17**, n° 10, pages 2407–2419.
- [HE et coll. 2008] Y. HE, Z. CHEN & A. EVANS (2008), «Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer’s disease», *The Journal of Neuroscience*, **28**, n° 18, pages 4756 –4766.
- [HE et coll. 2009] Y. HE, A. DAGHER, Z. CHEN, A. CHARIL, A. ZIJDENBOS, K. WORSLEY & A. EVANS (2009), «Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load», *Brain*, **132**, n° 12, pages 3366–3379.
- [HECKE et coll. 2011] W. V. HECKE, A. LEEMANS, C. A. SAGE, L. Emsell, J. VERAART, J. SIJBERS, S. SUNAERT & P. M. PARIZEL (2011), «The effect of template selection on diffusion tensor voxel-based analysis results», *NeuroImage*, **55**, n° 2, pages 566 - 573.
- [HESS et coll. 2006] C. P. HESS, P. MUKHERJEE, E. T. HAN, D. XU & D. B. VIGNERON (2006), «Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis», *Magnetic Resonance in Medicine*, **56**, n° 1, pages 104–117.

- [HOEFT et coll. 2007] F. HOEFT, N. BARNEA-GORALY, B. W. HAAS, G. GOLARAI, D. NG, D. MILLS, J. KORENBERG, U. BELLUGI, A. GALABURDA & A. L. REISS (2007), «More is not always better : increased fractional anisotropy of superior longitudinal fasciculus associated with poor visuospatial abilities in Williams syndrome», *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, **27**, n° 44, pages 11960–11965.
- [HONEY et coll. 2009] C. J. HONEY, O. SPORNS, L. CAMMOUN, X. GIGANDET, J. P. THIRAN, R. MEULI & P. HAGMANN (2009), «Predicting human resting-state functional connectivity from structural connectivity», *Proceedings of the National Academy of Sciences*, **106**, n° 6, pages 2035–2040.
- [HONG et coll. 2009] X. HONG, L. R. ARLINGHAUS & A. W. ANDERSON (2009), «Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging data», *Magnetic Resonance in Medicine*, **61**, n° 6, pages 1520–1527.
- [HUA et coll. 2008] K. HUA, J. ZHANG, S. WAKANA, H. JIANG, X. LI, D. S. REICH, P. A. CALABRESI, J. J. PEKAR, P. C. M. VAN ZIJL & S. MORI (2008), «Tract probability maps in stereotaxic spaces : analyses of white matter anatomy and tract-specific quantification», *NeuroImage*, **39**, n° 1, pages 336–347.
- [HUMPHRIES & GURNEY 2008] M. D. HUMPHRIES & K. GURNEY (2008), «Network ‘small-world-ness’ : a quantitative method for determining canonical network equivalence», *PLoS One*, **3**, n° 4.
- [ITURRIA-MEDINA et coll. 2008] Y. ITURRIA-MEDINA, R. C. SOTERO, E. J. CANALES-RODRÍGUEZ, Y. ALEMÁN-GÓMEZ & L. MELIE-GARÍCA (2008), «Studying the human brain anatomical network via diffusion-weighted MRI and graph theory», *NeuroImage*, **40**, n° 3, pages 1064 - 1076.
- [ITURRIA-MEDINA et coll. 2011] Y. ITURRIA-MEDINA, A. PÉREZ FERNÁNDEZ, P. VALDÉS HERNÁNDEZ, L. GARCÍA PENTÓN, E. J. CANALES-RODRÍGUEZ, L. MELIE-GARCIA, A. L. CASTELLANOS & M. O. ORTEGA (2011), «Automated discrimination of brain pathological state attending to complex structural brain network properties : the shiverer mutant mouse case», *PLoS ONE*, **6**, n° 5.
- [JANSONS & ALEXANDER 2003] K. M. JANSONS & D. C. ALEXANDER (2003), «Persistent angular structure : new insights from diffusion magnetic resonance imaging data», *Inverse Problems*, **19**, n° 5, pages 1031–1046.
- [JBABDI et coll. 2004] S. JBABDI, P. BELLEC, G. MARRELEC, V. PERLBARG & H. BENALI (2004), «A level set method for building anatomical connectivity paths between brain areas using DTI», *Proceedings of the IEEE ISBI*, pages 1024–1027.
- [JBABDI et coll. 2007] S. JBABDI, M. WOOLRICH, J. ANDERSSON & T. BEHRENS (2007), «A Bayesian framework for global tractography», *NeuroImage*, **37**, n° 1, pages 116–129.
- [JBABDI et coll. 2008] S. JBABDI, P. BELLEC, R. TORO, J. DAUNIZEAU, M. PÉLÉGRINI-ISSAC & H. BENALI (2008), «Accurate anisotropic fast marching for diffusion-based geodesic tractography», *International Journal of Biomedical Imaging*.
- [JENSEN et coll. 2005] J. H. JENSEN, J. A. HELPERN, A. RAMANI, H. LU & K. KACZYNSKI (2005), «Diffusional kurtosis imaging : the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging», *Magnetic Resonance in Medicine*, **53**, n° 6, pages 1432–1440.
- [JIAN & VEMURI 2007] B. JIAN & B. C. VEMURI (2007), «A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI», *IEEE Transactions on Medical Imaging*, **26**, n° 11, pages 1464–1471.
- [JONES 2008] D. JONES (2008), «Studying connections in the living human brain with diffusion MRI», *Cortex*, **44**, n° 8, pages 936–952.

- [JONES & CERCIGNANI 2010] D. K. JONES & M. CERCIGNANI (2010), «Twenty five pitfalls in the analysis of diffusion MRI data», *NMR in Biomedicine*, **23**, n° 7, pages 803–820.
- [JONES & PIERPAOLI 2005] D. K. JONES & C. PIERPAOLI (2005), «Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach», *Magnetic Resonance in Medicine*, **53**, n° 5, pages 1143–1149.
- [JONES et coll. 2002] D. K. JONES, L. D. GRIFFIN, D. C. ALEXANDER, M. CATANI, M. A. HORSFIELD, R. HOWARD & S. C. WILLIAMS (2002), «Spatial normalization and averaging of diffusion tensor MRI data sets», *NeuroImage*, **17**, n° 2, pages 592 - 617.
- [JONES et coll. 2007] D. K. JONES, X. A. CHITIS, D. JOB, P. L. KHONG, L. T. LEUNG, S. MARENCO, S. M. SMITH, M. R. SYMMS, D. K. JONES, X. A. CHITIS, D. JOB, P. L. KHONG, L. T. LEUNG, S. MARENCO, S. M. SMITH & M. R. SYMMS (2007), «What happens when nine different groups analyze the same DT-MRI data set using voxel-based methods?», *Proceedings of the International Society for Magnetic Resonance in Medicine*.
- [KHURD et coll. 2007] P. KHURD, R. VERMA & C. DAVATZIKOS (2007), «Kernel-based manifold learning for statistical analysis of diffusion tensor images», *Proceedings of the IPMI*, **4584**, pages 581–593.
- [KHURD et coll. 2007] P. KHURD, S. BALOCH, R. GUR, C. DAVATZIKOS & R. VERMA (2007), «Manifold learning techniques in image analysis of high-dimensional diffusion tensor magnetic resonance images», *Proceedings of the IEEE CVPR*, pages 1–7.
- [KOCH et coll. 2002] M. A. KOCH, D. G. NORRIS & M. HUND-GEORGIADIS (2002), «An investigation of functional and anatomical connectivity using magnetic resonance imaging», *NeuroImage*, **16**, n° 1, pages 241–250.
- [KÖTTER 2007] R. KÖTTER (2007), «Anatomical concepts of brain connectivity», dans V. JIRSA & A. R. MCINTOSH, éditeurs, *Handbook of Brain Connectivity*, volume 12 de la collection *Understanding Complex Systems*, pages 149–167, Springer.
- [KREHER et coll. 2005] B. W. KREHER, J. F. SCHNEIDER, I. MADER, E. MARTIN, J. HENNIG & K. A. IL'YASOV (2005), «Multitensor approach for analysis and tracking of complex fiber configurations», *Magnetic Resonance in Medicine*, **54**, n° 5, pages 1216–1225.
- [KREHER et coll. 2008] B. W. KREHER, I. MADER & V. G. KISELEV (2008), «Gibbs tracking : A novel approach for the reconstruction of neuronal pathways», *Magnetic Resonance in Medicine*, **60**, n° 4, pages 953–963.
- [KUBICKI et coll. 2002] M. KUBICKI, C. WESTIN, S. E. MAIER, H. MAMATA, M. FRUMIN, H. ERSNER-HERSHFIELD, R. KIKINIS, F. A. JOLESZ, R. MCCARLEY & M. E. SHENTON (2002), «Diffusion tensor imaging and its application to neuropsychiatric disorders», *Harvard Review of Psychiatry*, **10**, n° 6, pages 324–336.
- [LATORA & MARCHIORI 2001] V. LATORA & M. MARCHIORI (2001), «Efficient behavior of small-world networks», *Physical Review Letters*, **87**, n° 19, pages 198701.
- [LAWES et coll. 2008] I. N. C. LAWES, T. R. BARRICK, V. MURUGAM, N. SPIERINGS, D. R. EVANS, M. SONG & C. A. CLARK (2008), «Atlas-based segmentation of white matter tracts of the human brain using diffusion tensor tractography and comparison with classical dissection», *NeuroImage*, **39**, n° 1, pages 62–79.
- [LAZAR & ALEXANDER 2005] M. LAZAR & A. L. ALEXANDER (2005), «Bootstrap white matter tractography (BOOT-TRAC)», *NeuroImage*, **24**, n° 2, pages 524–532.
- [LAZAR et coll. 2003] M. LAZAR, D. M. WEINSTEIN, J. S. TSURUDA, K. M. HASAN, K. ARFANAKIS, M. E. MEYERAND, B. BADIE, H. A. ROWLEY, V. HAUGHTON, A. FIELD & A. L. ALEXANDER (2003), «White matter tractography using diffusion tensor deflection», *Human Brain Mapping*, **18**, n° 4, pages 306–321.

- [LE BIHAN & BRETON 1985] D. LE BIHAN & E. BRETON (1985), «Imagerie de diffusion in vivo par résonance magnétique nucléaire», *Comptes-Rendus de l'Académie des Sciences*, **93**, n° 5, pages 27-34.
- [LEE & VERLEYSSEN 2007] J. A. LEE & M. VERLEYSSEN (2007), *Nonlinear dimensionality reduction*, Springer.
- [LEEMANS et coll. 2005] A. LEEMANS, J. SIJBERS, S. D. BACKER, E. VANDERVLIELT & P. PARIZEL (2005), «Affine coregistration of diffusion tensor magnetic resonance images using mutual information», *IEEE Advanced Concepts for Intelligent Vision Systems (ACIVS)*, **3708**, pages 523–530.
- [LEEMANS et coll. 2006] A. LEEMANS, J. SIJBERS, S. DE BACKER, E. VANDERVLIELT & P. PARIZEL (2006), «Multiscale white matter fiber tract coregistration : a new feature-based approach to align diffusion tensor data», *Magnetic Resonance in Medicine*, **55**, n° 6, pages 1414–1423.
- [LENGLET 2006] C. LENGLET (2006), *Méthodes géométriques et variationnelles pour le traitement d'IRM du tenseur de diffusion*, Thèse de doctorat, Université de Nice Sophia-Antipolis.
- [LENGLET et coll. 2006] C. LENGLET, M. ROUSSON & R. DERICHE (2006), «DTI segmentation by statistical surface evolution», *IEEE Transactions on Medical Imaging*, **25**, n° 6, pages 685–700.
- [LENGLET et coll. 2009] C. LENGLET, J. S. W. CAMPBELL, M. DESCOTEAUX, G. HARO, P. SAVADJIEV, D. WASSERMANN, A. ANWANDER, R. DERICHE, G. B. PIKE, G. SAPIRO, K. SIDDIQI & P. M. THOMPSON (2009), «Mathematical methods for diffusion MRI processing», *NeuroImage*, **45**, n° 1 Suppl, pages S111–122.
- [LIU et coll. 2008] Y. LIU, M. LIANG, Y. ZHOU, Y. HE, Y. HAO, M. SONG, C. YU, H. LIU, Z. LIU & T. JIANG (2008), «Disrupted small-world networks in schizophrenia», *Brain*, **131**, n° 4, pages 945–961.
- [LU et coll. 2006] H. LU, J. H. JENSEN, A. RAMANI & J. A. HELPERN (2006), «Three dimensional characterization of non Gaussian water diffusion in humans using diffusion kurtosis imaging», *NMR in Biomedicine*, **19**, n° 2, pages 236–247.
- [LYNALL et coll. 2010] M. LYNALL, D. S. BASSETT, R. KERWIN, P. J. MCKENNA, M. KITZBICHLER, U. MULLER & E. BULLMORE (2010), «Functional connectivity and brain networks in schizophrenia», *Journal of Neuroscience*, **30**, n° 28, pages 9477–9487.
- [MACHADO 2007] A. MACHADO (2007), «Multiple testing correction in medical image analysis», *Journal of Mathematical Imaging and Vision*, **29**, n° 2-3, pages 107–117.
- [MADDAH et coll. 2008] M. MADDAH, W. E. L. GRIMSON, S. K. WARFIELD & W. M. WELLS (2008), «A unified framework for clustering and quantitative analysis of white matter fiber tracts», *Medical Image Analysis*, **12**, n° 2, pages 191–202.
- [MARKOU & SINGH 2003] M. MARKOU & S. SINGH (2003), «Novelty detection : a review—part 1 : statistical approaches», *Signal Processing*, **83**, n° 12, pages 2481–2497.
- [MARKOU & SINGH 2003] M. MARKOU & S. SINGH (2003), «Novelty detection : a review—part 2 : neural network based approaches», *Signal Processing*, **83**, n° 12, pages 2499–2521.
- [MAYER et coll. 2011] A. MAYER, G. ZIMMERMAN-MORENO, R. SHADMI, A. BATIKOFF & H. GREENSPAN (2011), «A supervised framework for the registration and segmentation of white matter fiber tracts», *IEEE Transactions on Medical Imaging*, **30**, n° 1, pages 131–145.
- [MAZZIOTTA et coll. 1995] J. C. MAZZIOTTA, A. W. TOGA, A. EVANS, P. FOX & J. LANCASTER (1995), «A probabilistic atlas of the human brain : theory and rationale for its

- development. The international consortium for brain mapping (ICBM)», *NeuroImage*, **2**, n° 2, pages 89–101.
- [MIKA et coll. 1999] S. MIKA, B. SCHÖLKOPF, A. SMOLA, K. MÜLLER, M. SCHOLZ & G. RÄTSCH (1999), «Kernel PCA and de-noising in feature spaces», *Proceedings of the 1998 conference on Advances in neural information processing systems*, **11**, pages 536–542.
- [MINATI & WĘGLARZ 2007] L. MINATI & W. P. WĘGLARZ (2007), «Physical foundations, models, and methods of diffusion magnetic resonance imaging of the brain : a review», *Concepts in Magnetic Resonance Part A*, **30A**, n° 5, pages 278–307.
- [MOAKHER 2008] M. MOAKHER (2008), «Fourth-order Cartesian tensors : old and new facts, notions and applications», *Quarterly Journal of Mechanics and Applied Mathematics*, **61**, n° 2, pages 181–203.
- [MOAKHER 2009] M. MOAKHER (2009), «The algebra of fourth-order tensors with application to diffusion MRI», dans D. LAIDLAW & J. WEICKERT, éditeurs, *Visualization and Processing of Tensor Fields*, Mathematics and Visualization, pages 57–80, Springer.
- [MORI et coll. 1999] S. MORI, B. J. CRAIN, V. P. CHACKO & P. C. M. V. ZIJL (1999), «Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging», *Annals of Neurology*, **45**, n° 2, pages 265–269.
- [MORI et coll. 2008] S. MORI, K. OISHI, H. JIANG, L. JIANG, X. LI, K. AKHTER, K. HUA, A. V. FARIA, A. MAHMOOD, R. WOODS, A. W. TOGA, G. B. PIKE, P. R. NETO, A. EVANS, J. ZHANG, H. HUANG, M. I. MILLER, P. VAN ZIJL & J. MAZZIOTTA (2008), «Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template», *NeuroImage*, **40**, n° 2, pages 570 - 582.
- [MOUNTCASTLE 1998] V. B. MOUNTCASTLE (1998), *Perceptual neuroscience : the cerebral cortex*, Harvard University Press.
- [MÚÑOZ-MORENO & MARTIN-FERNANDEZ 2009] E. MÚÑOZ-MORENO & M. MARTIN-FERNANDEZ (2009), «Characterization of the similarity between diffusion tensors for image registration», *Computers in Biology and Medicine*, **39**, pages 251–265.
- [MÚÑOZ-MORENO et coll. 2009] E. MÚÑOZ-MORENO, R. CÁRDENES-ALMEIDA & M. MARTIN-FERNANDEZ (2009), «Review of techniques for registration of diffusion tensor imaging», dans S. AJA-FERNÁNDEZ, R. L. GARCÍA, D. TAO & X. LI, éditeurs, *Tensors in Image Processing and Computer Vision*, Advances in Pattern Recognition, page 273–297.
- [NIENDORF et coll. 1996] T. NIENDORF, R. M. DIJKHUIZEN, D. G. NORRIS, M. VAN LOOKEREN CAMPAGNE & K. NICOLAY (1996), «Biexponential diffusion attenuation in various states of brain tissue : implications for diffusion weighted imaging», *Magnetic Resonance in Medicine*, **36**, n° 6, pages 847–857.
- [NUCIFORA et coll. 2007] P. G. P. NUCIFORA, R. VERMA, S.-K. LEE & E. R. MELHEM (2007), «Diffusion-tensor MR imaging and tractography : exploring brain microstructure and connectivity», *Radiology*, **245**, n° 2, pages 367–384.
- [O'DONNELL & WESTIN 2007] L. J. O'DONNELL & C. F. WESTIN (2007), «Automatic tractography segmentation using a high-dimensional white matter atlas», *IEEE Transactions on Medical Imaging*, **26**, n° 11, pages 1562–1575.
- [O'DONNELL et coll. 2009] L. J. O'DONNELL, C. WESTIN & A. J. GOLBY (2009), «Tract-based morphometry for white matter group analysis», *NeuroImage*, **45**, n° 3, pages 832–844.
- [OH et coll. 2007] J. S. OH, I. C. SONG, J. S. LEE, H. KANG, K. S. PARK, E. KANG & D. S. LEE (2007), «Tractography-guided statistics (TGIS) in diffusion tensor imaging for the detection of gender difference of fiber integrity in the midsagittal and parasagittal corpora callosa», *NeuroImage*, **36**, n° 3, pages 606–616.

- [OTSU 1979] N. OTSU (1979), «A threshold selection method from graylevel histograms», *IEEE Transactions on Systems, Man and Cybernetics*, **9**, n° 1, pages 62–66.
- [ÖZARSLAN & MARECI 2003] E. ÖZARSLAN & T. H. MARECI (2003), «Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging», *Magnetic Resonance in Medicine*, pages 955—965.
- [ÖZARSLAN et coll. 2005] E. ÖZARSLAN, B. C. VEMURI & T. H. MARECI (2005), «Generalized scalar measures for diffusion MRI using trace, variance, and entropy», *Magnetic Resonance in Medicine*, **53**, n° 4, pages 866–876.
- [ÖZARSLAN et coll. 2006] E. ÖZARSLAN, T. M. SHEPHERD, B. C. VEMURI, S. J. BLACKBAND & T. H. MARECI (2006), «Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT)», *NeuroImage*, **31**, n° 3, pages 1086–1103.
- [ÖZARSLAN et coll. 2006] E. ÖZARSLAN, T. M. SHEPHERD, B. C. VEMURI, S. J. BLACKBAND & T. H. MARECI (2006), «Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT)», *NeuroImage*, **31**, n° 3, pages 1086–1103.
- [PAJEVIC & PIERPAOLI 2000] S. PAJEVIC & C. PIERPAOLI (2000), «Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data : application to white matter fiber tract mapping in the human brain», *Magnetic Resonance in Medicine*, **43**, n° 6, pages 921–921.
- [PANNEK et coll. 2011] K. PANNEK, J. L. MATHIAS, E. D. BIGLER, G. BROWN, J. D. TAYLOR & S. E. ROSE (2011), «The average pathlength map : a diffusion MRI tractography-derived index for studying brain pathology», *NeuroImage*, **55**, n° 1, pages 133–141.
- [PARK et coll. 2008] C. H. PARK, S. Y. KIM, Y. KIM & K. KIM (2008), «Comparison of the small-world topology between anatomical and functional connectivity in the human brain», *Physica A : Statistical Mechanics and its Applications*, **387**, n° 23, pages 5958–5962.
- [PARK et coll. 2003] H. PARK, M. KUBICKI, M. E. SHENTON, A. GUIMOND, R. W. MCCARLEY, S. E. MAIER, R. KIKINIS, F. A. JOLESZ & C. WESTIN (2003), «Spatial normalization of diffusion tensor MRI using multiple channels», *NeuroImage*, **20**, n° 4, pages 1995–2009.
- [PARK et coll. 2007] J. PARK, D. KANG, J. KIM, J. T. KWOK & I. W. TSANG (2007), «SVDD-based pattern denoising», *Neural Computation*, **19**, n° 7, pages 1919–1938.
- [PARKER & ALEXANDER 2003] G. J. M. PARKER & D. C. ALEXANDER (2003), «Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information», *Proceedings of the IPMI*, **18**, pages 684–695.
- [PASTERNAK et coll. 2010] O. PASTERNAK, N. SOCHEN & P. J. BASSER (2010), «The effect of metric selection on the analysis of diffusion tensor MRI data», *NeuroImage*, **49**, n° 3, pages 2190–2204.
- [PÉCHAUD et coll. 2009] M. PÉCHAUD, M. DESCOTEAUX & R. KERIVEN (2009), «Brain connectivity using geodesics in HARDI.», *Proceedings of MICCAI*, **5762**, n° 2, pages 482–489.
- [PERRIN 2006] M. PERRIN (2006), *Imagerie de diffusion à haute résolution angulaire : étude du modèle q-ball par couplage simulations - fantômes et applications au suivi de fibres et à la parcellisation du cortex.*, Thèse de doctorat, Université Paris-Sud XI, Faculté des Sciences d’Orsay.
- [PERRIN et coll. 2005] M. PERRIN, C. POUPON, B. RIEUL, P. LEROUX, A. CONSTANTINESCO, J.-F. MANGIN & D. LEBIHAN (2005), «Validation of q-ball imaging with a diffusion fibre-crossing phantom on a clinical scanner», *Philosophical Transactions of the Royal Society B : Biological Sciences*, **360**, n° 1457, pages 881–891.

- [POLLARI et coll. 2006] M. POLLARI, T. NEUVONEN & J. LÖTJÖNEN (2006), «Affine registration of diffusion tensor MR images», *Proceedings of MICCAI*, **4191**, n° 2, pages 629–636.
- [POLLARI et coll. 2007] M. POLLARI, T. NEUVONEN, M. LILJA & J. LOTJONEN (2007), «Comparative evaluation of voxel similarity measures for affine registration of diffusion tensor MR images», *Proceedings of the IEEE ISBI*, pages 768–771.
- [POUPON 1999] C. POUPON (1999), *Détection des faisceaux de fibres de la substance blanche pour l'étude de la connectivité anatomique cérébrale.*, Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications.
- [POUPON et coll. 2000] C. POUPON, C. A. CLARK, V. FROUIN, J. RÉGIS, I. BLOCH, D. L. BIHAN & J. F. MANGIN (2000), «Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles», *NeuroImage*, **12**, n° 2, pages 184–195.
- [POUPON et coll. 2008] C. POUPON, B. RIEUL, I. KEZELE, M. PERRIN, F. POUPON & J.-F. MANGIN (2008), «New diffusion phantoms dedicated to the study and validation of high angular resolution diffusion imaging (HARDI) models», *Magnetic Resonance in Medicine*, **60**, n° 6, pages 1276–1283.
- [POUPON et coll. 2010] C. POUPON, L. LARIBIERE, G. TOURNIER, J. BERNARD, D. FOURNIER, P. FILLARD, M. DESCOTEAUX & J.-F. MANGIN (2010), «A diffusion hardware phantom looking like a coronal brain slice», *Proceedings of the International Society for Magnetic Resonance in Medicine*.
- [QI 2006] L. QI (2006), «Rank and eigenvalues of a supersymmetric tensor, the multivariate homogeneous polynomial and the algebraic hypersurface it defines», *Journal of Symbolic Computation*, **41**, n° 12, pages 1309–1327.
- [QI et coll. 2008] L. QI, Y. WANG & E. X. WU (2008), «D-eigenvalues of diffusion kurtosis tensors», *Journal of Computational and Applied Mathematics*, **221**, n° 1, pages 150 - 157.
- [QI et coll. 2009] L. QI, D. HAN & E. X. WU (2009), «Principal invariants and inherent parameters of diffusion kurtosis tensors», *Journal of Mathematical Analysis and Applications*, **349**, n° 1, pages 165–180.
- [RAFFELT et coll. 2011] D. RAFFELT, J. TOURNIER, J. FRIPP, S. CROZIER, A. CONNELLY & O. SALVADO (2011), «Symmetric diffeomorphic registration of fibre orientation distributions», *NeuroImage*, **56**, n° 3, pages 1171–1180.
- [RAMIREZ-MANZANARES & RIVERA 2006] A. RAMIREZ-MANZANARES & M. RIVERA (2006), «Basis tensor decomposition for restoring intra-voxel structure and stochastic walks for inferring brain connectivity in DT-MRI», *International Journal of Computer Vision*, **69**, pages 77–92.
- [REISERT et coll. 2011] M. REISERT, I. MADER, C. ANASTASOPOULOS, M. WEIGEL, S. SCHNELL & V. KISELEV (2011), «Global fiber reconstruction becomes practical», *NeuroImage*, **54**, pages 955–962.
- [RENARD et coll. 2010] F. RENARD, V. NOBLET, A. GRIGIS, C. HEINRICH & S. KREMER (2010), «Comparison of interpolation methods for angular resampling of diffusion weighted images», dans *Proceedings of International Conference on Image Processing Theory, Tools and Applications*, pages 207–211, juillet 2010.
- [RENARD et coll. 2010] F. RENARD, V. NOBLET, C. HEINRICH & S. KREMER (2010), «Re-orientation strategies for high order tensors», dans *Proceedings of the IEEE ISBI*, pages 1185–1188, avril 2010.
- [RISKEN & FRANK 1996] H. RISKEN & T. FRANK (1996), *The Fokker-Planck equation : methods of solutions and applications*, Springer.

- [ROBINSON et coll. 2010] E. C. ROBINSON, A. HAMMERS, A. ERICSSON, A. D. EDWARDS & D. RUECKERT (2010), «Identifying population differences in whole-brain structural networks : a machine learning approach», *NeuroImage*, **50**, n° 3, pages 910–919.
- [ROCA et coll. 2009] P. ROCA, D. RIVIÈRE, P. GUEVARA, C. POUPON & J.-F. MANGIN (2009), «Tractography-based parcellation of the cortex using a spatially-informed dimension reduction of the connectivity matrix», *Proceedings of MICCAI*, **5761**, n° 1, pages 935–942.
- [ROHDE et coll. 2003] G. K. ROHDE, S. PAJEVIC, C. PIERPAOLI & P. J. BASSER (2003), «A Comprehensive Approach for Multi-channel Image Registration», *Workshop on Biomedical Image Registration*, pages 214–223.
- [ROHDE et coll. 2004] G. K. ROHDE, S. PAJEVIC & C. PIERPAOLI (2004), «Multi-channel registration of diffusion tensor images using directional information», *Proceedings of the IEEE ISBI*, pages 712– 715.
- [ROMANIUK et coll. 2003] B. ROMANIUK, M. DESVIGNES & M. REVENU (2003), «Shape variability and spatial relationships modeling in statistical pattern recognition», *Pattern Recognition Letters*, **25**, pages 239-247.
- [ROYSAM et coll. 2009] B. ROYSAM, W. SHAIN & G. A. ASCOLI (2009), «The central role of neuroinformatics in the national academy of engineering’s grandest challenge : reverse engineer the brain», *Neuroinformatics*, **7**, n° 1, pages 1–5.
- [RUBINOV & SPORNS 2010] M. RUBINOV & O. SPORNS (2010), «Complex network measures of brain connectivity : uses and interpretations», *NeuroImage*, **52**, n° 3, pages 1059–1069.
- [RUBINOV et coll. 2009] M. RUBINOV, S. A. KNOCK, C. J. STAM, S. MICHELOYANNIS, A. W. F. HARRIS, L. M. WILLIAMS & M. BREAKSPEAR (2009), «Small-world properties of nonlinear brain activity in schizophrenia», *Human Brain Mapping*, **30**, n° 2, pages 403–416.
- [RUIZ-ALZOLA et coll. 2002] J. RUIZ-ALZOLA, C. WESTIN, S. K. WARFIELD, C. ALBEROLA, S. E. MAIER & R. KIKINIS (2002), «Nonrigid registration of 3D tensor medical data», *Medical Image Analysis*, **6**, n° 2, pages 143–161.
- [SABUNCU et coll. 2009] M. R. SABUNCU, S. K. BALCI, M. E. SHENTON & P. GOLLAND (2009), «Image-driven population analysis through mixture modeling», *IEEE Transactions on Medical Imaging*, **28**, n° 9, pages 1473–1487.
- [SCANNELL et coll. 1999] J. SCANNELL, G. BURNS, C. HILGETAG, M. O’NEIL & M. YOUNG (1999), «The connectional organization of the cortico-thalamic system of the cat», *Cerebral Cortex*, **9**, n° 3, pages 277–299.
- [SCHERRER & WARFIELD 2010] B. SCHERRER & S. K. WARFIELD (2010), «Why multiple b-values are required for multi-tensor models : evaluation with a constrained log-Euclidean model», *Proceedings of the IEEE ISBI*, pages 1389–1392.
- [SCHÖLKOPF et coll. 2001] B. SCHÖLKOPF, J. C. PLATT, J. SHAWE-TAYLOR, A. J. SMOLA & R. C. WILLIAMSON (2001), «Estimating the support of a high-dimensional distribution», *Neural Computation*, **13**, n° 7, pages 1443–1471.
- [SCHWARTZMAN 2010] A. SCHWARTZMAN (2010), «Group comparison of eigenvalues and eigenvectors of diffusion tensors», *Journal of the American Statistical Association*, **105**, n° 490, pages 588–599.
- [SCHWARTZMAN et coll. 2005] A. SCHWARTZMAN, R. F. DOUGHERTY & J. E. TAYLOR (2005), «Cross-subject comparison of principal diffusion direction maps», *Magnetic Resonance in Medicine*, **53**, n° 6, pages 1423–1431.
- [SHAFFER 1995] J. P. SHAFFER (1995), «Multiple hypothesis testing», *Annual Review of Psychology*, **46**, n° 1, pages 561–584.

- [SHAWE-TAYLOR & CRISTIANINI 2004] J. SHAWE-TAYLOR & N. CRISTIANINI (2004), *Kernel methods for pattern analysis*, Cambridge University Press.
- [SJÖSTRAND et coll. 2007] K. SJÖSTRAND, M. S. HANSEN, H. B. LARSSON & R. LARSEN (2007), «A path algorithm for the support vector domain description and its application to medical imaging», *Medical Image Analysis*, **11**, n° 5, pages 417–428.
- [SKUDLARSKI et coll. 2008] P. SKUDLARSKI, K. JAGANNATHAN, V. D. CALHOUN, M. HAMPSON, B. A. SKUDLARSKA & G. PEARLSON (2008), «Measuring brain connectivity : diffusion tensor imaging validates resting state temporal correlations», *NeuroImage*, **43**, n° 3, pages 554–561.
- [SMITH et coll. 2006] S. M. SMITH, M. JENKINSON, H. JOHANSEN-BERG, D. RUECKERT, T. E. NICHOLS, C. E. MACKAY, K. E. WATKINS, O. CICCARELLI, M. Z. CADER, P. M. MATTHEWS & T. E. BEHRENS (2006), «Tract-based spatial statistics : voxelwise analysis of multi-subject diffusion data», *NeuroImage*, **31**, n° 4, pages 1487–1505.
- [SPORNS 2007] O. SPORNS (2007), «Brain connectivity», *Scholarpedia*, **2**, n° 10, pages 4695.
- [SPORNS 2010] O. SPORNS (2010), «Connectome», *Scholarpedia*, **5**, n° 2, pages 5584.
- [SPORNS 2010] O. SPORNS (2010), *Networks of the brain*, MIT Press.
- [SPORNS et coll. 2005] O. SPORNS, G. TONONI & R. KÖTTER (2005), «The human connectome : a structural description of the human brain», *PLoS Computational Biology*, **1**, n° 4, pages e42.
- [STAM et coll. 2007] C. STAM, B. JONES, G. NOLTE, M. BREAKSPEAR & P. SCHELTENS (2007), «Small-world networks and functional connectivity in Alzheimer’s disease», *Cerebral Cortex*, **17**, n° 1, pages 92–99.
- [STEJSKAL & TANNER 1965] E. O. STEJSKAL & J. E. TANNER (1965), «Spin diffusion measurements : spin echoes in the presence of a time-dependent field gradient», *Journal of Chemical Physics*, **42**, n° 1, pages 288.
- [STROGATZ 2001] S. H. STROGATZ (2001), «Exploring complex networks», *Nature*, **410**, n° 6825, pages 268–276.
- [STUDHOLME 2007] C. STUDHOLME (2007), «Incorporating DTI data as a constraint in deformation tensor morphometry between T1 MR images», *Proceedings of the IPMI*, **20**, pages 223–232.
- [SUN et coll. 2009] Z. Y. SUN, M. PERROT, A. TUCHOLKA, D. RIVIÈRE & J.-F. MANGIN (2009), «Constructing a dictionary of human brain folding patterns», *Proceedings of MICCAI*, **5762**, n° 2, pages 117–124.
- [SUPEKAR et coll. 2008] K. SUPEKAR, V. MENON, D. RUBIN, M. MUSEN & M. D. GREICIUS (2008), «Network analysis of intrinsic functional brain connectivity in Alzheimer’s disease», *PLoS Computational Biology*, **4**, n° 6, pages e1000100.
- [TALAIRACH & TOURNOUX 1988] J. TALAIRACH & P. TOURNOUX (1988), *Co-planar stereotaxic atlas of the human brain : 3-D proportional system : an approach to cerebral imaging*, Thieme.
- [TAO & MILLER 2006] X. TAO & J. MILLER (2006), «A method for registering diffusion weighted magnetic resonance images», *Proceedings of MICCAI*, **4191**, pages 594–602.
- [TARANTOLA 2005] A. TARANTOLA (2005), *Elements for physics : quantities, qualities, and intrinsic theories*, Springer.
- [TAX & DUIN 1999] D. M. TAX & R. P. DUIN (1999), «Support vector domain description», *Pattern Recognition Letters*, **20**, n° 11-13, pages 1191–1199.
- [TAX & DUIN 2004] D. M. J. TAX & R. P. W. DUIN (2004), «Support vector data description», *Machine Learning*, **54**, n° 1, pages 45–66.

- [TENENBAUM et coll. 2000] J. B. TENENBAUM, V. DE SILVA & J. C. LANGFORD (2000), «A global geometric framework for nonlinear dimensionality reduction», *Science*, **290**, n° 5500, pages 2319–2323.
- [THOMAZ et coll. 2007] C. E. THOMAZ, F. L. DURAN, G. F. BUSATTO, D. F. GILLIES & D. RUECKERT (2007), «Multivariate statistical differences of MRI samples of the human brain», *Journal of Mathematical Imaging and Vision*, **29**, n° 2-3, pages 95–106.
- [TOGA & MAZZIOTTA 2002] A. W. TOGA & J. C. MAZZIOTTA (2002), *Brain mapping : The methods*, Academic Press, deuxième édition.
- [TOURNIER et coll. 2004] J. TOURNIER, F. CALAMANTE, D. G. GADIAN & A. CONNELLY (2004), «Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution», *NeuroImage*, **23**, n° 3, pages 1176–1185.
- [TUCH 2002] D. TUCH (2002), *Diffusion MRI of complex tissue structure*, Thèse de doctorat, Massachusetts Institute of Technology.
- [TUCH 2004] D. S. TUCH (2004), «Q-ball imaging», *Magnetic Resonance in Medicine*, **52**, n° 6, pages 1358–1372.
- [TUCH et coll. 2002] D. S. TUCH, T. G. REESE, M. R. WIEGELL, N. MAKRIS, J. W. BELLIVEAU & V. J. WEDEEN (2002), «High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity», *Magnetic Resonance in Medicine*, **48**, n° 4, pages 577–582.
- [TZOURIO-MAZOYER et coll. 2002] N. TZOURIO-MAZOYER, B. LANDEAU, D. PAPATHANASSIOU, F. CRIVELLO, O. ETARD, N. DELCROIX, B. MAZOYER & M. JOLIOT (2002), «Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain», *NeuroImage*, **15**, n° 1, pages 273–289.
- [VAN HECKE et coll. 2008] W. VAN HECKE, J. SIJBERS, E. D’AGOSTINO, F. MAES, S. DE BACKER, E. VANDERVLIEET, P. M. PARIZEL & A. LEEMANS (2008), «On the construction of an inter-subject diffusion tensor magnetic resonance atlas of the healthy human brain», *NeuroImage*, **43**, n° 1, pages 69–80.
- [VAPNIK 1999] V. VAPNIK (1999), *The nature of statistical learning theory*, Springer.
- [VAROQUAUX et coll. 2010] G. VAROQUAUX, F. BARONNET, A. KLEINSCHMIDT, P. FILLARD & B. THIRION (2010), «Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling», *Proceedings of MICCAI*, **6361**, pages 200–208.
- [VENKATARAMAN et coll. 2010] A. VENKATARAMAN, Y. RATHI, M. KUBICKI, C. WESTIN & P. GOLLAND (2010), «Joint generative model for fMRI/DWI and its application to population studies», *Proceedings of MICCAI*, **6361**, n° 1, pages 191–9.
- [VERAART et coll. 2010] J. VERAART, W. V. HECKE, I. BLOCKX, A. V. D. LINDEN, M. VERHOYE & J. SIJBERS (2010), «Non-rigid coregistration of diffusion kurtosis data», *Proceedings of the IEEE ISBI*, pages 392–395.
- [VERMA et coll. 2007] R. VERMA, P. KHURD & C. DAVATZIKOS (2007), «On analyzing diffusion tensor images by identifying manifold structure using isomaps», *IEEE Transactions on Medical Imaging*, **26**, n° 6, pages 772–778.
- [VIK 2004] T. VIK (2004), *Modèles statistiques d'apparence non gaussiens. Application à la création d'un atlas probabiliste de perfusion cérébrale en imagerie médicale.*, Thèse de doctorat, Université Louis Pasteur, Strasbourg 1.
- [VIK et coll. 2005] T. VIK, F. HEITZ, I. NAMER & J. ARMSPACH (2005), «On the modeling, construction, and evaluation of a probabilistic atlas of brain perfusion», *NeuroImage*, **24**, n° 4, pages 1088–1098.

- [VIK et coll. 2007] T. VIK, F. HEITZ & P. CHARBONNIER (2007), «Robust pose estimation and recognition using non-Gaussian modeling of appearance subspaces», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**, n° 5, pages 901–905.
- [WAKANA et coll. 2007] S. WAKANA, A. CAPRIHAN, M. M. PANZENBOECK, J. H. FALLON, M. PERRY, R. L. GOLLUB, K. HUA, J. ZHANG, H. JIANG, P. DUBEY, A. BLITZ, P. VAN ZIJL & S. MORI (2007), «Reproducibility of quantitative tractography methods applied to cerebral white matter», *NeuroImage*, **36**, n° 3, pages 630–644.
- [WANG et coll. 2009] J. WANG, L. WANG, Y. ZANG, H. YANG, H. TANG, Q. GONG, Z. CHEN, C. ZHU & Y. HE (2009), «Parcellation-dependent small-world brain functional networks : a resting-state fMRI study», *Human Brain Mapping*, **30**, n° 5, pages 1511–1523.
- [WANG et coll. 2011] Y. WANG, A. GUPTA, Z. LIU, H. ZHANG, M. L. ESCOLAR, J. H. GILMORE, S. GOUTTARD, P. FILLARD, E. MALTBIÉ, G. GERIG & M. STYNER (2011), «DTI registration in atlas based fiber analysis of infantile Krabbe disease», *NeuroImage*, **55**, n° 4, pages 1577–1586.
- [WASSERMANN et coll. 2008] D. WASSERMANN, M. DESCOTEAUX & R. DERICHE (2008), «Diffusion maps clustering for magnetic resonance q-ball imaging segmentation», *Journal of Biomedical Imaging*, **2008**, pages 5 :1–5 :12.
- [WASSERMANN et coll. 2010] D. WASSERMANN, L. BLOY, E. KANTERAKIS, R. VERMA & R. DERICHE (2010), «Unsupervised white matter fiber clustering and tract probability map generation : applications of a Gaussian process framework for white matter fibers», *NeuroImage*, **51**, n° 1, pages 228–241.
- [WATTS & STROGATZ 1998] D. J. WATTS & S. H. STROGATZ (1998), «Collective dynamics of 'small-world' networks», *Nature*, **393**, n° 6684, pages 440–442.
- [WEDEEN et coll. 2005] V. J. WEDEEN, P. HAGMANN, W. I. TSENG, T. G. REESE & R. M. WEISSKOFF (2005), «Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging», *Magnetic Resonance in Medicine*, **54**, n° 6, pages 1377–1386.
- [WEE et coll. 2010] C. WEE, P. YAP, J. N. BROWNDYKE, G. G. POTTER, D. C. STEFFENS, K. WELSH-BOHMER, L. WANG & D. SHEN (2010), «Accurate identification of MCI patients via enriched white-matter connectivity network», *Proceedings of the MLMI*, pages 140–147.
- [WEE et coll. 2011] C. WEE, P. YAP, W. LI, K. DENNY, J. N. BROWNDYKE, G. G. POTTER, K. A. WELSH-BOHMER, L. WANG & D. SHEN (2011), «Enriched white matter connectivity networks for accurate identification of MCI patients», *NeuroImage*, **54**, n° 3, pages 1812–1822.
- [WESTIN et coll. 2002] C. F. WESTIN, S. E. MAIER, H. MAMATA, A. NABAVI, F. A. JOLESZ & R. KIKINIS (2002), «Processing and visualization for diffusion tensor MRI», *Medical Image Analysis*, **6**, n° 2, pages 93 – 108.
- [WHITCHER et coll. 2007] B. WHITCHER, J. J. WISCO, N. HADJIKHANI & D. S. TUCH (2007), «Statistical group comparison of diffusion tensors via multivariate hypothesis testing», *Magnetic Resonance in Medicine*, **57**, n° 6, pages 1065–1074.
- [WHITE et coll. 1986] J. G. WHITE, E. SOUTHGATE, J. N. THOMSON & S. BRENNER (1986), «The structure of the nervous system of the nematode *Caenorhabditis elegans*», *Philosophical Transactions of the Royal Society B, Biological Sciences*, **314**, n° 1165, pages 1–340.
- [WIEST-DAESSLÉ et coll. 2008] N. WIEST-DAESSLÉ, S. PRIMA, P. COUPÉ, S. P. MORRISSEY & C. BARILLOT (2008), «Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI : applications to DT-MRI», *Proceedings of MICCAI*, **5242**, n° 2, pages 171–179.

- [WOLZ et coll. 2010] R. WOLZ, P. ALJABAR, J. V. HAJNAL, A. HAMMERS & D. RUECKERT (2010), «LEAP : learning embeddings for atlas propagation», *NeuroImage*, **49**, n° 2, pages 1316–1325.
- [XU et coll. 2003] D. XU, S. MORI, D. SHEN, P. C. M. VAN ZIJL & C. DAVATZIKOS (2003), «Spatial normalization of diffusion tensor fields», *Magnetic Resonance in Medicine*, **50**, n° 1, pages 175–182.
- [XU et coll. 2009] Q. XU, A. W. ANDERSON, J. C. GORE & Z. DING (2009), «Unified bundling and registration of brain white matter fibers», *IEEE Transactions on Medical Imaging*, **28**, n° 9, pages 1399–1411.
- [YAMADA et coll. 2009] K. YAMADA, K. SAKAI, K. AKAZAWA, S. YUEN & T. NISHIMURA (2009), «MR tractography : a review of its clinical applications», *Magnetic Resonance in Medical Sciences*, **8**, n° 4, pages 165–174.
- [YAP et coll. 2011] P. YAP, Y. CHEN, H. AN, Y. YANG, J. H. GILMORE, W. LIN & D. SHEN (2011), «SPHERE : SPHERICAL harmonic elastic REGISTRATION of HARDI data», *NeuroImage*, **55**, n° 2, pages 545–556.
- [YEO et coll. 2008] B. T. YEO, T. VERCAUTEREN, P. FILLARD, X. PENNEC, P. GOLLAND, N. AYACHE & O. CLATZ (2008), «DTI Registration with exact finite-strain differential», *Proceedings of the IEEE ISBI*, pages 700–703.
- [YO et coll. 2009] T. YO, A. ANWANDER, M. DESCOTEAUX, P. FILLARD, C. POUPON & T. R. KNÖSCHE (2009), «Quantifying brain connectivity : a comparative tractography study», *Proceedings of MICCAI*, **5761**, n° 1, pages 886–893.
- [YUSHKEVICH et coll. 2008] P. A. YUSHKEVICH, H. ZHANG, T. J. SIMON & J. C. GEE (2008), «Structure-specific statistical mapping of white matter tracts», *NeuroImage*, **41**, n° 2, pages 448–461.
- [ZALESKY & FORNITO 2009] A. ZALESKY & A. FORNITO (2009), «A DTI-derived measure of cortico-cortical connectivity», *IEEE Transactions on Medical Imaging*, **28**, n° 7, pages 1023–1036.
- [ZALESKY et coll. 2010] A. ZALESKY, A. FORNITO, I. H. HARDING, L. COCCHI, M. YÜCEL, C. PANTELIS & E. T. BULLMORE (2010), «Whole-brain anatomical networks : does the choice of nodes matter?», *NeuroImage*, **50**, n° 3, pages 970–983.
- [ZHANG et coll. 2006] H. ZHANG, P. A. YUSHKEVICH, D. C. ALEXANDER & J. C. GEE (2006), «Deformable registration of diffusion tensor MR images with explicit orientation optimization», *Medical Image Analysis*, **10**, n° 5, pages 764–785.
- [ZHANG et coll. 2007] H. ZHANG, B. B. AVANTS, P. A. YUSHKEVICH, J. H. WOO, S. WANG, L. F. MCCLUSKEY, L. B. ELMAN, E. R. MELHEM & J. C. GEE (2007), «High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences : an example study using amyotrophic lateral sclerosis», *IEEE Transactions on Medical Imaging*, **26**, n° 11, pages 1585–1597.
- [ZHANG et coll. 2009] H. ZHANG, P. A. YUSHKEVICH, D. RUECKERT & J. C. GEE (2009), «Tensor-based morphometry of fibrous structures with application to human brain white matter», *Proceedings of MICCAI*, **5762**, n° 2, pages 466–473.
- [ZHANG et coll. 2010] Y. ZHANG, A. DU, S. HAYASAKA, G.-H. JAHNG, J. HLAVIN, W. ZHAN, M. W. WEINER & N. SCHUFF (2010), «Patterns of age-related water diffusion changes in human brain by concordance and discordance analysis», *Neurobiology of Aging*, **31**, n° 11, pages 1991–2001.
- [ZHU et coll. 2010] H. ZHU, M. STYNER, Y. LI, L. KONG, Y. SHI, W. LIN, C. COE & J. H. GILMORE (2010), «Multivariate varying coefficient models for DTI tract statistics», *Proceedings of MICCAI*, **6361**, n° 1, pages 690–697.

- [ZVITIA et coll. 2008] O. ZVITIA, A. MAYER & H. GREENSPAN (2008), «Adaptive mean-shift registration of white matter tractographies», *Proceedings of the IEEE ISBI*, pages 692–695.
- [ZVITIA et coll. 2010] O. ZVITIA, A. MAYER, R. SHADMI, S. MIRON & H. K. GREENSPAN (2010), «Co-registration of white matter tractographies by adaptive-mean-shift and Gaussian mixture modeling», *IEEE Transactions on Medical Imaging*, **29**, n° 1, pages 132–145.