Diffusion directions imaging : reconstruction haute résolution des faisceaux de matière blanche par IRM de diffusion basse résolution angulaire

Aymeric Stamm

ANNÉE 2013

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Informatique

Ecole doctorale Matisse

présentée par

Aymeric Stamm

préparée à l'unité de recherche UMR CNRS 6074 / IRISA
Équipe d'accueil: VisAGeS - INSERM U746
Composante universitaire: Université de Rennes 1

Diffusion Directions Imaging: High resolution reconstruction of white matter fascicles from low angular resolution diffusion MRI

Thèse soutenue à Rennes
le 29 novembre 2013

devant le jury composé de :

Daniel ALEXANDER Professor of Imaging Science, University College of London / rapporteur

Rachid DERICHE Directeur de Recherche, INRIA Sophia-Antipolis / rapporteur

Xavier PENNEC Directeur de Recherche, INRIA Sophia-Antipolis / examinateur

Alexandra GOLBY Associate Professor, Harvard Medical School, Associate Neurosurgeon, Brigham and Women's Hospital / examinateur

Bernard DELYON Enseignant-Chercheur en Mathématiques, IRMAR / examinateur

Christian BARILLOT Directeur de Recherche, CNRS / directeur de thèse

Patrick PEREZ Directeur de Recherche, INRIA Bretagne-Atlantique / co-directeur de thèse

Acknowledgments

First, I would like to thank Pf. Daniel Alexander et Dr. Rachid Deriche for having reviewed this work and provided me with useful comments. I also thank Pr. Alexandra Golby, Dr. Xavier Pennec and Dr. Bernard Delyon for participating to the committee of my PhD defense.

I also address all my gratitude to my official advisors, Dr. Christian Barillot and Dr. Patrick Pérez, for putting their trust in me on this complicated topic like they did. They guided me all along this work providing me with a certain amount of independence while being there at crucial moments during these three years.

I express a deep and sincere thank you to Olivier Commowick, a colleague and very good friend, who has been a C++ mentor to me and, thanks to the numerous scientific discussions that we had, has largely contributed to this work. Now, I would like to spend a few words in french to thank my kin.

Je tiens tout d'abord à remercier ma femme Caroline qui a toujours eu foi en moi et m'a énormément soutenu durant la dure période de rédaction: un grand merci du fond du coeur, mon Amour! Je voudrais également à remercier mes parents et mes frères pour leurs encouragements et leur disponibilité pour discuter dans les moments plus sombres de l'existence d'un chercheur. Enfin, je voudrais remercier mes beaux-parents, Agnès et Laurent, ma belle-soeur Sophie et mon beau-frère Maël, pour leur convivialité et leur bonne humeur indispensables, notamment en période de rédaction, pour la santé du ciboulot!

Contents

Résumé en français

Sommaire

1 Contexte

L'IRM de diffusion permet d'imager *in* vivo et de manière non-invasive la microstructure des tissues en inférant les déplacements moléculaires induits par diffusion. La matière blanche est une des deux composantes du système nerveux central; elle est composée majoritairement de paquets d'axones de forme cylindrique, appelés *faisceaux*, qui assurent la transmission des messages nerveux et (ii) de cellules gliales de forme sphérique dont le rôle principal est de nourrir et protéger les axones. La diffusion dans de tels environnements est contrainte par leur géométrie et son étude est donc particulièrement pertinente dans les faisceaux où les directions de diffusion correspondent aux directions des faisceaux.

La tâche n'est cependant pas aisée. En effet, la résolution spatiale des images de diffusion est généralement basse, surtout en utilisation clinique, ce qui implique que de multiples faisceaux et cellules gliales coexistent en chaque voxel. Il est donc nécessaire de proposer une modélisation locale qui permette de dissocier la diffusion dans ces différents environnements. Mathématiquement, cela revient à dire que la distribution des déplacements moléculaires dans un voxel donné est un mélange de

distributions qui caractérisent les déplacements moléculaires dans chaque environnement (i.e., faiseaux et cellules gliales). Dans le contexte de l'imagerie de diffusion, ces modèles sont souvent dénommés modèles à multiples compartiments (MCM).

Plusieurs MCMs ont été proposés au cours de ces dix dernières années. Le premier d'entre eux est intitulé Diffusion Tensor Imaging (DTI). Ce modèle repose sur l'hypothèse qu'un voxel ne contient qu'un seul environnement microstructurel ce qui n'est pas le cas dans 80% de la matière blanche. En conséquence, les biomarqueurs dérivés du DTI ne sont pas spécifiques et les algorithmes de reconstruction de faisceaux dépendant du DTI ne sont pas précis. D'autres MCMs ont été proposés depuis pour pallier à ce défaut et ont donné lieu à des résultats prometteurs. Malheureusement, ces modèles nécessitent des acquisitions spécifiques et longues qui ne sont pas ou ne peuvent pas être pratique courante en utilisation clinique, à l'heure actuelle.

Dans ce contexte, cette thèse propose une chaine de traitement qui permet une reconstruction appropriée des faisceaux de matière blanche. Cette chaine inclut principalement deux éléments:

- un nouvel MCM qui peut être estimé à partir d'images de diffusion cliniques et qui fournit des biomarqueurs spécifiques et un algorithme de tractographie; ceci contraste les MCMs actuels dont l'estimation à partir de données cliniques est un problème mal posé;
- un algorithme de tractographie qui approxime une distribution de faisceaux à partir d'un point donné dans la matière blanche; ceci contraste la plupart des algorithmes traditionnels qui prennent en compte uniquement l'incertitude sur l'estimation des directions locales de diffusion.

2 Etat de l'art

L'imagerie médicale a fait un bon en avant lors de l'avènement de l'IRM [Lauterbur 1973, Damadian 1974, Mansfield 1977] dans les années soixante-dix. En effet, l'IRM est une technique d'imagerie non-invasive qui peut être employée *in* vivo. L'IRM de diffusion fut introduit dans les années quatre-vingt [Le Bihan 1985, Merboldt 1985, Taylor 1985] et permet la quantification de la diffusion des molécules d'eau dans un corps. Depuis les travaux de [Moseley 1990] qui ont prouvé l'anisotropie de la diffusion dans la matière blanche, l'IRM de diffusion suscite un grand intérêt car il semble être l'outil approprié pour étudier la microstructure et l'architecture de la matière blanche. En témoignent les multiples projets internationaux actuels relatifs à l'étude du cerveau par IRM: le projet "Brain Molecular Anatomy", le projet CONNECT, le projet "Connectome", le projet "One Mind For Research initiative", le projet "Human Brain", le projet "Projectome" ainsi que les deux projets "Human connectome" financés par le National Institute of Health (NIH) (NIH et NIH Blueprint. Pour mieux comprendre la complexité de l'étude du cerveau, la prochaine section est dédiée à une brève description du cerveau.

2.1 Le cerveau et la matière blanche

Le cerveau, aussi appelé *cerebrum*, est une partie du système nerveaux central principalement divisée en quatre régions appelées *lobes*:

- le lobe frontal responsable de la conscience;
- le lobe temporal responsable de la reconnaissance vocale;
- le lobe pariétal responsable du mouvement et de la perception des stimuli;
- le lobe occipital responsable de la vision.

Le cerveau comprend également le tronc cérébral responsable des fonctions vitales de base et le cerebellum responsable de la coordination de mouvement. Un schéma macroscopique du cerveau et proposé en fig. 1.

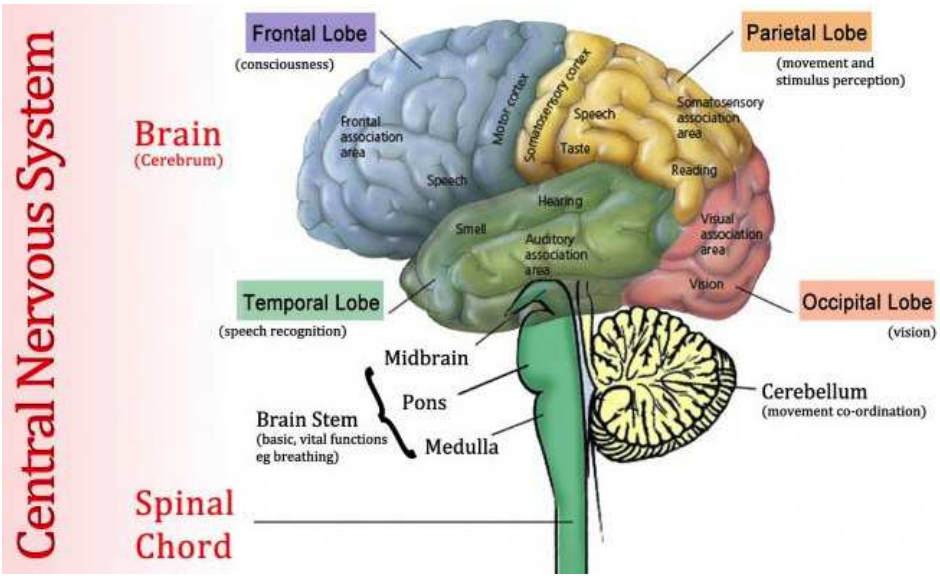


Figure 1: **Schematic view of the Human CNS.** Image courtesy of `http://climatereview.net/ChewTheFat/?attachment_id=1061`.

Ces différentes régions contiennent des millions de neurones qui communiquent entre eux par les faisceaux de matière blanche. Spécifiquement, un neurone est composé d'un corps cellulaire, de multiples petites branches appelées *dendrites* qui reçoivent les informations et d'un axone entouré d'une gaine de myéline qui transmet les informations (Figure 2). Ces axones sont regroupés par "paquets" appelés *faisceaux* et l'ensemble des faisceaux constitue une des deux parties de la matière blanche. Le second constituant de la matière blanche est la glie composée de cellules gliales. Il y a trois types de cellules gliales: les astrocytes qui protègent les axones, les oligodendrocytes qui produisent la myéline et les microglies qui réalisent la phagocytose des cellules mortes. Enfin,l'ensemble du cerveau baigne dans le liquide cérébro-spinal (CSF).

Les molécules d'eau dans la matière blanche sont donc réparties dans ces trois types d'environnement et donc sujettes à différents profils de diffusion selon la géometrie de ces environnements [Pfeuffer 1998]:

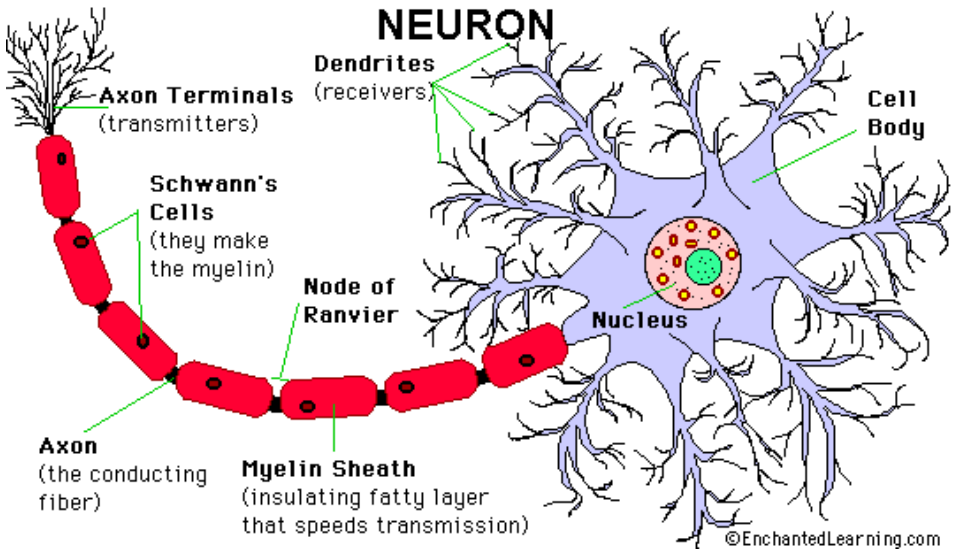


Figure 2: **Typical structure of a neuron.** (© EnchantedLearning.com)

• la diffusion dans les faisceaux est anisotrope et principalement contrainte le long des faisceaux;

• la diffusion dans les cellules gliales est isotrope et contrainte par la géométrie sphérique de ces cellules;

• la diffusion dans le CSF est isotrope et soumise à aucune contrainte.

Le réseau de transmission neuronal qui comprend l'ensemble des faisceaux a été beaucoup étudié [Catani 2002, Jellison 2004, Catani 2008, Thiebaut de Schotten 2011]. On s'accorde généralement pour classer les faisceaux en trois catégories:

• les faisceaux associatifs qui établissent des connections entre les régions corticales d'un même hemisphère cérébral;

• les faisceaux de projection qui lient les régions corticales aux noyaux profonds, tronc cérébral, cerebellum et moelle épinière;

• les faisceaux commissuraux qui interconnectent les régions corticales équivalentes dans les deux hémisphères.

La connaissance de cette architecture est résumée par la fig. 3.

2.2 L'IRM de diffusion

Dans cette section, nous décrivons succintement le principe de l'IRM de diffusion afin de comprendre quelles sont les données que nous manipulons et en quoi elles sont liées au phénomène de diffusion.

Le principe de l'IRM de diffusion est de mesurer l'atténuation de signal provoquée par l'application d'un gradient spatial de champ magnétique (DSG) par rapport au signal mesuré en l'absence d'un tel gradient. Un DSG est caractérisé par sa magnitude G, sa direction $\boldsymbol{g}$ et sa durée δ, ce qui conduisit [Callaghan 1991] à introduire la notation compacte $\boldsymbol{q} = \gamma\delta G\boldsymbol{g}$, où γ est la constante gyromagnétique du

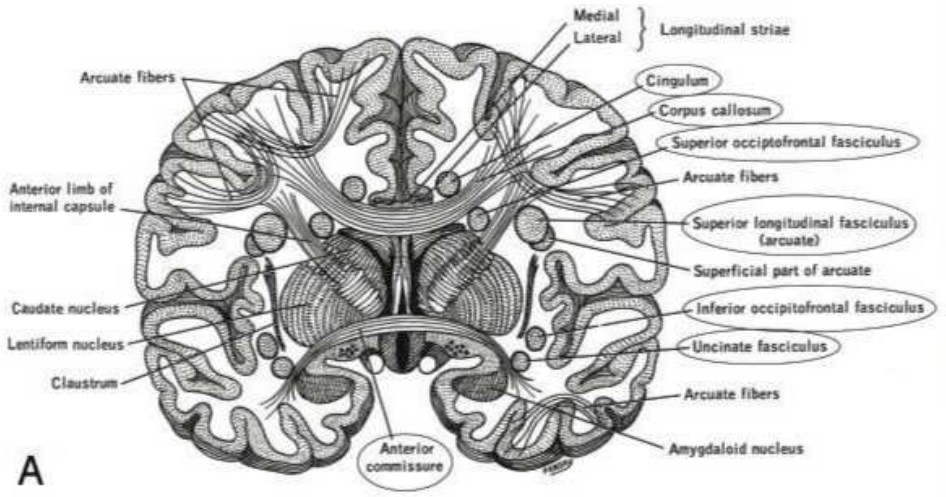


Figure 3: **Global view of the WM fascicles.** Illustrative coronal view of the relationship between several WM fascicles. Image courtesy of [Jellison 2004].

proton, pour désigner un DSG.

Concrètement, une IRM de diffusion est un examen durant lequel on acquiert une ou plusieurs images sans DSG et une collection d'images pour plusieurs DSGs séparés d'un temps Δ, appelé *temps de diffusion*. L'atténuation de signal provoquée par chaque DSG est lié à la distribution des déplacements moléculaires induits par diffusion. En effet, d'après [Einstein 1905], ces déplacements moléculaires sont aléatoires et on peut définir une fonction de densité de probabilité (PDF) $p_{\mathbf{x}_2|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_2;\Delta)$, qui donne la probabilité pour qu'un proton, initialement localisé en $\boldsymbol{x}_1$ se soit déplacé à la position $\boldsymbol{x}_2$ sous l'effet de la diffusion dans le temps Δ. L'atténuation de signal observée est alors reliée à cette PDF par la relation suivante:

$$A(\boldsymbol{q},\Delta) = \left| \int_V p_{\mathbf{x}_1}(\boldsymbol{x}_1) \left(\int_{\mathbb{R}^3} p_{\mathbf{x}_2|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_2;\Delta) \exp\{-i < \boldsymbol{q}, \boldsymbol{x}_2 - \boldsymbol{x}_1 >\} \mathrm{d}\boldsymbol{x}_2 \right) \mathrm{d}\boldsymbol{x}_1 \right| , \quad (1)$$

où $< \cdot, \cdot >$ est le produit scalaire dans $\mathbb{R}^3$, V est le volume du voxel et $p_{\mathbf{x}_1}(\boldsymbol{x}_1)$ est la PDF qui donne la probabilité pour qu'un proton soit localisé en position $\boldsymbol{x}_1$ dans le voxel.

L'eq. (1) peut être simplifiée par l'introduction d'un "ensemble-averaged propagator" (EAP) dans le voxel:

$$p_{\mathbf{x}}(\boldsymbol{x};\Delta) = \int_V p_{\mathbf{x}_1}(\boldsymbol{x}_1) p_{\mathbf{x}_1+\mathbf{x}|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_1+\boldsymbol{x};\Delta) \mathrm{d}\boldsymbol{x}_1 , \quad (2)$$

où $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$ est le vecteur aléatoire représentant le déplacement moléculaires induit par diffusion. L'atténuation de signal se simplifie alors de la manière suivante:

$$A(\boldsymbol{q},\Delta) = \left| \int_{\mathbb{R}^3} p_{\mathbf{x}}(\boldsymbol{x};\Delta) \exp\{-i < \boldsymbol{q}, \boldsymbol{x} >\} \mathrm{d}\boldsymbol{x} \right| . \quad (3)$$

Si aucun obstacle n'entrave le déplacement moléculaire, on dit que la diffusion est libre. Dans ce cas, l'EAP est une PDF Gaussienne de moyenne 0 et matrice de

covariance $2D(\Delta-\delta/3)I_3$, où D est le coefficient de diffusivité [Stejskal 1965b]. Ceci simplifie grandement l'eq. (3) qui devient:

$$A(G,\delta,\Delta) = \exp\left\{-\gamma^2\delta^2G^2D(\Delta-\delta/3)\right\} . \tag{4}$$

Afin de simplifier encore plus cette équation, le concept de b-value fut introduit par [Le Bihan 1991]:

$$b = \gamma^2\delta^2G^2(\Delta-\delta/3) . \tag{5}$$

La b-value résume en une seule valeur la contribution de la magnitude G du DSG, de sa direction $\boldsymbol{g}$, de sa durée δ et de la durée Δ entre 2 DSG successifs. L'eq. (1.11) se lit donc:

$$A(b) = \exp\{-bD\} . \tag{6}$$

Une IRM de diffusion est déterminée par un nombre n_g de directions de DSGs et un nombre n_b de b)values.

2.3 L'imagerie du tenseur de diffusion

L'imagerie du tenseur de diffusion, dénotée DTI [Basser 1994b], est une généralisation du modèle de diffusion libre à la diffusion dans un faisceau. Le modèle est identique à l'exception de la matrice de covariance qui n'est plus proportionnelle à l'identité mais se lit $2(\Delta-\delta/3)D$, où D est un tenseur appelé *tenseur de diffusion*. L'estimation de ce tenseur D donne accès à de nombreux indicateurs utiles pour les cliniciens [Basser 1996a, Westin 2002]. Les deux principaux sont la diffusivité moyenne (MD) et la fraction d'anisotropie (FA). La MD est définie comme la moyenne des valeurs propres et la FA résume la déviation de chaque valeur propre à la MD en un coefficient compris entre 0 et 1. Une FA proche de 1 indique une valeur propre dominante et donc une orientation de diffusion prépondérante alors qu'une FA proche de 0 indique que toutes les valeurs propres sont semblables et que la diffusion est isotrope (dans ce cas, cela revient au modèle de diffusion libre). Ce modèle a suscité beaucoup d'enthousiasme pour étudier l'architecture de la matière blanche. Son estimation a été l'objet de nombreux travaux qui aboutirent au développement du cadre log-euclidien qui fournit une estimation rapide et robuste du tenseur de diffusion [Fillard 2007]. Beaucoup ont également étudié le nombre optimal de b-values (n_b) et de directions de DSGs (n_g) pour que cette estimation soit effectivement robuste. Dans une étude comparative, [Jones 2004] arrive à la conclusion qu'il suffit de prendre $n_b = 1$ et $n_g = 30$, ce qui permet d'obtenir une estimation très robuste du tenseur de diffusion à partir d'une acquisition qui ne dure que 5 minutes. Ce temps d'acquisition et la qualité des résultats ont fait du DTI le modèle de référence en utilisation clinique. En particulier, de nombreuses pathologies cérébrales ont été étudiées sous un angle nouveau grâce au DTI, à la fois en terme de différences de MD et FA par rapport à des sujets sains [Horsfield 2002, Sundgren 2004, Sexton 2009] et en termes de tractographie [Ciccarelli 2008]. Devant les promesses d'un tel modèle, de nombreuses bases de données ont donc été constituées par le passé spécifiquement

pour l'estimation du DTI (i.e., avec une seule b-value et moins de 30 directions de DSG).

Cependant, au fur et à mesure, les chercheurs et cliniciens se sont aperçus que la MD et la FA ne sont pas des biomarqueurs spécifiques dans le sens où, par exemple, une baisse de FA constatée chez des patients par rapport à des sujets sains ne reflète pas forcément une cause pathologique [Pierpaoli 1996a]. Les deux marqueurs sont également hautement corrélés [Shenton 2012]. La tractographie basée DTI a aussi ses limites car elle semble créer des faisceaux qui n'existent pas vraiment [Ciccarelli 2008]. Tous ces problèmes sont liés au fait que le DTI n'est valable que dans un voxel homogène, c-à-d un voxel qui ne contient qu'un seul environnement microstructurel tel qu'un faisceau ou un ensemble de cellules gliales identiques ou des molécules d'eau diffusant libre. La résolution spatiale typique d'une IRM de diffusion étant $2\times2\times2$ mm^3 et l'échelle d'un axone étant le μm, un voxel en général contient plusieurs faisceaux, des cellules gliales et des molécules d'eau diffusant librement. En d'autres termes, le DTI résume toutes ces contributions en une seule PDF qui, par exemple, sera isotrope dans le cas où deux faisceaux orthogonaux coexistent, résultant en une baisse de FA qui n'a pas lieu d'être. Le modèle Gaussien pour la diffusion au sein d'un seul faisceau a également été récemment remis en question par [Cheung 2009, Kunz 2011] qui ont observé des atténuations de signal non-monoexponentielles dans des faisceaux.

Un bon modèle de diffusion est donc un MCM qui inclut plusieurs compartiments pour prendre en compte les multiples environnements microstructurels dans un voxel et ainsi fournir des biomarqueurs plus spécifiques et des orientations de diffusion plus fidèles. De plus, la diffusion dans les faisceaux doit être modélisée par une distribution non-Gaussienne. Chaque environnement microstructurel a une forme géométrique particulière: les faisceaux sont cylindriques et les cellules gliales sphériques. Ces propriétés géométriques doivent donc guider la modélisation de la diffusion au sein de ces environnements. Par la suite, nous présentons succintement les MCMs existants qui sont basés sur la géométrie des cellules et leurs limitations et proposons un nouvel MCM qui permet de remédier à ces limitations. Nous élaborons ensuite un nouvel algorithme de tractographie qui utilise ce nouveau modèle pour reconstruire les faisceaux de matière blanche.

3 Modélisation locale de la diffusion

3.1 Les MCM existants

La façon la plus simplifiée de représenter un faisceau est de considérer que c'est un cylindre de rayon nulle, appelé "stick". Par ailleurs, dans un cerveau non pathologique, les membranes des cellules gliales sont perméables si bien que l'on peut considérer que les molécules d'eau dans cet environnement diffusent librement. Le modèle "ball-and-stick" [Behrens 2007] propose donc de considérer M faisceaux non-collinéaires et un environnement spécifique pour les molécules d'eau diffusant librement. La diffusion dans la direction des faisceaux $\pm\boldsymbol{\mu}$ est considérée libre et la

diffusivité dans ces directions est donc la même que la diffusivité de l'environnement libre d. La distribution des déplacements moléculaires dans un faisceau est modélisée par un tenseur de diffusion n'ayant qu'une seule valeur propre positive dans la direction du faisceau ($D^{\text{fascicle}} = d\boldsymbol{\mu}\boldsymbol{\mu}$) tandis que celle dans l'environnement libre est modélisée par un tenseur de diffusion proportionnel à l'identité ($D^{\text{iso}} = dI_3$). Malheureusement, ce modèle simplifie trop la géométrie des faisceaux. La diffusion n'est pas autorisée dans une direction autre que celle du faisceau ce qui n'est pas réaliste. Une version généralisée du modèle est proposée par [Hosey 2005] qui estime une diffusivité non nulle dans les directions perpendiculaires à celles des faisceaux. Cependant, les MCMs de la famille exponentielle avec des tenseurs dont la partie isotrope n'est pas nulle ne peuvent pas être estimés à partir de données à une seule b-value [Scherrer 2010]. De plus, d'après les observations de [Cheung 2009, Kunz 2011], une distribution Gaussienne n'est pas adéquate dans les faisceaux de toute manière.

Des modèles plus réalistes ont été proposés pour décrire la diffusion dans les faisceaux. [Assaf 2004] proposent le "composite hindered and restricted model of diffusion" (CHARMED) qui modélise de manière indépendente la diffusion le long de l'orientation du faisceau et la diffusion dans l'espace perpendiculaire à cette orientation. Spécifiquement, la diffusion le long de l'orientation du faisceau est modélisée par une Gaussienne à 1 dimension et la diffusion dans l'espace perpendiculaire est décrite par le modèle de Neumann [Neuman 1974], spécialement conçu pour décrire la diffusion contrainte dans un cylindre de rayon donné. Ce modèle n'est valide que lorsque le temps de diffusion Δ est du même ordre que la durée δ d'un DSG ($\Delta \sim \delta$), ce qui n'est pas le cas en pratique. [Assaf 2005b] corrige ce problème en utilisant le modèle de van Gelderen [van Gelderen 1994] valide pour $\delta < \Delta$, en lieu et place de celui de Neumann. Le CHARMED original nécessite de préspécifier la valeur du rayon des axones et n'intégre pas d'environnement pour la diffusion libre. La première limitation est comblée par [Assaf 2008a] qui montrent comment estimer une distribution de rayons d'axones en supposant qu'elle appartient à la famille des distributions Gamma. La seconde limitation est comblée par [Barazany 2009] qui montrent que l'on peut également incorporer un environnement de diffusion libre. Différemment, [Zhang 2012] proposent "neurite orientation dispersion and density imaging" (NODDI) qui modélise la diffusion dans un faisceau comme résultante de l'intégration de la diffusion dans une infinité de "sticks" dont l'orientation serait distribuée autour de celle du faisceau auquel ils appartiennent avec une certaine incertitude appelée dispersion en orientation (OD). Même si cela mène a une diffusion non-Gaussienne au sein du faisceau, il n'en demeure pas moins que les axones sont représentés comme des cylindres de rayon nul ce qui ne reflète pas la réalité. De plus, que ce soit CHARMED ou NODDI, ces modèles nécessitent des acquisitions spécifiques à multiples b-values qui sont pour l'instant trop longue pour un usage clinique et ne peuvent donc pas non plus être utilisés pour des études rétrospectives.

Notre but est donc de proposer un MCM incluant un environnement pour la diffusion libre et un environnement pour chaque faisceau dans lequel la diffusion est modélisée par une distribution non-Gaussienne de sorte que l'ensemble soit physiquement plausible et mathématiquement estimable à partir d'acquisitions à une seule

b-value et peu de directions de DSG.

3.2 Diffusion Directions Imaging

3.2.1 Diffusion with a fascicle

Dans un environnement microstructurel spécifique (e.g. faisceau ou molécules non contraintes), l'atténuation de signal est traditionnellement décrite par une monoexponentielle [Stejskal 1965b]:

$$A(b, \boldsymbol{g}) = \exp\{-b < \boldsymbol{g}, D\boldsymbol{g} >\} \ .$$

Cela revient à modéliser les déplacements moléculaires $\mathbf{x}$ comme $\mathbf{x} = \sqrt{2(\Delta - \delta/3)}\mathbf{y}$, où $\mathbf{y}$ suit une distribution Gaussienne de moyenne nulle et matrice de covariance D. Pour décrire une atténuation non-monoexponentielle, nous proposons une nouvelle distribution statistique non-Gaussienne pour modéliser $\mathbf{y}$.

Un faisceau est caractérisé, entre autres, par son orientation $\pm\boldsymbol{\mu}$. La diffusion dans un faisceau possède la propriété de symmétrie antipodale, ce qui signifie que les molécules diffusent dans les deux sens en proportions égales, ce qui s'écrit:

$$p_{\mathbf{y}}(\boldsymbol{y}|\pm\boldsymbol{\mu}, \Theta) = \frac{1}{2}p_{\mathbf{w}}(\boldsymbol{y}|+\boldsymbol{\mu}, \Theta) + \frac{1}{2}p_{\mathbf{w}}(\boldsymbol{y}|-\boldsymbol{\mu}, \Theta) \ , \tag{7}$$

où $p_{\mathbf{y}}$ est la PDF de $\mathbf{y}$, Θ est un ensemble de paramètres qui décrit le faisceau en dehors de son orientation et $\mathbf{w}$ est un vecteur aléatoire qui décrit le déplacement moléculaire dans une direction et un sens donnés.

Nous exprimons $\mathbf{w}$ comme la somme de deux variables aléatoires indépendantes, $\mathbf{w} = \mathbf{v} + \mathbf{z}$, où [5, 6]:

- $\mathbf{v}$ suit une distribution de von Mises & Fisher (cf. appendix A.2) sur la sphère de rayon $\sqrt{\nu d}$ ($\nu \in [0,1]$, $d > 0$) de direction moyenne $+\boldsymbol{\mu}$ ($\|\boldsymbol{\mu}\| = 1$) et paramètre de concentration $\kappa \geq 0$; sa norme est constante: $\|\mathbf{v}\|^2 = \nu d$.
- $\mathbf{z}$ suit une distribution Gaussienne de moyenne nulle et de matrice de covariance $\Sigma = \frac{(1-\nu)d}{\kappa+1}(I_3 + \kappa\boldsymbol{\mu}\boldsymbol{\mu}')$; son orientation est constante pour $\kappa \to \infty$ et est égale à $\pm\boldsymbol{\mu}$.
- $\mathbf{v}$ et $\mathbf{z}$ sont statistiquement indépendants.

Pour des valeurs grandes de κ, cette définition de $\mathbf{w}$ sépare complètement la variability directionnelle (capturée par $\mathbf{v}$) de la variabilité en amplitude (capturée par $\mathbf{z}$). La *direction moyenne* des déplacements s'identifie à $+\boldsymbol{\mu}$. Le paramètre $\kappa \geq 0$ mesure la *dispersion directionnelle* (plus κ est grand, plus la dispersion est faible). Le paramètre $d > 0$ mesure l'*échelle* des déplacements et le paramètre $\nu \in [0,1]$ est un *indice de non-Gaussianité* (plus il est près de 0, plus la diffusion est Gaussienne).

La PDF des déplacements dans le faisceau peut ensuite être obtenue analytiquement et est égale à:

$$\begin{aligned} p_{\mathbf{y}}(\boldsymbol{y}; \pm\boldsymbol{\mu}, \kappa, d, \nu) &= C(\kappa, d, \nu)\exp\left\{-\frac{(\kappa+1)y_\perp^2 + y_{/\!/}^2}{2(1-\nu)d}\right\} \\ &\times \int_{-1}^{1}\exp\left\{\frac{r_\nu\kappa}{2}t^2 + \kappa t\right\}\cosh\left(\frac{r_\nu y_{/\!/} t}{\sqrt{\nu d}}\right) I_0\left(\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}y_\perp\sqrt{1-t^2}\right)\mathrm{d}t \ , \end{aligned} \tag{8}$$

quel que soit $\boldsymbol{y} \in \mathbb{R}^3$, où:

$$\begin{aligned} r_\nu &:= \frac{\nu}{1-\nu}\,, \\ C(\kappa, d, \nu) &:= \frac{\kappa(\kappa+1)\exp\left\{-\frac{r_\nu(\kappa+1)}{2}\right\}}{2\left(2\pi(1-\nu)d\right)^{3/2}\sinh\kappa}\,, \\ (y_{/\!/}, y_\perp) &:= \left(<\boldsymbol{\mu}, \boldsymbol{y}>, \sqrt{\|\boldsymbol{y}\|^2 - <\boldsymbol{\mu}, \boldsymbol{y}>^2}\right)\,, \end{aligned}$$

et I_0 est la fonction de Bessel modifiée d'ordre 0 [Abramowitz 1972], avec la convention $<\boldsymbol{\mu}, \boldsymbol{y}> = \|\boldsymbol{y}\|$ quel que soit $\boldsymbol{y} \in \mathbb{R}^3$ quand $\kappa = 0$.

Les paramètres de cette distribution ont une interprétation intuitive:

- $\pm\boldsymbol{\mu}$ est l'*orientation* du faisceau,
- κ est un *indice de disperson en orientation*, qui quantifie la concentration de molécules autour de l'orientation du faisceau,
- d est la *diffusivité libre* le long du faisceau,
- ν est la *fraction d'occupation de l'espace extra-axonal* dans le faisceau.

Cette PDF peut être insérée dans l'eq. (3), dont l'intégrale peut alors être calculée analytiquement. Nous obtenons l'atténuation de signal attendue en fonction des paramètres du modèle:

$$\begin{aligned} A(b, \boldsymbol{g}) &= \exp\left\{-\frac{b(1-\nu)d}{\kappa+1}\left(1+\kappa<\boldsymbol{\mu}, \boldsymbol{g}>^2\right)\right\} \\ &\times \frac{\kappa}{\sinh\kappa}\begin{cases} \left|\operatorname{sinc}\left(\sqrt{2b\nu d - \kappa^2}\right)\right| & \text{, if } 2b\nu d \geq \kappa^2 \text{ and } \boldsymbol{g} \perp \boldsymbol{\mu}\,, \\ \dfrac{|\alpha\sinh\alpha\cos\beta + \beta\cosh\alpha\sin\beta|}{\alpha^2+\beta^2} & \text{, otherwise,} \end{cases} \end{aligned} \tag{9}$$

where:

$$\alpha = \sqrt{\frac{\operatorname{Re}(z)+|z|}{2}}\,,\ \beta = \frac{\operatorname{Im}(z)}{\sqrt{2(\operatorname{Re}(z)+|z|)}} \text{ and } z = \kappa^2 - 2b\nu d + 2i\kappa\sqrt{2b\nu d}<\boldsymbol{\mu}, \boldsymbol{g}>\,.$$

La PDF donnée par l'eq. (8) permet également de calculer analytiquement tous les moments de la distribution des déplacements moléculaires le long de l'orientation du faisceau et dans l'espace perpendiculaire à cette orientation.

En particulier, la diffusivité parallèle dans le faisceau vaut $D_{/\!/}^{\text{fascicle}} = d\left[1 - 2\nu\xi(\kappa)\right]$ et n'est pas sensible au modèle de diffusion employé. En effet, [Pierpaoli 1996b] rapportent $D_{/\!/} = 1.7 \times 10^{-3}$ mm^2/s en utilisant le DTI et [Veraart 2011b] rapportent $D_{/\!/} = 1.71 \times 10^{-3}$ mm^2/s en généralisant le DTI avec une kurtosis non nulle. Nous proposons donc la contrainte suivante que nous qualifions de *physiquement plausible*:

$$D_{/\!/}^{\text{fascicle}} = d\left[1 - 2\nu\xi(\kappa)\right] = 1.71 \times 10^{-3} \text{ mm}^2/\text{s}\,. \tag{10}$$

3.2.2 Modèle de diffusion complet

Pour prendre en compte de multiples environnements microstructurels, nous utilisons un MCM dans lequel nous intégrons

• un environnement de diffusion libre modélisée par une Gaussienne de moyenne nulle et de matrice de covariance $D^{\text{iso}} = d_{\text{iso}} I_3$ donnant lieu à une atténuation de signal attendue donnée par l'eq. (6);

• M faisceaux à l'intérieur desquels la diffusion est modélisée selon l'eq. (8) donnant lieu à une atténuation de signal attendue donnée par l'eq. (9).

Afin de rendre le modèle estimable à partir d'acquisitions à une seule b-value, nous appliquons quelques contraintes physiquement plausibles à ce modèle:

• nous appliquons la contrainte décrite par l'eq. (10);

• la diffusivité de l'environnement de diffusion libre est fixée à la valeur de diffusivité libre à 37°C: $d_{\text{iso}} = 3.0 \times 10^{-3}$ mm^2/s;

• la dispersion en orientation est entendue au niveau du voxel, ce qui implique que le paramètre de concentration κ est identique dans tous les faisceaux;

• la proportion d'espace extra-axonal est supposée être la même dans tous les faisceaux du voxel.

En résumé, nous proposons un nouvel MCM, appelé Diffusion Directions Imaging (DDI), qui contient $3M + 2$ paramètres indépendants, où M est le nombre de faisceaux dans le voxel. Spécifiquement, les paramètres indépendants du modèle sont:

• les angles $(\theta_i, \phi_i) \in [0, \pi] \times [0, 2\pi]$ $(i = 1, \ldots, M)$ qui définissent les orientations des faisceaux,
• les proportions $w_i \in [0, 1]$ $(i = 1, \ldots, M)$ de chaque faisceau dans le voxel,
• la dispersion moyenne en orientation $\kappa \geq 0$ dans les faisceaux,
• la proportion moyenne $\nu \in [0, 1]$ d'espace extra-axonal dans les faisceaux.

Nous avons conçu, en utilisant le software Camino[1], des données simulées de croisement de faisceaux à différents angles de séparation, utilisant différentes b-values, différents nombres de directions de DSG, différents rapports signal-sur-bruit (SNR). Nous avons validé sur ces données le DDI et nous l'avons comparé au modèle ball-and-stick et au modèle CHARMED tel qu'implémenté dans Camino. Les résultats montrent que nous sommes capables d'estimer le DDI à partir d'une seule b-value. La valeur de cette b-value ainsi que le nombre de directions de DSG importent peu, pourvu que le SNR soit élevé. La proportion moyenne d'espace extra-axonal dans les faisceaux est également estimée correctement.

Nous avons également estimé le DDI sur le cerveau d'un patient souffrant de tumeurs. Nous montrons que notre modèle propose des biomarqueurs qui semblent spécifiques et qui pourraient être d'un grand intérêt pour une étude plus détaillés de la microstructure de l'espace péritumoral.

Il reste un problème à traiter qui est inhérent tout MCM: le nombre de faisceaux dans un voxel donné est supposé être connu ce qui n'est pas le cas en pratique. Il s'agit d'un problème de sélection de modèle qui fait l'objet de la section suivante.

[1] http://cmic.cs.ucl.ac.uk/camino/

3.3 Détermination du nombre optimal de faisceaux

Deux types d'approches ont été envisagés pour répondre à la question du nombre optimal de faisceaux dans un voxel: les approches dites “brute-force" et le moyennage Bayesien. Dans la première approche, un ensemble de MCMs avec nombre de faisceaux croissants (de 0 à $M_{\max}$) est estimé et le nombre de faisceaux du modèle qui ajuste le mieux les atténuations de signal attendues à celles observées est retenu comme nombre optimal. La comparaison entre modèles reposent généralement sur un F-test [Alexander 2002, Kreher 2005, Scherrer 2012]. Puisque plus un modèle est complexe, mieux il s'ajuste aux données, le F-test favorise trop souvent les MCMs les plus complexes: cela s'appelle de l'“overfitting". Pour le limiter, [Schultz 2010] utilisent le critère d'information de Bayes (BIC) pour pénaliser la complexité du MCM qui croit avec le nombre de faisceaux. Récemment, [Scherrer 2013] proposent d'utiliser l'erreur de généralisation pour choisir le MCM sur la base de sa capacité à prédire de nouvelles données, ce qui contourne le problème de l'overfitting. Néanmoins, toutes ces méthodes recherchent le meilleur MCM uniquement parmi ceux estimés.

Le moyennage Bayésien, à l'inverse, consiste à intégrer la sélection de modèle en même temps que l'estimation. L'idée est d'estimer directement le meilleur MCM possible comme le MCM qui maximise une distribution a posteriori sur les modèles. Cela repose sur un choix approprié de distribution a priori sur les MCMs: [Poupon 2001, Demiralp 2011] utilisent des champs aléatoires de Markov et [Behrens 2007] utilisent l'“automatic relevance detection" (ARD) qui affecte des distributions a priori non informative à tous les paramètres du modèle sauf les poids de mélange qui sont distribués selon une distribution Beta. Une composante du mélange non supportée par les données est ainsi annulée du mélange. Ces méthodes ne sont cependant pas transférables en usage clinique car elles nécessitent des temps de calculs extrêmement longs et requiert une grosse quantité de données pour obtenir une distribution a posteriori qui ne dépende pas fortement de l'a priori.

Nous proposons donc une solution entre les deux précédentes options. A partir d'un ensemble de MCMs estimés, nous calculons la probabilité de chaque MCM d'être le modèle le plus proche du vrai modèle inconnu au sens de la divergence de Kullback-Leibler. Cette approche est très efficace car elle ne nécessite de ne connaitre que les critères d'information d'Akaike pour chaque modèle [Posada 2004]. Ces probabilités sont ensuite utilisées pour produire moyenner chaque paramètre entre les différents MCMs estimés et ainsi obtenir un modèle moyen avec des paramètres qui ne dépendent plus du nombre de faisceaux initialement choisi.

Nous avons comparé notre approche à l'ARD. Dans cette étude, nous fournissons une preuve empirique qu'il est peu utile de considérer des régions à 4 faisceaux dans la matière blanche. Nous montrons également que nous identifions de manière robuste les régions à 3 faisceaux telles que la corona radiata ou le centrum semi-ovale à partir de données cliniques, là où l'ARD échoue.

4 Tractographie de la matière blanche du cerveau

4.1 Algorithmes existants

Le dernier maillon de la chaine pour pouvoir reconstruire les faisceaux de matière blanche à partir des images de diffusion est justement l'algorithm de reconstruction en lui-même qui utilise le modèle de diffusion estimé pour dépeindre les faisceaux. C'est ce que l'on appelle la *tractographie*. Les premiers algorithmes de tractographie, qui reposent principalement sur le DTI, reconstruisent un faisceau en suivant, à partir d'un point donné, les orientations de diffusion localement estimées par un modèle de manière déterministe, c-à-d sans soucier de la variabilité de cette estimation. Ce sont les reconstructions dites "déterministe streamline" [Mori 1999, Xue 1999, Conturo 1999, Basser 2000].

Pour prendre en compte l'incertitude sur l'estimation des orientations locales de diffusion, les algorithmes plus récents peuvent être classés selon deux approches:

- ceux qui reconstruisent un faisceau globalement optimal à partir d'un point donné en minimisant une fonction de coût appropriée: parmi eux, les algorithmes de type "fast-marching" [Parker 2002, Prados 2006, Staempfli 2006], ceux basés sur la résolution de problèmes de type Hamilton-Jacobi-Bellman [Parker 2002, O'Donnell 2002, Jackowski 2005], ceux de type "spin glass" [Mangin 2002, Kreher 2008, Fillard 2009, Reisert 2011];
- ceux qui reconstruisent un faisceau en approximant localement une distribution d'orientations autour des orientations estimées par le modèle de diffusion [Parker 2003a, Parker 2003b, Parker 2005, Berman 2008].

Nous proposons une approche entre les deux solutions précédentes: nous cherchons un faisceau globalement optimal mais sans imposer de connaissances anatomiques a priori qui peuvent biaiser la reconstruction mais nous souhaitons nous concentrons sur l'incertitude au niveau du faisceau lui-même plutôt que sur l'incertitude locale de chaque orientation estimée. Notre approche repose sur l'utilisation du filtrage particulaire, introduite dans ce contexte par [Brun 2002, Björnemo 2002, Zhang 2009, Pontabry 2013].

Nous avons donc élaboré un filtre particulaire qui permet d'approximer une distribution de faisceaux à partir d'un point donné. Il est aisé d'imaginer qu'une telle distribution peut avoir de nombreux modes. Conséquemment, les filtres particulaires traditionnels ne peut convenir car ils sont connus pour ne pas bien maintenir la multimodalité de la distribution de filtrage qu'ils essaient d'approximer. Nous avons donc amélioré le filtre particulaire traditionnel.

4.2 Modélisation de la distribution de faisceaux

Un faisceau de longueur ρk, où k est un pas donné et $\rho > 0$ est la taille du pas supposée constante, est une séquence $X^k = [(\mathbf{x}_0, \mathbf{v}_{-1}), \dots, (\mathbf{x}_k, \mathbf{v}_{k-1})] \in \left(\Omega, \mathbb{S}^2\right)^{k+1}$ de paires successives de positions et de directions d'arrivée, reliées par $\mathbf{x}_{k+1} = \mathbf{x}_k + \rho \mathbf{v}_k$. L'espace d'état au pas k est l'ensemble des faisceaux possibles X^k initialisés à une position donnée $\mathbf{x}_0$ avec une direction d'arrivée donnée $\mathbf{v}_{-1}$.

La distribution de filtrage au pas k est la distribution de X^k, conditionnellement aux données de diffusion $\mathcal{Y}_k = \{y_{\mathbf{x}} : \|\mathbf{x} - \mathbf{x}_0\| \leq k\rho\}$ où $y_{\mathbf{x}} = \{S_{\mathbf{x},i}, \Psi_{\mathbf{x},i}\}_{i=1}^n$. S_i désigne une image d'atténuations de signal observées and Ψ_i une image d'atténuations de signal attendues. Au pas $k+1$, cette distribution est séquentiellement mise à jour en calculant la distribution de prédiction $p(X^{k+1}|\mathcal{Y}_k)$ à partir de la distribution de filtrage au pas précédent k et ensuite en utilisant la règle de Bayes pour obtenir $p(X^{k+1}|\mathcal{Y}_k)$. Ces étapes de prédiction et de mise à jour nécessitent de spécifier respectivement le modèle d'évolution $p(X^{k+1}|X^k)$ et la vraisemblance $p(\mathcal{Y}_{k+1}|X^{k+1})$.

Nous faisons l'hypothèse que les faisceaux peuvent être modélisés par des chaîne de Markov d'ordre 1. Conséquemment, le modèle d'évolution se simplifie en $p(\mathbf{v}_k|\mathbf{v}_{k-1})$. Suivant [Zhang 2009, Pontabry 2013], nous utilistons le modèle d'évolution suivant:

$$p(\mathbf{v}_k|\mathbf{v}_{k-1}) = \mathrm{vMF}(\mathbf{v}_k; \mathbf{v}_{k-1}, \kappa) = \frac{\kappa}{4\pi \sinh \kappa} \exp\{\kappa < \mathbf{v}_{k-1}, \mathbf{v}_k >\}\,, \tag{11}$$

où $\mathrm{vMF}(\cdot; \mathbf{v}_{k-1}, \kappa)$ est la distribution de von Mises & Fisher [Jupp 1989] sur la sphère unité ayant pour direction moyenne $\mathbf{v}_{k-1} \in \mathbb{S}^2$ et pour paramètre de concentration $\kappa \geq 0$. Le paramètre κ du modèle d'évolution contrôle la régularité des faisceaux reconstruits.

Désignons $\mathcal{B}_k = \{\mathbf{x} \in \Omega : \|\mathbf{x} - \mathbf{x}_0\| \leq k\rho\}$ et supposons que les données de diffusion soient spatialement indépendantes conditionnellement au faisceau, la vraisemblance $p(\mathcal{Y}_{k+1}|X^{k+1})$ peut alors être décomposée comme suit:

$$p(\mathcal{Y}_{k+1}|X^{k+1}) = \prod_{\mathbf{x} \in \mathcal{B}_{k+1} \setminus \mathbf{x}_{0:k+1}} p_0(y_{\mathbf{x}}) \prod_{j=0}^{k} p_1(y_{\mathbf{x}_{j+1}}|\mathbf{v}_j) \propto \prod_{j=0}^{k} \frac{p_1(y_{\mathbf{x}_{j+1}}|\mathbf{v}_j)}{p_0(y_{\mathbf{x}_{j+1}})}\,, \tag{12}$$

où $\mathbf{x}_{0:k+1}$ est la chaine polygonale définie par X^{k+1}, $p_1(\cdot|\mathbf{v}_j)$ est la distribution de bruit de mesure en supposant que $\mathbf{v}_j$ est la direction locale du faisceau et p_0 est la distribution du bruit de mesure.

Afin de mieux maintenir la multi-modalité de la distribution de filtrage, nous empruntons les idées de [Vermaak 2003] et la formulons comme un mélange à M_k composantes:

$$p(X^k|\mathcal{Y}_k) = \sum_{m=1}^{M_k} \pi_{m,k} p_m(X^k|\mathcal{Y}_k)\,, \tag{13}$$

où $\sum_{m=1}^{M_k} \pi_{m,k} = 1$. Une telle formulation nous permet d'accomplir la récursion de filtrage pour chaque composante p_m séparément, à supposer que chaque poids du mélange soit mis à jour comme la vraisemblance pondérée normalisée pour la composante correspondante.

4.3 Filtre particulaire multi-modal

En régle général, il n'y a pas d'expression analytique pour les équations de la récursion de filtrage. Une stratégie fréquente consiste à recourir aux filtres particulaires. Ils approximent la distribution de filtrage par un ensemble d'échantillons appelés *particules* à qui sont attribués des poids appropriés de façon à ce que l'ensemble reflète la distribution de filtrage à chaque pas.

Les filtres traditionnels approximent les distibutions comme celle donnée en eq. (13) avec $M_k = 1$. La seule différence lorsque $M_k > 1$ est que les particules sont regroupés dans des “clusters" et le ré-échantillonnage est fait dans chaque cluster, ce qui donnent plus de chances aux faisceaux secondaires. Deux sortes de poids sont donc calculés: ceux des particules et ceux des clusters, avec des règles de mises à jour en adéquation avec le “sequential importance sampling". Ainsi, on peut voir chaque cluster comme une super-particule. L'ensemble des particules dans un cluster donné m a pour but d'approximer la composante p_m de distribution de filtrage donnée en eq. (13).

Utilisons les notations de [Vermaak 2003]. Soit $\mathcal{P}_k = \{M_k, \Pi_k, \mathcal{X}_k, \mathcal{W}_k, \mathcal{C}_k\}$ la représentation en particules de la distribution de filtrage où M_k est le nombre de composantes, $\Pi_k = \{\pi_{m,k}\}_{m=1}^{M_k}$ l'ensemble des poids de composantes, $\mathcal{X}_k = \{\mathbf{x}_k^{(\ell)}\}_{\ell=1}^{N}$ l'ensemble des N particules, $\mathcal{W}_k = \{w_k^{(\ell)}\}_{\ell=1}^{N}$ l'ensemble des poids de particules et $\mathcal{C}_k = \{c_k^{(\ell)}\}_{\ell=1}^{N}$ l'ensemble des indicateurs d'appartenance aux composantes (i.e., $c_k^{(\ell)} = m$ si la particule ℓ apparient à la composante m). Etant donné $\mathcal{P}_k$, la représentation particulaire évolue au pas $k+1$ en 5 étapes:

Proposition de nouveaux échantillons: Nous générons aléatoirement des nouvelles directions pour prolonger les faisceaux, selon une densité de proposition $q(\cdot|\mathbf{v}_{k-1}^{(\ell)}, \mathcal{Y}_k)$ qui dépend de la direction précédente et du modèle de diffusion au pas k:

$$\mathbf{v}_k^{(\ell)} \sim q(\mathbf{v}_k|\mathbf{v}_{k-1}^{(\ell)}, \mathcal{Y}_k) \text{ and } \mathbf{x}_{k+1}^{(\ell)} = \mathbf{x}_k^{(\ell)} + \rho\mathbf{v}_k^{(\ell)}\,. \tag{14}$$

Mise à jour des poids de particules: Les faisceaux sont prolongés et leur poids sont mis à jour de sorte que le nouvel échantillon soit représentatif de la distribution de filtrage au pas $k+1$ [Doucet 2000]:

$$\tilde{w}_{k+1}^{(\ell)} = \frac{w_k^{(\ell)} p_1(y_{\mathbf{x}_{k+1}^{(\ell)}}|\mathbf{v}_k^{(\ell)}) p(\mathbf{v}_k^{(\ell)}|\mathbf{v}_{k-1}^{(\ell)})}{p_0(y_{\mathbf{x}_{k+1}^{(\ell)}}) q(\mathbf{v}_k^{(\ell)}|\mathbf{v}_{k-1}^{(\ell)}, \mathcal{Y}_k^{(\ell)})}\,. \tag{15}$$

La normalisation de ces poids est faite dans chaque composante séparément:

$$w_{k+1}^{(\ell)} = \frac{\tilde{w}_{k+1}^{(\ell)}}{\sum_{j\in\mathcal{I}_{m,k}} \tilde{w}_{k+1}^{(j)}}\,, \tag{16}$$

où $\mathcal{I}_{m,k} = \{\ell \in [\![1, N]\!] : c_k^{(\ell)} = m\}$ est l'ensemble des indices des particles qui appartiennent à la m-ième composante de la distribution de filtrage au pas k.

Mise à jour des poids de composantes: Les poids de composantes doivent être mis à jour de manière correcte pour s'assurer que le filtre particulaire agit toujours sur chaque composante individuellement:

$$\pi_{m,k+1} = \frac{\pi_{m,k}\tilde{w}_{m,k+1}}{\sum_{i=1}^{M} \pi_{i,k}\tilde{w}_{i,k+1}} \text{ with } \tilde{w}_{m,k+1} = \sum_{\ell\in\mathcal{I}_{m,k}} \tilde{w}_{k+1}^{(\ell)}\,. \tag{17}$$

Ré-échantillonnage dans chaque composante: Afin d'éviter la dégénérescence des poids de particules, un ré-échantillonnage peut être nécessaire [Doucet 2000].

Cette étape peut être effectuée dans chaque sous-ensemble de particules associé à une composante de la distribution de filtrage indépendemment, selon les poids des particules de ce sous-ensemble [Vermaak 2003]: nous calculons le nombre effectif de particules dans chaque sous-ensemble comme:

$$\mathrm{ESS}_m = \left(\sum_{\ell \epsilon \mathcal{I}_{m,k}} \left(w_{k+1}^{(\ell)} \right)^2 \right)^{-1}, \tag{18}$$

et nous déclenchons le ré-échantillonnage dans un sous-ensemble de particules si ESS_m est en-dessous d'un seuil fixé $\alpha|\mathcal{I}_{m,k}|$, où $|\cdot|$ est l'opérateur *taille d'un ensemble*. Dans ce cas, le ré-échantionnage est effectué à partir d'une distribution catégorique comprenant $|\mathcal{I}_{m,k}|$ catégories dont les probabilités d'événement sont données par l'eq. (4.7).

Classification des particules "prolongées" dans de nouveaux clusters: Le nombre de composantes M_k de la distribution de filtrage est inconnu. Ainsi, à la fin de chaque pas, il est dynamiquement estimé en fusionnant et/ou en séparant en deux certaines composantes: M_k, $\mathcal{C}_k$ et $\mathcal{I}_{m,k}$ sont alors mis à jour en M_{k+1}, $\mathcal{C}_{k+1}$ et $\mathcal{I}_{m,k+1}$ en conséquence.

Nous avons implémenté cet algorithme de tractographie avec le DDI mais aussi avec le DTI pour comparaison. Dans la reconstruction du faisceau cortico-spinal, nous concluons que notre algorithme permet de retracer toute son étendue. Nous démontrons également par le biais de la tractographie que le DDI permet une reconstruction plus détaillée des faisceaux que le DTI. Toutes ces conclusions sont une nouvelle fois tirées sur des données cliniques.

5 Tractographie péri-tumorale préchirurgicale & Conclusions

Les outils développés dans cette thèse ont pour but d'être utilisés en pratique clinique. Nous avons en particulier élaboré un nouveau modèle de diffusion, que nous avons dénommé Diffusion Directions Imaging, qui permet d'estimer d'inférer de manière détaillée et précise la microstructure de la matière blanche à partir d'images de diffusion couramment acquises en clinique. Nous avons combiné ce modèle avec un algorithme de tractographie qui effectue la reconstruction à proprement parler des faisceaux de la matière blanche. Cet algorithme repose sur l'estimation d'une distribution de faisceaux à partir d'un point donné qui prend donc compte l'incertitude commise sur chaque faisceau généré. A partir de données simulées et d'images cliniques acquises surdes sujets sains, nous avons montré que l'ensemble de ces outils offrent des biomarquers spécifiques et une très bonne reconstruction du CST.

Pour renforcer ces résultats, nous avons participé aux trois éditions du challenge MICCAI intitulé "DTI Tractography Challenge". L'idée est d'étudier la disposition des faisceaux de matière blanche dans les régions péri-tumorales pour le pré-planning chirurgical. Dans ce contexte, les images de diffusion acquises en pré-op sont encore plus contraintes qu'à l'accoutumée. Nous avons donc eu à travailler avec des images à

très basse résolution angulaire (20 directions de DSG), très basse résolution spatiale ($1.2 \times 1.2 \times 5.2$ mm^3) et une seule b-value b = 1000 s/mm^2. L'objectif pratique était de reconstruire le CST à partir de telles données. Chaque année, nous avons obtenus des résultats en accord avec l'anatomie connue de ce faisceau malgré la mauvaise qualité inhérente aux images acquises en pré-op. De plus, nous montrons l'amélioration de nos outils dans le temps en comparant les résultats de l'année 2012 avec les outils dans leur version de l'époque et les outils actuelles. Le résultat de la comparaison montre que nous délivrons ajourd'hui des outils de qualité prêts à l'utilisation clinique. Les temps de calculs sont aussi raisonnables puisqu'il faut compter deux heures pour l'estimation du modèle de diffusion qui peut être effectuée off-line et seulement cinq minutes pour la tractographie online.

Introduction

Context

Diffusion MRI (dMRI) is a modality in MRI that is able to depict the microstructure of the tissues in vivo and non-invasively. The key idea is to study how the Diffusion-Induced Molecular Displacements (DIMD) are constrained by the geometry of the environment in which they occur. The **White Matter (WM)** is one of the two components of the CNS and is mainly composed of glial cells that can be thought as spheres and bundles of myelinated axons, referred to as **fascicles**, that can be represented by cylinders. Diffusion in the WM has been proven to be anisotropic, which mainly reflects diffusion in the fascicles. dMRI thus appears as a particularly appealing tool for the study, **reconstruction** and analysis of the WM fascicles.

These tasks are however non-trivial. On one hand, the *study* of WM fascicles requires to extract the intra-fascicle DIMD out of the diffusion data. dMRI only provides an undirect measure of the random DIMD through the magnetization of water spins. Estimating the probability distribution of the random DIMD out of the diffusion data is known as a **local diffusion modeling** problem. Many diffusion models have been devised to this end. Microstructural parameters describing the different environments in a given voxel such as fascicle orientations, fractions of occupancy of the fascicles, proportion of free water or proportions of intra- and extra-axonal spaces can now be quantified. On the other hand, the *reconstruction* of WM fascicles uses *inter alia* the fascicle orientations locally identified by a diffusion model as inputs to generate global tracts composing a given fascicle: this is often dubbed **tractography**. Many tractography algorithms have been proposed in the literature. While the first proposals merely reconstructed tracts from a single voxel by following the local orientations provided by the underlying diffusion model, the most recent alogrithms try to adequately account for the uncertainty on the estimated local orientations using either global optimization methods or probabilistic frameworks.

Notwithstanding the recent promising diffusion models, dMRI is not a standard modality in clinical MRI examination protocols. Clinicians are limited by acquisition time because they are dealing with real patients who cannot stay long in the scanner, leading to low spatial and angular resolution diffusion data. Their expectations about local diffusion modeling are to find, from this low quality data, **specific** biomarkers for brain disorders, i.e., microstructural parameters whose variations in patients with respect to healthy subjects are only the reflection of the underlying pathology and not the translation of a flaw in the model. Diffusion models that can be fitted on such data reduce to Diffusion Tensor Imaging (DTI), which unfortunately, has been shown to provide non-specific biomarkers. Some recent diffusion

models provide specific biomarkers for some brain disorders but require diffusion data that are not affordable in clinics. Moreover, resorting to such new models prevents one from retrospectively analyzing already acquired clinical diffusion data. Last but not least, clinicians are used to DTI. Most recent models depart quite a lot from this modeling and would require some training for the medical staff, which makes them hardly transferable to clinical practice. For its part, tractography led to one of the biggest disappointments in the field. Even if it was impressive at first glance, neurosurgeons are still not convinced of its benefits. WM fascicles are not visible to the naked eye and, thus, we do not know the *ground truth* about their structural organization. However, neurosurgeons do have some knowledge about it and, at least, some known fascicles are expected to be accurately reconstructed from tractography. Neurosurgeons thought they will learn more than they already know about WM fascicles thanks to tractography. It turns out that, in fact, tractography hardly confirmed, so far, what is already established.

Notations

Throughout the manuscript, the following notations and terminologies are adopted:

- $\Omega \subseteq \mathbb{R}^3$ denotes the image spatial domain,
- An image over this domain is a set $I = \{i_{\boldsymbol{x}}, \boldsymbol{x} \in \Omega\}$; for sake of clarity, the image signal $i_{\boldsymbol{x}}$ at location $\boldsymbol{x}$ will be denoted by the entire image symbol I unless it makes the discussion unclear,
- $\boldsymbol{q} = \gamma\delta G\boldsymbol{g}$ denotes a Diffusion-Sensitizing Gradient (DSG) applied in direction $\boldsymbol{g}$ with magnitude G during a period δ (the **pulse duration**); γ is the gyromagnetic ratio specific to the excited spin ($\gamma = 2.68 \times 10^8$ rad/s/T for the proton spin),
- $S = S(\boldsymbol{q})$ denotes a **measured Diffusion-Weighted (DW) image** or **signal** under the application of a DSG $\boldsymbol{q}$,
- $\Psi = \Psi(\boldsymbol{q})$ denotes the corresponding **predicted DW image** or **signal** from a specific diffusion model or in absence of measurement noise,
- S_0 denotes a **measured unweighted image** or **signal**,
- Ψ_0 denotes the **true unweighted image** or **signal** in absence of measurement noise, which, in this work, is always the average of all the measured unweighted images/signals,
- $A = A(\boldsymbol{q})$ denotes the **Signal Intensity Decay (SID)** predicted by a specific diffusion model or in absence of measurement noise under the application of a DSG $\boldsymbol{q}$.

Organization and contributions of the thesis

The work of this thesis aims at providing tools for analyzing clinical diffusion data that yield specific biomarkers for brain disorders and faithful reconstruction of the known WM fascicles. The contributions are twofold: (i) we worked on local diffusion modeling to propose a diffusion model that provides as many microstructurally

relevant parameters as possible given single-shell low angular resolution clinical diffusion data and (ii) we developed tractography algorithms that are able to accurately reconstruct complex known fascicles such as the Cortico-Spinal Tract (CST). Such analysis tools could be then used to go beyond the known architecture of the WM in clinics. The manuscript is divided into five chapters.

Chapter 1: State of the art

In this chapter, we first outline a brief historical introduction to dMRI from its origin to the most recent developments. We motivate the growing enthusiasm that this imaging technique generated towards the understanding of the human brain, pointing out a number of ongoing research projects on the subject that have received substantial financial supports from many medical care institutes. Then, in order to better appreciate the ins and outs of the contributions of this thesis, we summarize the necessary background about (i) the anatomy of the brain and the CNS and (ii) the principles of dMRI from the very definition of the diffusion phenomenon to its undirect measurement. Finally, we present the theory and practice behind DTI, which is currently the means to analyze diffusion data in clinics. We explain what has been learned so far on both healthy and pathological human brains thanks to DTI. We also stress the limitations of such an approach and we discuss alternative tractography-driven analysis tools.

Chapter 2: Local modeling of the diffusion

In this chapter, we discuss a particular class of diffusion models, termed Multi-Compartment Models (MCMs), which seem to be well suited to characterize the diffusion process within specific microstructural environments. Indeed, these models express the voxelwise distribution of DIMD as a finite mixture of different distributions, each of them characterizing the DIMD in particular tissues, such as fascicles, glial cells, etc. In particular, we focus on MCMs that constrain the DIMD by taking into account the geometrical properties of the environment in which molecules are trapped: we refer to these models as *geometry-based* MCMs. Geometrical constraints can be set either by explicitly including parameters that describe the geometry of the tissues (e.g., fascicles can be modeled as cylinders whose radius explicitly takes part to the predicted DW signal) or by designing an appropriate shape for the diffusion profile that fits within a specific tissue (e.g., cigar-shaped diffusion is suited to fascicles modeled as cylinders, with no explicit intervention of the radius of the cylinders). We describe the geometry-based MCMs proposed in the literature and we discuss their advantages and drawbacks. We finally propose a new MCM for the WM, which, under few physically plausible constraints, can be estimated from single-shell low angular resolution dMRI. This model provides many microstructural information regarding the fascicles: their orientation, their fraction of occupancy in the voxel, the resulting global orientational dispersion in the voxel, the proportion of surrounding Cerebro-Spinal Fluid (CSF) and, when supported by the data, the

fraction of occupancy of extra-axonal space. We show that these parameters may be good biomarkers for patients with tumors. For example, we demonstrate that the usual Fractional Anisotropy (FA) is the result of the combination of a given orientational dispersion, a given extra-axonal occupancy and a given CSF occupancy and we show that the FA is low but homogeneous in the peritumoral region whereas parameters defining the FA vary quite significantly.

Chapter 3: Selection-free estimation of Multi-Compartment Models

Despite their ability to model the diffusion at a sub-voxel scale, MCMs suffer from a major drawback: they require one to know the right number of compartments in each voxel. In the WM, the question often reduces to "how many non-collinear fascicles coexist in a given voxel?" This is a model selection problem that precludes the use of MCMs. In this chapter, we review the different approaches that have been employed in the dMRI literature to overcome this problem. Two main methods can be distinguished: **brute force** approaches and **averaging** approaches. The first ones aim at providing a map with the optimal number of fascicles in each voxel while the second ones rather assign to each model of a candidate set its probability to fit the data and they output an averaged model in each voxel given the data. Brute force approaches have been widely used in the dMRI community and we briefly summarize the different contributions. On the contrary, the only averaging approach to date used for the selection of the optimal number of fascicles is Automatic Relevance Detection (ARD). We discuss this method, which performs an implicit model averaging within a fully Bayesian framework. We then propose an explicit way of averaging models based on the Akaike Information Criterion (AIC) with no need to resort to time-consuming Markov Chain Monte Carlo (MCMC) procedures. By applying this explicit averaging approach to the model proposed in Chapter 2, we show in particular that we do not need to depend on an ad-hoc b-value threshold under which it is pointless to try to estimate proportions of intra- and extra-axonal spaces: the averaging automatically prunes the extra-axonal spaces if the data do not support their existence.

Chapter 4: Brain White Matter tractography

In the previous chapters, we designed a new MCM that provides many microstructural parameters even from single-shell low angular resolution dMRI (typical characteristics of clinical data). In this chapter, we aim at presenting tractography algorithms that are suited to multi-fascicle models (and, hence, to MCMs) and that perform the actual reconstruction of the fascicles fast enough and accurately enough to be translated to clinics. We review the plethora of tractography algorithms in the literature. Most of them model a **fascicle** as a number of **tracts** that are successive pairs of positions and arrival directions. We observe that (i) deterministic streamline algorithms have mainly been defined for DTI but have hardly been extended to multi-fascicle models and (ii) the objects of interest in most algorithms are

the local orientations of the tracts that are sometimes adjusted by global optimization or probabilistic tools. As a consequence to point (i), we propose a deterministic streamline algorithm based on the FACT algorithm. We extended it to multi-fascicle models and we show its great ability to reconstruct the hand motor tracts of the CST in conjunction with the MCM proposed in Chapter 2. As a consequence to point (ii), we propose a probabilistic tractography algorithm that approximates the distribution of all possible tracts of a fascicle. It relies on a particle filter in which samples ("particles") from entire tracts are propagated. The objects of interest are thus directly the tracts rather than their local orientations. This tractography algorithm is not *stricto sensu* new in the dMRI community. However, previous published algorithms have hard time dealing with the fact that this distribution of tracts can have many modes (e.g., the CST can be seen as a multi-modal distribution of tracts with one mode per different cortical areas). Traditional particle filters proposed in the literature often keep only one of them, failing to reconstruct the full extent of the CST. We demonstrate that, in conjunction to the MCM proposed in Chapter 2, our particle filter proves better at maintaining the multi-modality of the distribution of interest.

Chapter 5: MICCAI DTI Tractography Challenges

In this last chapter, we employ all the previously developed tools to reconstruct the CST in patients with tumors of different natures. We volonteered to participate to all three editions of the MICCAI DTI tractography challenge since 2011 to face the problem of diffusion modeling and tractography on the field, with all the constraints that it implies. The organizers provided us with diffusion data acquired in the context of presurgical planning of tumor removal. At the proximity of the surgery, the neuroradiologists can only afford short-time dMRI, resulting in single-shell low angular resolution diffusion data. Spatial resolution is also limited with a trade-off between in-plane resolution and slice thickness to achieve very short acquisition times. We present the results that we obtained using such data for the reconstruction of the CST. In the first edition, we used a very preliminary version of the MCM introduced in Chapter 2 without model selection/averaging (we fitted two fascicles everywhere). In the second edition, we proposed an ad-hoc model selection based on the value of an index derived from the eigenvalues of the diffusion tensor. In these first two editions, we used our deterministic streamline algorithm. In the third edition, we used the complete MCM described in Chapter 2, we averaged it according to the explicit averaging method proposed in Chapter 3 and we performed the tractography using the adaptive multi-modal particle filter presented in Chapter 4. The quality of the diffusion data was even poorer for the third edition but the use of the complete model with the probabilistic tractography yield the best results over all three editions, providing the full extent of the CST away from the tumors and depicting the rearrangement of the tracts in the vicinity of the tumor.

Chapter 1

State of the art

Contents

1.1 Context

1.1.1 Advent and development of Magnetic Resonance Imaging

In vivo imaging of the human body was made possible by the discovery of X-rays by [Rontgen 1896]. This gave rise, inter alia, to the *cerebral angiography* pioneered

in 1927 by Egas Moniza at the university of Lisbon, which is an imaging technique that provides images of blood vessels in the brain using a contrast agent and to the *Computed Tomography (CT) scan* introduced by [Hounsfield 1973], which is an imaging technique that aims at reconstructing images of anatomical structures of the human body from the measure of X-rays absorption by the tissues. X-ray-based imaging techniques remain however all invasive for the human body and rely on ionizing radiations that happen to have adverse effects [De Oliveira 1987].

A major leap forward *in-vivo* non-invasive imaging of the human body has been the introduction of MRI by [Lauterbur 1973, Damadian 1974, Mansfield 1977], for which both Paul Lauterbur and Sir Peter Mansfield received the 2003 Nobel price in Physiology or Medicine. This imaging technique is based on the phenomenon of resonance of the protons in presence of a magnetic field [Bloch 1946, Purcell 1946]. MRI has the advantage, with respect to Xrays and CT scans, to be non-invasive and there are no known scientific evidence of harmful effects of a magnetic field on the human body at the current range (until now always less than 3T). It also provides a spatial resolution comparable to the previous invasive imaging techniques with a far better contrast resolution due to the large number of different modalities that have been devised since its advent: T1-weighted, T2-weighted, FLuid Attenuated Inversion Recovery (FLAIR), proton density, functional and diffusion MRI are some examples of possible modalities (see [Kastler 2011] for a review on the different MRI modalities). Each modality has been developed to obtain a good contrast for a specific pathology: for instance, T2-weighted MRI is particularly sensitive to the development of edemas in patients with tumors.

1.1.2 Diffusion MRI & Human Brain

Although it will be described in depth in section 1.3, we would like here to spend a few words on the origin of dMRI. It has been introduced by [Le Bihan 1985, Merboldt 1985, Taylor 1985] to quantify the diffusion of water in a body. It became rapidly of great interest for its very good contrast resolutions for various pathologies therefore providing valuable complementary information to other MRI modalities. For example, it helps in better delineating lesions in patients with acute stroke (fig. 1.1,right), providing additional contrasts with respect to the traditional FLAIR sequence (fig. 1.1,left), which, said briefly, images the tranverse relaxation time of the tissues. dMRI also supports the diagnosis of many other cerebral pathologies such as tumors [Kono 2001], abscesses [Chang 2002], Creutzfeldt-Jakob Disease (CJD) [Shiga 2004], Multiple Sclerosis disease [Larsson 1992] and so on.

It also became tremendously popular and can be considered as an MRI modality of reference for the study and analysis of the microstructure of the brain tissues when [Moseley 1990] showed that diffusion in the brain is anisotropic, i.e., that it is constrained by the geometry of the tissues. This means that dMRI is a tool of choice for inferring brain tissues microstructure [Le Bihan 2003].

Due to the lack of appropriate imaging techniques, the brain remains one of the most enigmatic organ of the human body. The work of [Moseley 1990] has stimu-

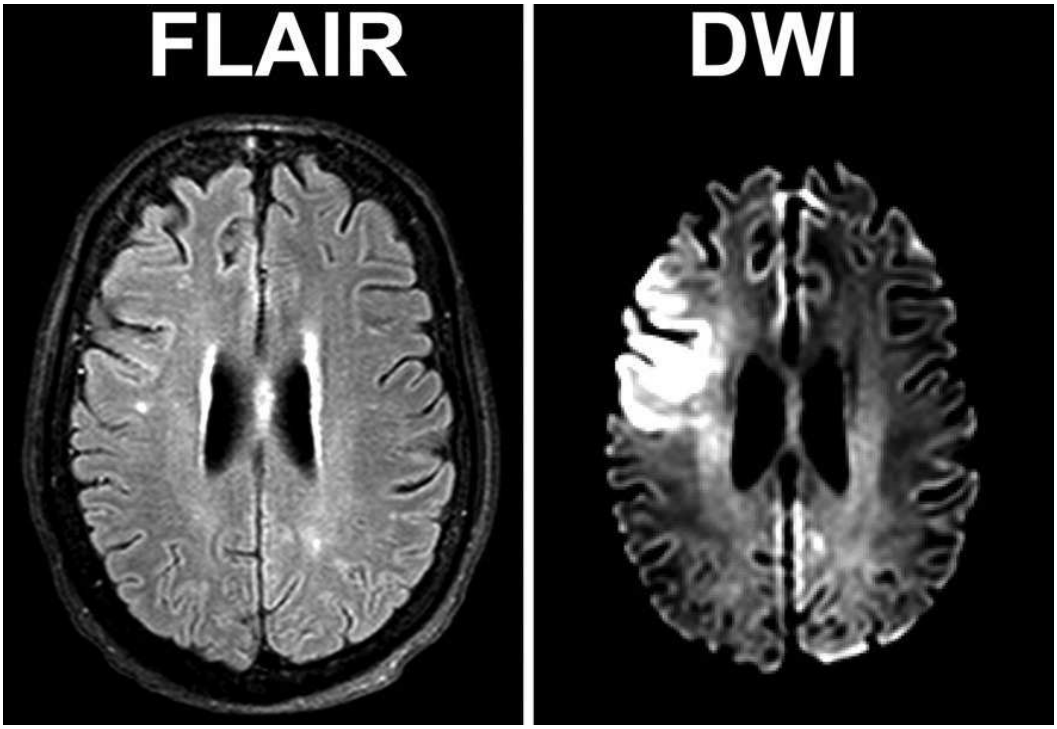


Figure 1.1: **Comparison of FLAIR and dMRI.** Comparison of FLAIR and dMRI (DWI stands for diffusion-weighted imaging) modalities in the detection of tissue changes in a patient with acute stroke (image courtesy of [Bokkers 2012]): dMRI (right) provides additional valuable contrasts with respect to the traditional FLAIR sequence (left).

lated the scientific community for understanding its architecture. Many progresses towards this goal have been achieved over the past two decades. So far, the scientific community managed to gather enough information on the brain in order to design the appropriate analysis tools. This knowledge together with the advent and constant evolution of the MRI scanners have triggered a series of projects that aim at solving the mysteries of the human brain, among which:

The Brain Molecular Anatomy project. This project includes many different projects and has two main goals: (i) catalog all genes in both the healthy and pathological nervous systems and (ii) identify the different types of neurons in the brain.

The CONNECT project (2010). This project aims at studying the structural connectivities at both the micro- and macroscropic levels in the brain. The underlying hypothesis is that micro-structural and macro-structural parameters extracted using dMRI can help in charaterizing brain morphology and connectivity.

The Connectome project (2010). This project focuses on mapping neuron connectivity in different species. Using two imaging modalities that are laser-scanning light microscopy and semi-automated electron microscpopy, they seek to

obtain high-resolution imaging of the neural network.

The One Mind For Research initiative (2011). The idea here is to gather data and knowledge for the understanding of brain disorders. It is an effort to reduce the social and economic effects of brain injury and mental illness.

The Human Brain project (2012). This project encourages a better study of brain pathologies by attempting to create new biologically grounded classifications, based on key similarities and differences between previously published isolated studies of individual disorders. The main goal of the project is to provide the scientific community with a brain simulation platform.

The Projectome project (2012). This project proposes to set up high performance computational infrastructure to process brain data and to design the corresponding processing tools for analyzing and visualizing the data.

The NIH Human Connectome project (2013). This is one of the two most recent projects that put efforts in providing a compilation of neural data together with a graphical interface to navigate this data.

The NIH Blueprint Human Connectome Project (2013). This project is similar to its homonym and aims at *mapping the human connectome as accurately as possible in a large number of normal adults* and at *making this data freely available to the scientific community using a powerful, user-friendly informatics platform.*

All of these projects were financed substantially by various institutes, which hightlights the importance that is granted to this field of research. In this context, the content of this manuscript aims at playing a part in this important and challenging initiative that gather many scientists and we hope that it will make a significant contribution towards brain microstructure understanding.

In the remainder of this chapter, we will give a more in-depth description of the brain (Section 1.2) and we will go into details regarding the dMRI modality (Section 1.3). We will also explain how the acquired dMRI datasets are processed in order to be feasible for understanding the microstructure of brain tissues (Section 1.4). We will see that this approach not only provided new insights into brain disorders but also yielded big collections of data that are shared in the community to try to understand even better some pathologies. Finally, we will expose the limitations of the current way of analyzing dMRI datasets and some convenient solutions for the reconstruction of the neural network (Section 1.5).

1.2 Description of the brain

The brain along with the spinal cord form what is called the CNS (see fig. 1.2), which is the part of the nervous system that coordinates the activity of the different part of the human body. It receives and treats the information, which is then sent to the various parts of the body by the Peripheral Nervous System (PNS).

The brain, or *cerebrum*, is one part of the CNS. Its outermost sheet of neural tissues is the **cerebral cortex**, which is made of four main lobes:

- the **frontal lobe** responsible for the consciousness,

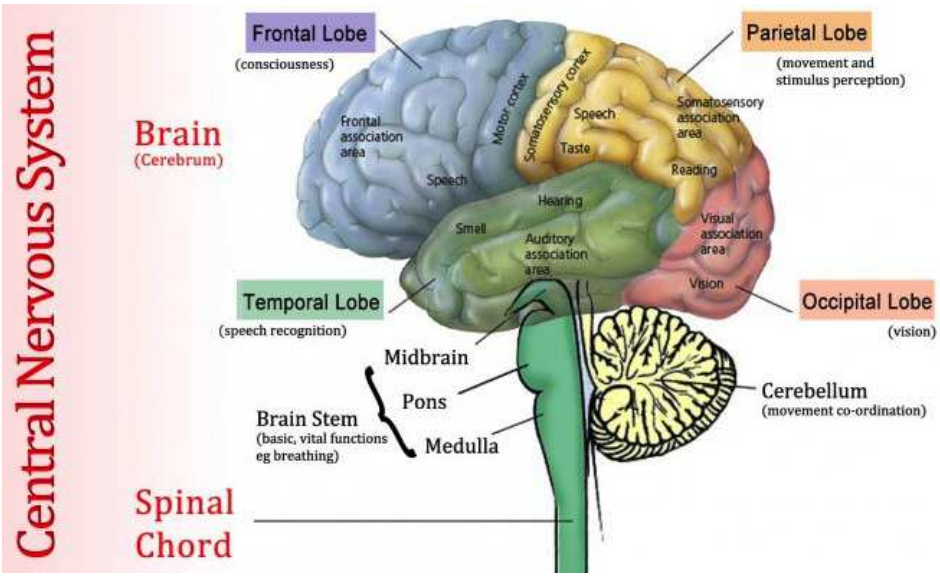


Figure 1.2: **Schematic view of the Human CNS.** Image courtesy of http://climatereview.net/ChewTheFat/?attachment_id=1061.

- the **temporal lobe** responsible for speech recognition,
- the **parietal lobe** responsible for movement and stimulus perception,
- the **occipital lobe** responsible for vision.

Other parts of the brain include the **brain stem** responsible for basic vital functions like breathing and the **cerebellum** responsible for movement coordination. This structure is illustrated in the upper part of fig. 1.2.

This section is dedicated to a brief presentation of the CNS. The interested reader can refer to [Duvernoy 1999] for a more in-depth description. The CNS can be depicted either at the microscopic level (Section 1.2.1), which describes the different cells composing it or at the macroscopic level (Section 1.2.2), which consists in classifying the cells into two categories whether they are meant to *transmit* or *process* messages. The brain neural transmission network is rather complicated in the brain and will be detailed in section 1.2.3.

1.2.1 Microscopic description

There are two types of cells that compose the CNS: the neurons and the glial cells. The neurons are responsible for the transmission of nerve impulses whereas glial cells have more a supporting and protecting role. There are a lot more of glial cells than neurons in the brain (factor of 3).

The neuron. The main function of a neuron is the transmission of nerve impulses. To this purpose, its typical structure is given by fig. 1.3 and comprises:

- a **cell body** at the heart of which lies a nucleus,
- multiple short branches called the **dendrites**, the role of which is to receive the information,

• one (or two in some cases when there are no dendrites) long branch called the **axon**, which is the conducting fiber between different neurons,
• in some cases, a **myelin sheath** that surrounds the axon, which is a thick layer of a greasy substance (the myelin) that speeds up the transmission,
• axon terminals that transmit the nerve impulses across a **synapse**, which is a gap between an axon terminal and the receiving dendrites of another neuron.

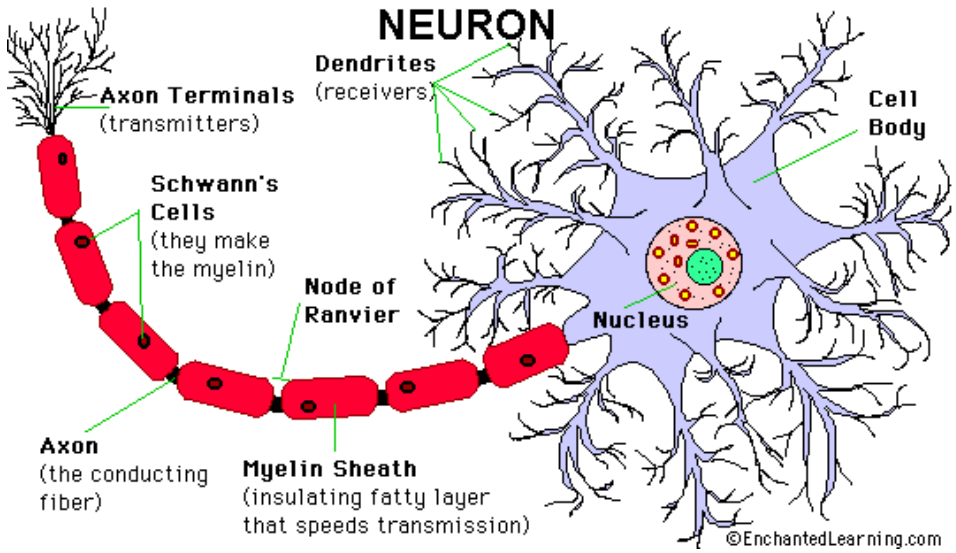


Figure 1.3: **Typical structure of a neuron.** (© EnchantedLearning.com)

Bundles of axons form the nerves, which are often termed nerve tracts or pathways within the CNS. The dendrites form a very large population of small fibers due to their numerous ramifications (up to 100000 per neuron). The dimension of the cell body varies from a diameter of 5 to 120μm whereas the axon is much smaller with a diameter that is generally between 0.5 and 3μm. However, axons that are surrounded with a myelin sheath can reach diameters up to 20μm [Bossy 1990]. Axons length has a very large range of variations from 1mm to 1m. The myelin sheath is produced by some glial cells: the Schwann's cells in the peripheral nervous system (i.e., the nervous system outside of the brain and the spinal cord) as illustrated by fig. 1.3 or the oligodendrocytes in the CNS. Nodes of Ranvier correspond to the periodic absences of myelin sheath along the axon.
We can distinguish three types of neurons:
• *sensory neurons* that ensure the transmission from sense receptors such as eyes or ears to the CNS (1% of all neurons),
• *motor neurons* the ensure the transmission from the CNS towards muscles (9% of all neurons),
• *interneurons cells* have no dendrites but two axons: one related to the spinal cord and the other to the muscle (90% of all neurons).

The glial cells. Their main function is to support, feed and protect the neurons. In particular, they are responsible for the phagocytosis of dead cells and

myelinogenesis. Their are three main types of glial cells in the CNS:
• The astroglia are the bulkiest cells. Their role is to support neurons: they can, to some extent, heal damaged nerve tissues as shown in fig. 1.5 and they create ramifications towards blood vessels to feed nerve cells as illustrated by fig. 1.4.
• The oligodendrocytes are satellite cells to the neuron. They are aligned between the nerve tracts and product the myelin (see fig. 1.4).
• The microglia are the smallest glial cells in the CNS, which allows them to circumnavigate in and between nerve tissues and thus to perform the phagocytosis of dead cells (see fig. 1.5).

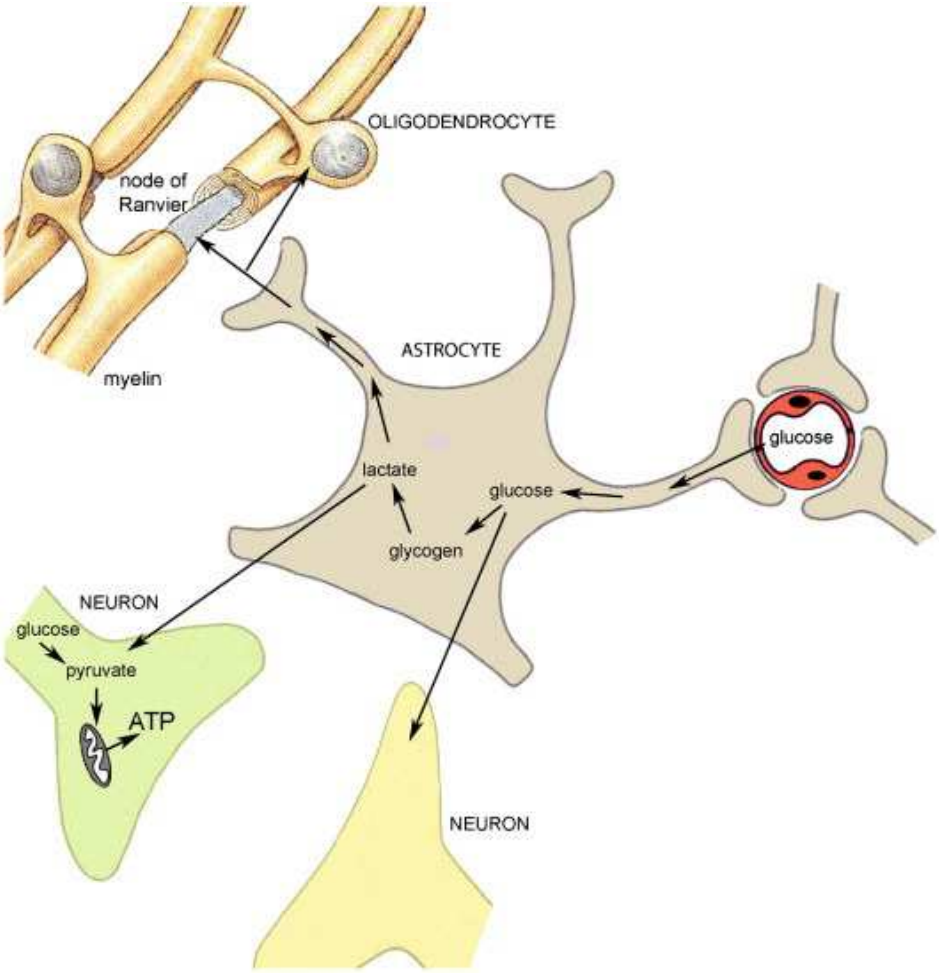


Figure 1.4: **Astroglia and Oligodendrocytes.** Illustration of the role of astroglia as energy suppliers to neurons and the role of oligodendrocytes as myelin producers (image courtesy of [De Keyser 2008]).

A recent study performed forebrain engraftment of human glial cells into neonatal immunodeficient mice and concluded that these mice learned faster and had a greater memory capability, which demonstrates that the elaborated cognitive capacities of the human brain are not only due to our complex neural network but also, and

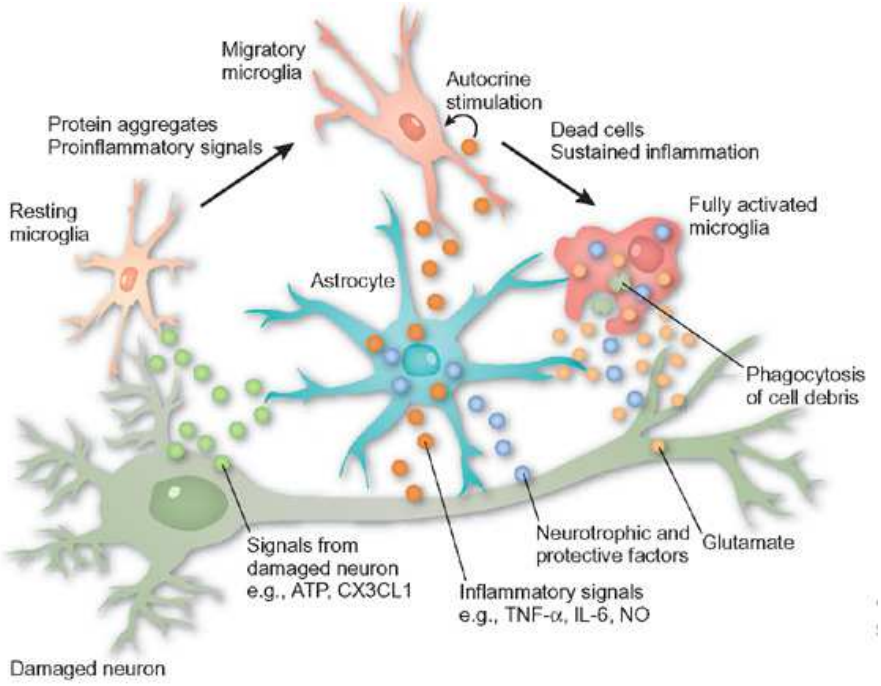


Figure 1.5: **Astroglia and Microglia.** Illustration of the protecting role of astroglia and the cleaning role of microglia (image courtesy of [Monk 2006]).

maybe mainly, to the huge number and diversity of our glial cells with respect to the other species [Han 2013].

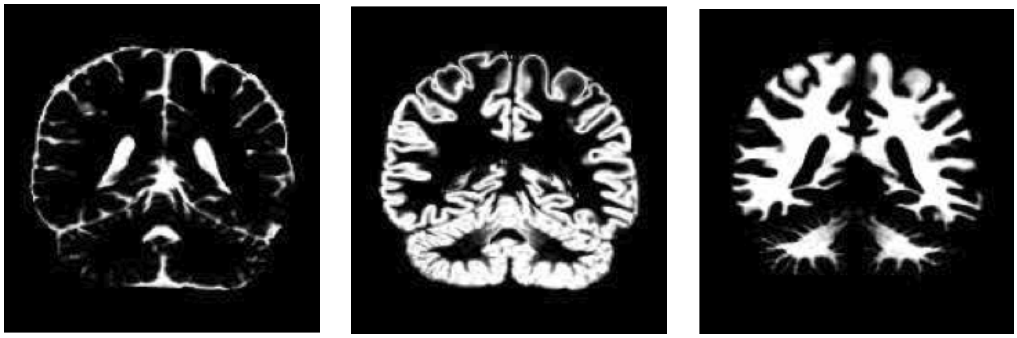
1.2.2 Macroscopic description

A practical way to describe the CNS from a macroscopic point of view can be found in histology, which studies the anatomy of cells and tissues by sectioning and staining. Macroscopic histology divides the CNS into three parts: (i) the CSF, (ii) the Gray Matter (GM) and (iii) the WM, which are illustrated by fig. 1.6.

The Cerebro-Spinal Fluid. The CNS is immersed in the CSF. This fluid navigates through communicative cavities called the **ventricles** (Figure 1.7). It contributes to absord traumatisms and also decreases the pressure in the brain. The CSF is producted by internal membranes of the ventricles, called the **choroid plexi**. There are 4 ventricles in the human brain: two lateral ventricles, a third ventricle under the previous two, related to them via the interventricular foramen and a fourth ventricle under the thrid one, related to it via the cerebral aqueduct also called Sylvius' aqueduct.

The Gray Matter. It essentially contains the body cells of the neurons with their dendrites, the unmyelinated axons and a part of all three types of glial cells. This particular composition confers the GM a grey color.

The White Matter. The main cells composing the WM are myelinated axons



(a) (b) (c)

Figure 1.6: **Cerebro-Spinal Fluid, Gray Matter and WM in the Human brain.** Illustration of the location of the CSF (a), the GM (b) and the WM (c) in the human brain on a coronal view (image courtesy of [Ciofolo 2009]).

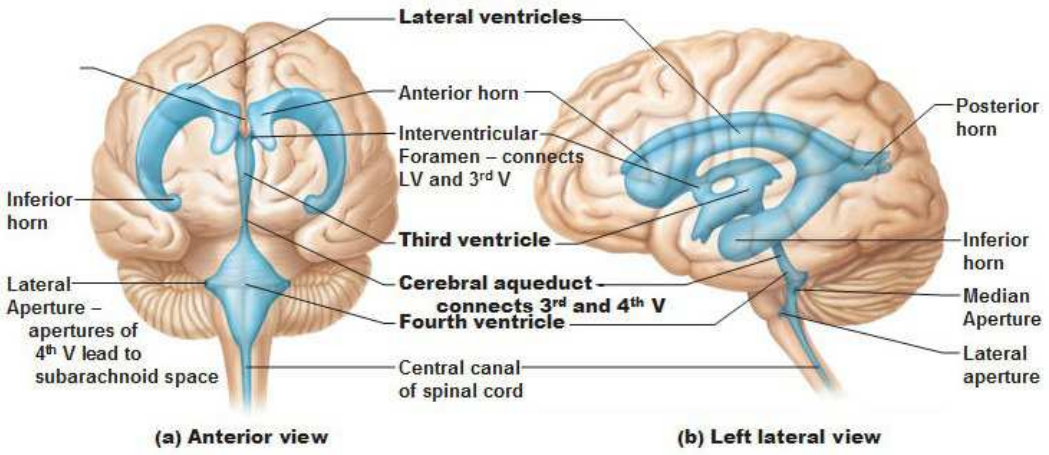


Figure 1.7: **Ventricles.** Schematic views of the position of the four ventricles in the human brain (© 2001 Benjamin Cummings, an imprint of Addison Wesley Longman, Inc.)

along with all three types of glial cells. The latter however have a slightly different function than in the GM. For instance, astroglia in the GM are coined *protoplasmic* astroglia and have little ramifications whereas astroglia in the WM are coined *fibrous* astroglia and have many ramifications that allow them to better protect and feed myelinated axons. Also, oligodendrocytes within the GM do not produce myelin as the GM axons are unmyelinated. The predominance in the WM of myelinated axons confers it a white color.

1.2.3 Brain neural transmission network

The brain neural transmission network resides in the WM. In this section, we aim at reviewing the normal anatomy of the WM nerve tracts, grouped into **fascicles**. The

following review is based on the works of [Catani 2002, Jellison 2004, Catani 2008, Thiebaut de Schotten 2011] that the reader wishing to delve deeper into the organization of the brain neural transmission network should refer to.

A global coronal view illustrating the currently known anatomic relationships between different fascicles is given in fig. 1.8. Circled fascicles correspond to the ones that the scientific community has been able to consistently identify. The other fascicles have only been occasionally identified but the current knowledge on the brain anatomy suggests their existence. The fascicles within the WM can be divided into three categories:

- the **association fascicles** that establish the connection between areas of the cerebral cortex (cortical areas) in each hemisphere of the brain,
- the **projection fascicles** that link the cortical areas to the deep nuclei (represented in fig. 1.8), brain stem, cerebellum and spinal cord (represented in fig. 1.2),
- the **commissural fascicles** that interconnect equivalent cortical areas between the two hemispheres.

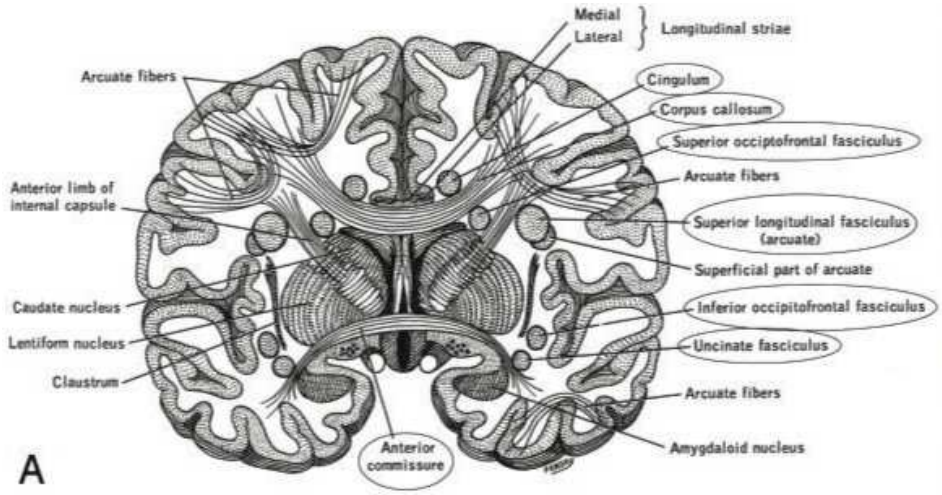


Figure 1.8: **Global view of the WM fascicles.** Illustrative coronal view of the relationship between several WM fascicles. Image courtesy of [Jellison 2004].

The association fascicles are illustated in fig. 1.9. They include the cingulum fascicle (fig. 1.9,a) that interconnects portions of the frontal, parietal and temporal lobes, the superior and inferior occipitofrontal fascicles (fig. 1.9,b-c) that interconnect the occipital and frontal lobes, the uncinate fascicle (fig. 1.9,b-c) that interconnects the frontal and temporal lobes, the superior and inferior arcuate fascicles (fig. 1.9,a-c) that connect the frontal lobe to the other three lobes.

The commissural fascicles are the Corpus Callosum (CC) fascicle illustrated in fig. 1.9(d), which is the largest of all fascicles and interconnects anterior and posterior poles of the hemispheres and the anterior commissure fascicle.

The projection fascicles are illustrated in fig. 1.10. They include the CST (fig. 1.10,a) that interconnects the motor cortex to the brain stem and spinal cord

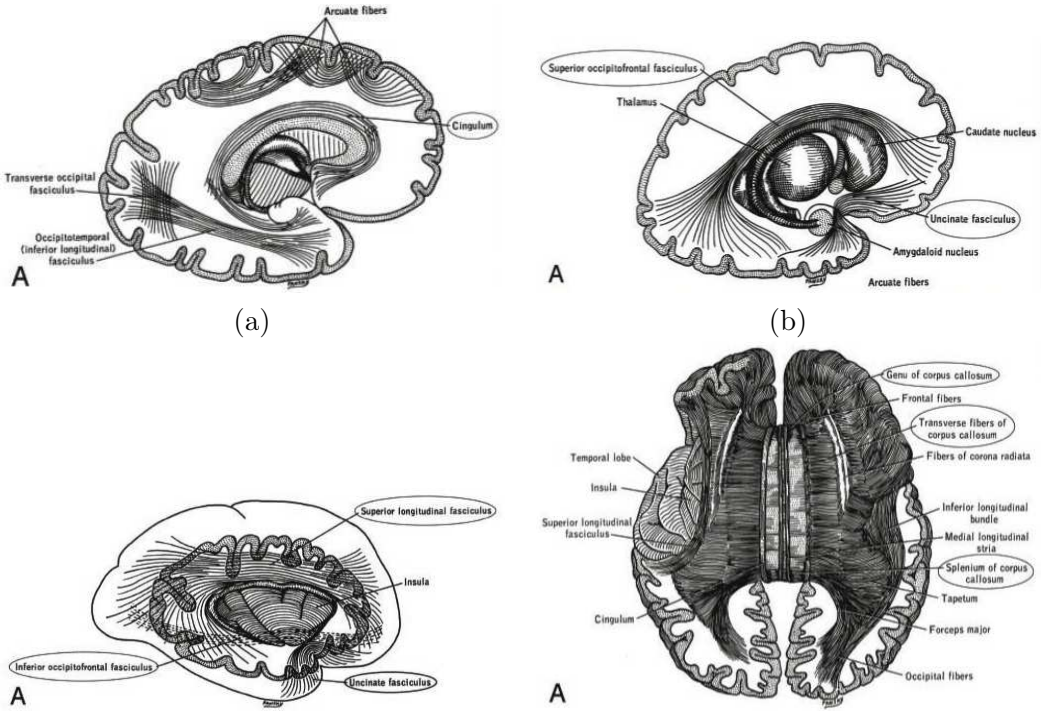


Figure 1.9: **Principal association and commissural fascicles.** Sagittal view of the cingulum association fascicle (a), the superior and inferior occipitofrontal association fascicles (b) and the uncinate and superior longitudinal association fascicles (c), axial view of the CC commissural fascicle (d). Image courtesy of [Jellison 2004].

(see fig. 1.2 for reference anatomy), the optic radiation fascicle (fig. 1.10,b) that connects the lateral geniculate nucleus to the occipital cortex, the internal capsule fascicle (fig. 1.10,c) that is a major access for in- and out-messages of the cerebral cortex and the corona radiata (fig. 1.10,d), which is not a fascicle but rather a structure in which fascicles from and to cortical areas get through.

This section was a brief introduction to the known anatomy of the brain as a part of the CNS. The reader should be convinced by now about the complexity of the architecture of the human brain. In particular, the WM that vehicules the information in and out the brain is composed of many different crossing fascicles that have not all been yet consistently identified. Also, the myelinated axons are not the only cells that constitute the WM. Since the work of [Moseley 1990] on dMRI applied to the analysis of the brain, many researchers focused on quantifying the microstruture of both WM and GM and on reconstructing the fascicles within the WM, which are invisible to the naked eye. The work presented in this manuscript goes in this direction too. In the next section, we will introduce the basics of dMRI.

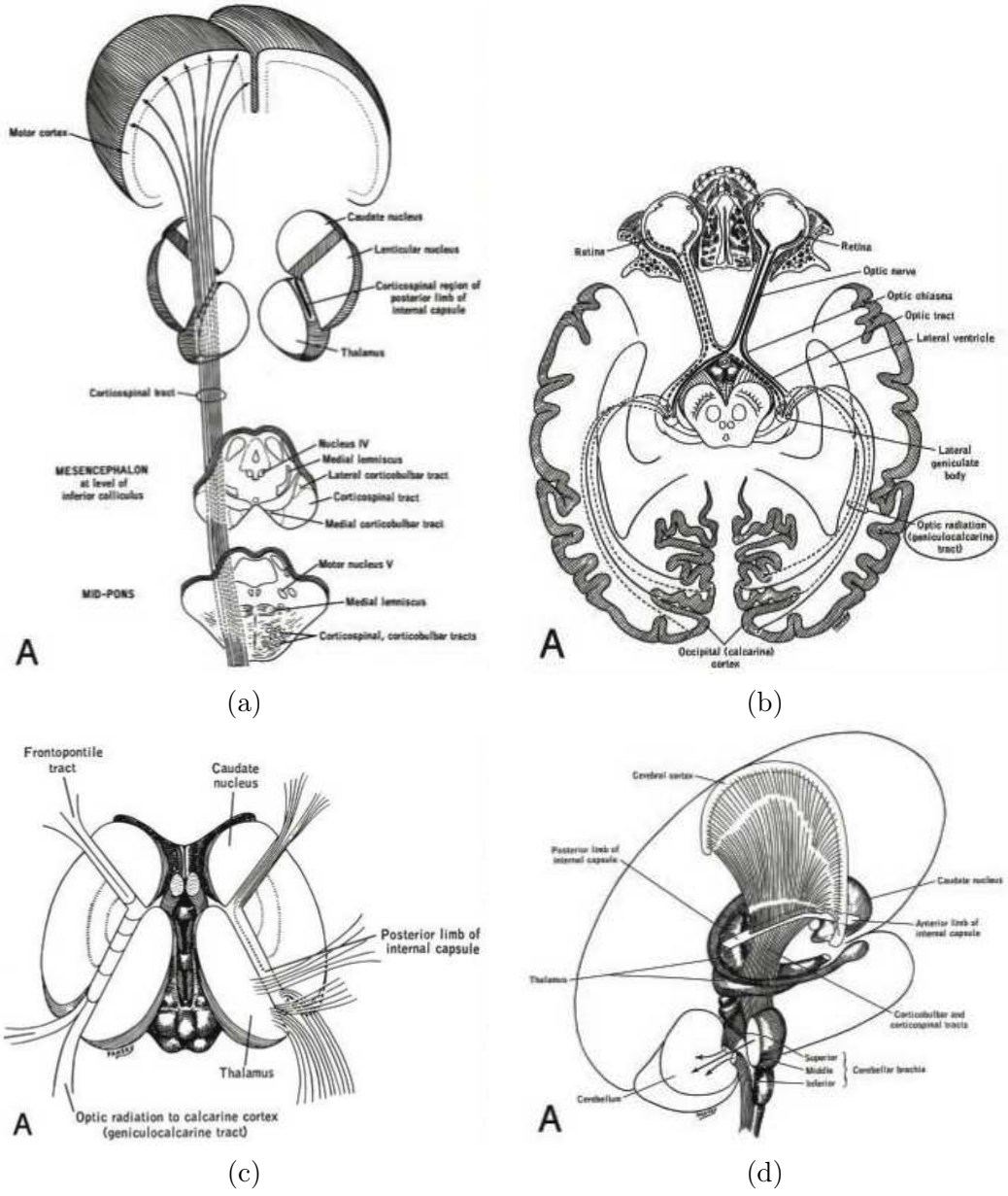


Figure 1.10: **Principal projection fascicles.** Illustration of the CST (a), axial view of the geniculocalcarine tract or optic radiation (b) and the internal capsule (c), medial view of the corona radiata (d). Image courtesy of [Jellison 2004].

1.3 Principles of Diffusion MRI

This section is dedicated to a simple introduction to dMRI and its application to measure diffusion of water in the human body. It is based on the works of [Basser 2009, Pipe 2009] that the reader can refer to for a deeper immersion into dMRI.

1.3.1 What is the diffusion?

Diffusion is a natural phenomenon of mass transportation that occurs in nature. The macroscopic idea behind the diffusion process is that, in a body that contains molecules that are distributed in several regions with different concentrations, a mass transport process will arise creating a *diffusive flux* that will tend to migrate the molecules from the most concentrated regions to the less concentrated regions as illustrated by fig. 1.11. In the physics community, this is known as Fick's first law [Fick 1855a, Fick 1855b]. It is an equation that links the diffusive flux $\mathbf{J}$ to any concentration difference $\mathbf{C}$ through

$$\mathbf{J} = -D\nabla\mathbf{C}\,, \tag{1.1}$$

where D is termed the *diffusion coefficient* and depends on (i) the size of diffusing molecules, (ii) the temperature in the medium in which the molecules are and (iii) the geometry of the medium. The sensitivity of the diffusion coefficient to the microstructure of the considered environment (cells in the medium and geometry of the medium) makes it very appealing for the study of the brain, which is full of water molecules subject to the diffusion process.

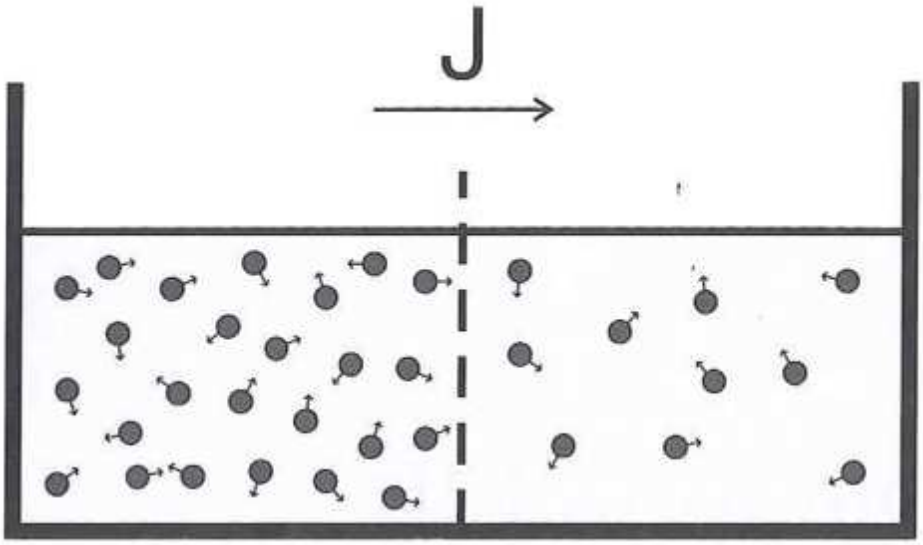


Figure 1.11: **Illustrative example of Fick's first law.** (image courtesy of [Basser 2009]).

The work of [Brown 1828] characterized the diffusion process at the microscopic level. It turns out that, contrary to what one may think, even in absence of a concentration difference, molecules can be subject to the diffusion process but the DIMD, in this case, will be invisible to the naked eye because, on average, there is indeed no net diffusive flux. Robert Brown observed at the microscope these displacements, successively on pollen grains and dust. The notion of a *diffusion displacement distribution* with zero mean (no net displacement on average) was

pioneered by [Einstein 1905] to quantify the probability for a given molecule to undergo a specific displacement.

1.3.2 How can it be measured using Magnetic Resonance Imaging?

MRI exploits the magnetization property of the hydrogen nuclei, also called the **protons**. Any proton exhibits a non zero magnetic dipole known as **spin**. Spins align themselves under the effect of an external magnetic field $\mathbf{B}_0$. The magnitude of a magnetic field is measured in tesla (T). First generation of clinical MRI scanners produced 0.5T (or even less) magnetic fields, current generation achieves 3T and, very recently, some MRI scanners for research purposes can generate 7 to 11T magnetic fields. To give an idea, the current clinical scanners achieve a magnetic field that is 60000 times stronger than the natural magnetic field that everyone experiences everyday without feeling it.

Spins under the effect of a magnetic field can undergo a phenomenon called **Larmor precession** if they are excited, for a brief period of time, by a Radio Frequency (RF) pulse with a frequency equal to the **Larmor frequency** of the spins given by

$$\omega = \gamma B_0 \,, \tag{1.2}$$

where γ is the gyromagnetic ratio, which is specific to the excited spin. In our case, the gyromagnetic ratio of the protons spin is $\gamma = 2.68 \times 10^8 \text{rad/s/T}$. Such an RF pulse is called **excitation RF pulse**. The excited spins are suddenly tilted in the plane of normal $\mathbf{B}_0$ where they start precessing (i.e., rotating around the axis defined by $\mathbf{B}_0$) at the Larmor frequency. Precessing spins generate a net magnetization that is measured as a current by receiver coils of the MRI scanner. The magnetization, and thus the measured signal, will be stronger if spins are in phase (i.e., all aligned together even if not with the magnetic field). Phase shifts will decrease the magnetization. Once the brief excitation RF pulse stops, the spins will tend to progressively realign themselves with the magnetic field $\mathbf{B}_0$. During the process, the magnetization in normal plane of $\mathbf{B}_0$, termed the **transverse magnetization**, decreases exponentially with a time constant T_2 called the **spin-spin relaxation time** ($\approx$ 100ms, depending on the tissue in which the molecules are).

However, the magnetic field $\mathbf{B}_0$ is not strictly homogeneous due to various factors [Abragam 1961]. The excitation RF pulse being computed according to eq. (1.2) thus results in excited spins at different precessing frequencies which provokes a loss in phase coherence and therefore, a faster decay of the transverse magnetization. In order to bypass the problems induced by inhomogeneities of the magnetic field, [Hahn 1950] proposed to apply, at some time t after the excitation RF pulse, a second RF pulse, termed **refocusing RF pulse**, to eliminate spin dephasing by flipping the spins in the plane of normal $-\mathbf{B}_0$ where they will precess the other way around. At time $2t$ precisely, the so-called **spin echo** will occur, that is spins will have regain total phase coherence. With the spin echo principle, the magnetic field $\mathbf{B}_0$ can then be considered as spatially homogeneous.

This is interesting as now one can volontarily make the magnetic field spatially varying by applying **magnetic field spatial gradients G** that create linear changes in the magnetic field as follows:

$$B(x, y, z) = B_0 + G_x x + G_y y + G_z z \,. \tag{1.3}$$

This is the basis of slice selection, image encoding and diffusion weighting. For example, let say that one aims at exciting the spins that are located at particular slice z of the brain. The application of a magnetic field spatial gradient in the z-direction modifies the magnetic field to $B(z) = B_0 + G_z z$. Spins of protons located in the plane (x, y) at slice z will then be the only ones excited by an excitation RF pulse of frequency given by eq. (1.2): $\omega(z) = \gamma(B_0 + G_z z)$. The prodecure for image encoding is not detailed here and we refer the reader to [Pipe 2009] for further details.

The spin-echo sequence was modified by [Carr 1954] who applied a DSG after the excitation RF pulse to make it sensitive only to the DIMD: this was the origin of DW MRI or dMRI. Consequently, spins of protons that were not subject to diffusion motion will have precessed at the same frequency all along the sequence (fig. 1.12,a), whereas protons that underwent diffusion motion will have changed location and thus their spins will have precessed at different frequencies, resulting, at the end of the sequence, in a persistent phase incoherence (fig. 1.12,b) and thus a decay of the measured signal which is only due to the DIMD.

Despite the great promises of this approach, it lacked clinical feasibility. Indeed, keeping a DSG active during the entire sequence is impractical. A feasible sequence was thus needed to fully exploit the spin-echo principle.

1.3.3 Pulse Gradient Spin Echo sequence

The Pulse Gradient Spin Echo (PGSE) sequence was exactly designed to make the most out of the spin-echo principle described in the previous section. It is due to [Stejskal 1965b]. They replaced the application of a single DSG during the entire sequence by the successive application of two DSGs of short pulse duration δ, with equal magnitude G and direction $\boldsymbol{g}$ and separated by a time Δ, called the **diffusion time**. One DSG is applied immediately after the excitation RF pulse and the other one immediately after the subsequent refocusing RF pulse. The sequence is illustrated in fig. 1.13. The time to reach the spin echo is called the Echo Time (TE). The time between two successive excitation RF pulses is called the Replication Time (TR).

Let then consider a molecule that is at location $\mathbf{x}_1$ during and after the application of the first DSG and at location $\mathbf{x}_2$ during and after the application of the second DSG. The magnitude of the magnetic field is thus $B_1 = B_0 + G < \boldsymbol{g}, \mathbf{x}_1 >$ at location $\mathbf{x}_1$, which makes spins at $\mathbf{x}_1$ precess at the frequency $\omega(\mathbf{x}_1) = \gamma B_1$, and $B_2 = B_0 + G < \boldsymbol{g}, \mathbf{x}_2 >$ at location $\mathbf{x}_2$, which makes spins at $\mathbf{x}_2$ precess at the frequency $\omega(\mathbf{x}_2) = \gamma B_2$. Note that $< \cdot, \cdot >$ denotes the inner product in $\mathbb{R}^3$.

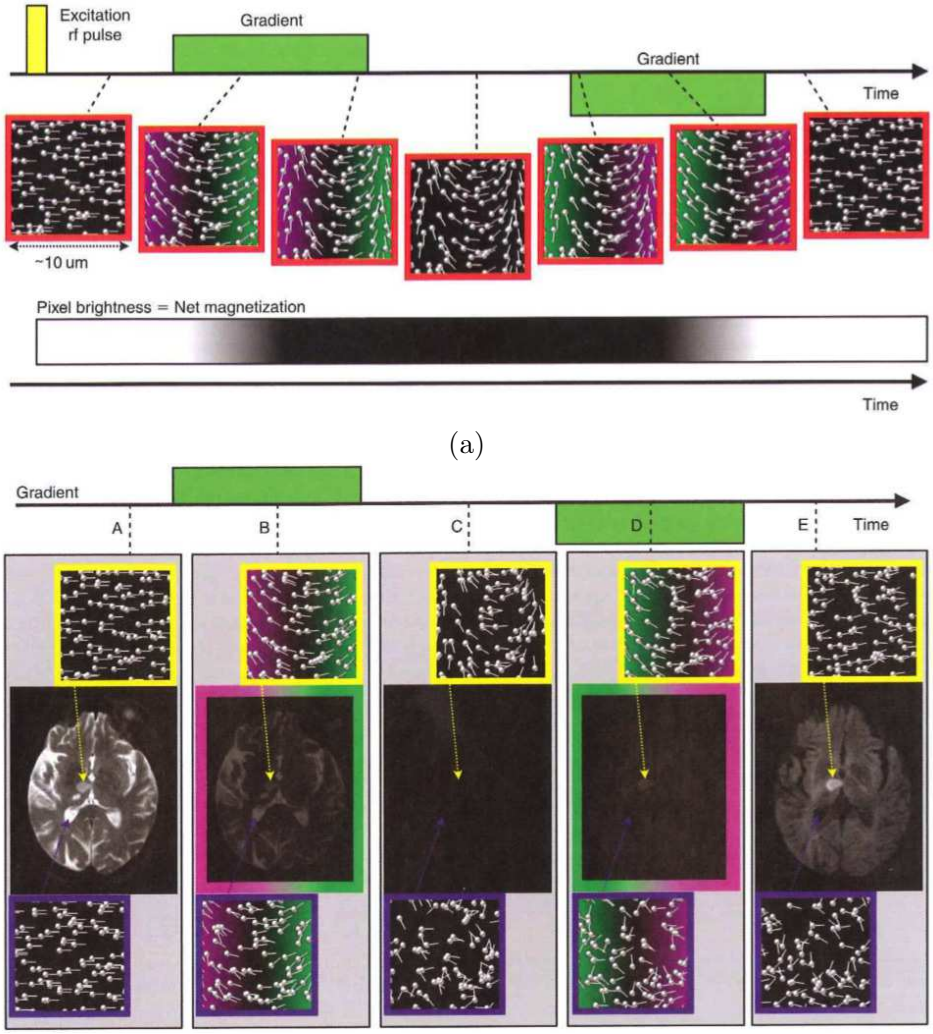


Figure 1.12: **Spin phase changes during the spin-echo sequence.** Top row image (a) represents the evolution of spin phase changes for molecules that do not undergo diffusive motion. The excitation pulse aligns all spins (first image). When the first DSG ends, the spins precess at the same frequencies but with different phases (middle image). When the second DSG ends, spins are in phase (last image). Bottom row image (b) represents the evolution of spin phase changes for diffusing molecules. The excitation pulse aligns all spins (first image). When the first DSG ends, the spins precess at different frequencies because of the location changes and thus have different phases (middle image). When the second DSG ends, most of the dephasing is removed but spins still are dephased due to the DIMD (last image). Image courtesy of [Pipe 2009].

Assuming that the pulse duration δ of the DSGs is negligible with respect to the diffusion time Δ and that the magnitude of the DSGs is constant over its application

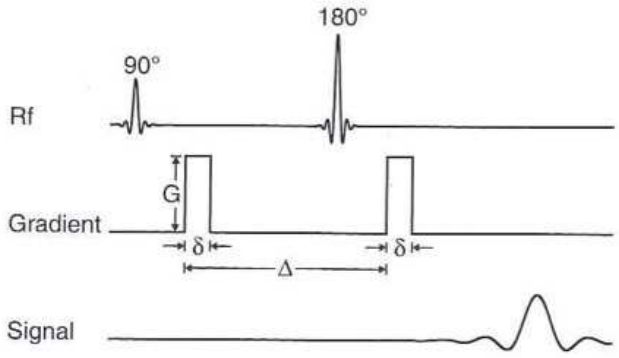


Figure 1.13: **Illustration of the PGSE sequence.** (Image courtesy of [Basser 2009])

pulse, which is known as the **narrow pulse regime**, the phase change of the molecule spins induced by the application of the first (respectively, second) DSG is then given by $\phi_1 = -\omega(\mathbf{x}_1)\delta$ (respectively, $\phi_2 = -\omega(\mathbf{x}_2)\delta$). The remaining phase shift after the application of the two DSGs can thus be quantified and amounts to:

$$\phi_2 - \phi_1 = -\gamma\delta G < \boldsymbol{g}, \mathbf{x}_2 - \mathbf{x}_1 > . \tag{1.4}$$

Equation (1.4) implies that the remaining phase shift is proportional to the DIMD that occurred between the application of the first and second DSG. In particular, if the molecule did not undergo diffusive motion, the phase shift is null and the measured magnetization does not experience a decay. This is the fondamental relation that explains how diffusion can be measured using dMRI.

1.3.4 Relation between the magnetization and the random diffusion process

Even if the spin echo principle compensates for the spatial inhomogeneities of the static magnetic field $\mathbf{B}_0$ generated by the MRI scanner, the exponential decay of the transverse magnetization with time constant T_2 (spin-spin relaxation time) still occurs and affects the magnetization in the same manner, whether or not DSGs are applied. Let M_0 be the magnetization immediately after the application of the excitation RF pulse. In absence of DSG, the magnetization $S_0^\star$ that we measure at time TE does not match M_0, but is given by:

$$S_0^\star = M_0 \exp\left\{-\frac{\mathrm{TE}}{T_2}\right\} + \text{ measurement noise}\,. \tag{1.5}$$

The upperscript * indicates that corresponding quantities are complex. The measured unweighted signal S_0 is given by $S_0 = |S_0^\star|$ and the true unweighted signal Ψ_0 in absence of noise is assumed to be the average of all the measured unweighted signals.

When a DSG $\boldsymbol{q} = \gamma\delta G\boldsymbol{g}$ [Callaghan 1991] of magnitude G and direction $\boldsymbol{g}$ is applied during a period δ to measure the DIMD, a way to bypass the influence

of the spin-spin relaxation time on the magnetization is to focus on the diffusion-induced magnetization decay $A^\star = A^\star(\boldsymbol{q})$ rather than the magnetization $S^\star = S^\star(\boldsymbol{q})$ itself, defining $A^\star = \frac{S^\star}{S_0^\star}$. However, eq. (1.5) highlights the importance of choosing the lowest possible TE for the PGSE sequence to ensure that the magnetization decay will mainly be due to diffusive motion and not to T_2 relaxation [Scherrer 2012].

From now on, we will refer to random quantities with a sanserif type font and to realizations of random quantities or deterministic quantities with curvilinear type font. Following [Einstein 1905], the DIMD are seen as not deterministic but random and one can define a Probability Density Function (PDF) $p_{\mathbf{x}_2|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_2;\Delta)$, which gives the probability that a given proton at location $\boldsymbol{x}_1$ will undergo a diffusive motion towards location $\boldsymbol{x_2}$ after a diffusion time Δ. The diffusion-induced magnetization decay $A^\star$ in a given voxel (3-dimensional pixel) is then given by:

$$A^\star(\boldsymbol{q},\Delta) = \int_V p_{\mathbf{x}_1}(\boldsymbol{x}_1)\left(\int_{\mathbb{R}^3} p_{\mathbf{x}_2|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_2;\Delta)\exp\{-i<\boldsymbol{q},\boldsymbol{x}_2-\boldsymbol{x}_1>\}\,\mathrm{d}\boldsymbol{x}_2\right)\mathrm{d}\boldsymbol{x}_1\,, \tag{1.6}$$

where $<\cdot,\cdot>$ denotes the inner product in $\mathbb{R}^3$, V is the volume of the voxel and $p_{\mathbf{x}_1}(\boldsymbol{x}_1)$ is the PDF that gives the probability for a proton in the voxel to be located at $\boldsymbol{x}_1$.

Equation (1.6) can be simplified by introducing the Ensemble Average Propagator (EAP) in the voxel, given by:

$$p_{\mathbf{x}}(\boldsymbol{x};\Delta) = \int_V p_{\mathbf{x}_1}(\boldsymbol{x}_1)p_{\mathbf{x}_1+\mathbf{x}|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_1+\boldsymbol{x};\Delta)\mathrm{d}\boldsymbol{x}_1\,, \tag{1.7}$$

where $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$ is the random vector representing the DIMD. The complex diffusion-induced magnetization decay then reads:

$$A^\star(\boldsymbol{q},\Delta) = \int_{\mathbb{R}^3} p_{\mathbf{x}}(\boldsymbol{x};\Delta)\exp\{-i<\boldsymbol{q},\boldsymbol{x}>\}\,\mathrm{d}\boldsymbol{x}\,. \tag{1.8}$$

Often in dMRI, the focus is on the magnitude of the complex diffusion-induced magnetization decay, hereafter referred to as the SID, which reads:

$$A(\boldsymbol{q},\Delta) = |A^\star(\boldsymbol{q},\Delta)| = \left|\int_{\mathbb{R}^3} p_{\mathbf{x}}(\boldsymbol{x};\Delta)\exp\{-i<\boldsymbol{q},\boldsymbol{x}>\}\,\mathrm{d}\boldsymbol{x}\right|\,. \tag{1.9}$$

The integral involved in eq. (1.9) is, by definition, the characteristic function $\varphi_{\mathbf{x}}$ of the distribution of the DIMD $\mathbf{x}$. The SID thus simply reads:

$$A(\boldsymbol{q},\Delta) = |\varphi_{\mathbf{x}}(\boldsymbol{q};\Delta)|\,. \tag{1.10}$$

The magnitude of the magnetization $S^\star(\boldsymbol{q},\Delta)$ is referred to as the measured DW signal and denoted $S(\boldsymbol{q},\Delta)$. The predicted DW signal under a specific diffusion model or in absence of measurement noise is denoted $\Psi(\boldsymbol{q},\Delta)$ and amounts to $\Psi(\boldsymbol{q},\Delta) = \Psi_0 A(\boldsymbol{q},\Delta)$, where $A(\boldsymbol{q},\Delta)$ is given by eq. (1.10) and Ψ_0 is the true unweighted signal.

The problem of mathematically defining the relationship between the magnetization and the random diffusion process was tackled differently in [Torrey 1956]. In this work, Bloch equations that describe the behaviour of the magnetization under relaxation phenomena [Bloch 1946] are generalized to integrate the effects of diffusion. We will not detail these equations here.

1.3.5 Summary

We have introduced the principles of dMRI. In section 1.3.1, we described the phenomenon of diffusion highlighting its great sensitivity for the study of microstructural environments in which it occurs. We illustrated in sections 1.3.2 and 1.3.3 how dMRI can be used to measure the magnetization of proton spins reflecting the underlying diffusive motions and we mathematically derived, in section 1.3.4, the relationship between this measured magnetization and the EAP that defines the random distribution of the DIMD.

The key relation is given by eq. (1.10), which shows that the magnitude of the magnetization is the magnitude of the Fourier transform of the EAP. In the next section, we will explain how the EAP is actually computed from the magnitude of the magnetization using this relationship in the MRI scanner.

1.4 Diffusion Tensor Imaging

1.4.1 Description of the model

If there are no obstacles that may hindered the diffusion, then we say that the diffusion is **free**. In this case, the EAP is a zero-mean isotropic Gaussian PDF whose covariance matrix is given by $2D\Delta I_3$ where D is the diffusivity coefficient in the voxel and I_3 is the 3×3 identity matrix. The expression of the covariance matrix derives from [Einstein 1905] who proved that the mean-squared displacement of the ensemble is $2D\Delta$. With the PGSE sequence, [Stejskal 1965b] indicate that this definition needs a correction to account for the duration δ of the DSGs: they show that the correct covariance matrix is $2D(\Delta - \delta/3)I_3$. Assuming such a diffusion profile leads to a very elegant simplification of eq. (1.10):

$$A(G, \delta, \Delta) = \exp\left\{-\gamma^2\delta^2G^2D(\Delta - \delta/3)\right\} . \tag{1.11}$$

In order to simplify eq. (1.11), the concept of b-value has been introduced in [Le Bihan 1991]. For a given DSG $\boldsymbol{q} = \gamma\delta G\boldsymbol{g}$ in a PGSE sequence in which it is applied twice with pulse duration δ and diffusion time Δ between its two applications, the b-value is defined as follows:

$$b = \gamma^2\delta^2G^2(\Delta - \delta/3) . \tag{1.12}$$

The b-value summarizes in one value the contribution of the DSG strength G, the DSG duration δ and the diffusion time Δ. Equation (1.11) then reads:

$$A(b) = \exp\{-bD\} . \tag{1.13}$$

In particular, the diffusivity of freely diffusing water at the normal human body temperature 37°C has been shown to be approximately $D = 3.0{\times}10^{-3}\text{mm}^2/\text{s}$ [Le Bihan 2006].

This description of the diffusion process is however only valid for free diffusion in isotropic media. The work of [Stejskal 1965a] extends eq. (1.13) to account for free anistropic diffusion and is the object of this section. A simple way to account for

anisotropic diffusion is to generalize the EAP to a zero-mean 3-dimensional Gaussian PDF with covariance matrix $\Sigma = 2(\Delta - \delta/3)D$, where D is not a scalar anymore but a DT, i.e. a 3×3 symmetric definite positive matrix. In other words, the EAP is assume to take the following form:

$$p_{\mathbf{x}}(\boldsymbol{x}; \Delta, \delta) = \left((4\pi(\Delta - \delta/3))^3 \det D\right)^{-1/2} \exp\left\{-\frac{< \boldsymbol{x}, D^{-1}\boldsymbol{x} >}{4(\Delta - \delta/3)}\right\},$$

where $< \cdot, \cdot >$ denotes the inner product in $\mathbb{R}^3$. Without loss of generality, the DIMD $\mathbf{x}$ are often rescaled and confounded with the random variable $\mathbf{y}$ defined as:

$$\mathbf{y} = (2(\Delta - \delta/3))^{-1/2}\,\mathbf{x}\,, \tag{1.14}$$

such that the covariance matrix of the distribution of $\mathbf{y}$ matches the DT D. Indeed, the PDF of the distribution of the random variable $\mathbf{y}$ amounts to:

$$p_{\mathbf{y}}(\boldsymbol{y}) = \left((2\pi)^3 \det D\right)^{-1/2} \exp\left\{-\frac{1}{2} < \boldsymbol{x}, D^{-1}\boldsymbol{x} >\right\}. \tag{1.15}$$

Equation (1.10) then reads:

$$A(\boldsymbol{q}, \Delta) = |\varphi_{\mathbf{x}}(\boldsymbol{q}; \Delta)| = \left|\varphi_{\mathbf{y}}\left(\sqrt{2(\Delta - \delta/3)}\boldsymbol{q}\right)\right| = \left|\varphi_{\mathbf{y}}\left(\sqrt{2\gamma^2\delta^2G^2(\Delta - \delta/3)}\boldsymbol{g}\right)\right|,$$

which, using eq. (1.12), simplifies to:

$$A(b, \boldsymbol{g}) = \left|\varphi_{\mathbf{y}}(\sqrt{2b}\boldsymbol{g})\right|. \tag{1.16}$$

Modeling the random variable $\mathbf{y}$ in eq. (1.16) with eq. (1.15) yield the Stejskal & Tanner equation [Stejskal 1965b]:

$$A(b, \boldsymbol{g}) = \exp\{-b < \boldsymbol{g}, D\boldsymbol{g} >\}. \tag{1.17}$$

This modeling of the diffusion is known as Diffusion Tensor Imaging (DTI) and is able to describe the diffusion in anisotropic media, through the eigensystem $\{\boldsymbol{e}_i, d_i\}_{i=1,\dots,3}$ of the DT ($d_1 \geq d_2 \geq d_3 > 0$). The orientation of diffusion is given by $\pm\boldsymbol{\mu} = \pm\boldsymbol{e}_1$ and one can compute the Apparent Diffusion Coefficient (ADC) $D(\boldsymbol{c})$ in any given orientation $\pm\boldsymbol{c}$, $\|\boldsymbol{c}\| = 1$ as $D(\boldsymbol{c}) = \boldsymbol{c}'D\boldsymbol{c}$. Many diffusivity-based scalar maps can be then produced from the DT, as, for instance, the ADC parallel to the orientation of diffusion, termed *parallel diffusivity* $D_{/\!/}$ and the mean ADC over the plane orthogonal to the orientation of diffusion, termed *perpendicular diffusivity* $D_\perp$:

$$D_{/\!/} = d_1 \text{ and } D_\perp = (d_2 + d_3)/2\,, \tag{1.18}$$

or the mean ADC over the unit 2-sphere, termed the Mean Diffusivity (MD):

$$\text{MD} = \frac{\text{tr}\,(D)}{3}. \tag{1.19}$$

Anisotropy measures can also be derived from the DT. The two most popular rotationally invariant anisotropy indices were introduced in [Basser 1996a], namely the FA and the Relative Anisotropy (RA), and are defined as:

$$\text{FA} = \sqrt{\frac{3}{2} \cdot \frac{\text{tr}\left((D - \text{MD} \cdot I_3)^2\right)}{\text{tr}\left(D^2\right)}} \text{ and } \text{RA} = \sqrt{\frac{1}{3} \cdot \frac{\text{tr}\left((D - \text{MD} \cdot I_3)^2\right)}{\left(\text{tr}\left(D\right)\right)^2}}. \tag{1.20}$$

The same authors also proposed the Volume Ratio (VR) and Lattice Index (LI), which are less commonly used [Pierpaoli 1996a]. Finally, shape indices have been proposed to distinguish between *prolate* (cigar-shaped), *oblate* (Smarties-shaped) and *spherical* (ball-shaped) tensors. In details, they are respectively termed linear coefficient (c_ℓ), planar coefficient (c_p) and spherical coefficient (c_s) and are defined as [Westin 2002]:

$$c_\ell = \frac{d_1 - d_2}{\text{tr}\left(D\right)}, \; c_p = \frac{2(d_2 - d_3)}{\text{tr}\left(D\right)} \text{ and } c_s = \frac{3d_3}{\text{tr}\left(D\right)}. \tag{1.21}$$

Note that, sometimes, these three coefficients are normalized by $\sqrt{\text{tr}\left(D^2\right)}$ instead of $\text{tr}\left(D\right)$.

Assuming that this description of the diffusion process is an appropriate diffusion model, the problem of designing an optimal dMRI sequence to accurately estimate the DT is needed and will be the object of the next section.

1.4.2 Optimal DT-MRI acquisition strategy and estimation framework

1.4.2.1 Acquisition strategy

A dMRI sequence is defined by several parameters: its TR and TE and parameters that define the DSGs, namely the b-value given by eq. (1.12) and the direction $\boldsymbol{g}$ of the DSG. The TR determines, together with the number of applied DSGs, the duration of the entire dMRI examination that should be under the 10 minutes for enabling its use in routine clinical practice. The sampling of the DSGs is thus very important. Two sampling strategies have been mainly used in the literature: **cartesian** sampling and **spherical** sampling. Cartesian sampling is generally referred to as Diffusion Spectrum Imaging (DSI) [Wedeen 2000, Wedeen 2005], which is an imaging technique designed to numerically compute the inverse Fourier transform of eq. (1.10) to obtain the EAP but requires a number of measurements in the range 200 - 512, which makes it impractical in the clinical routine practice. A popular sampling strategy for the DSGs has thus been adopted by the community, namely spherical sampling. It consists in radially sampling the DSGs on a few different spheres called the **shells** using a few number of different b-values and in sampling the directions of the DSGs on these shells. Single b-value sequences are often called **single-shell acquisitions** while multiple b-value sequences are called **multi-shell acquisitions**.

Since the TE is the time of spin echo and the TR is the time between two excitation RF pulses, the greater the TE the greater the TR. Consequently, larger TEs will lead to longer acquisitions. We detailed in section 1.3.4, and particularly with eq. (1.5), that the TE needs to be as small as possible to minimize confounding effects from T_2 relaxation of the tissues. Also, from the general definition of the b-value given by eq. (1.12), we can define the **nominal b-value**, i.e. the b-value for a unit-norm DSG:

$$b_{\text{nominal}} = \gamma^2 \delta^2 (\Delta - \delta/3) \,. \tag{1.22}$$

It has been shown that the minimum achievable TE, which leads to the shortest acquisition and to negligible T_2 relaxation effects, is linked to this nominal b-value via a complex relationship that can be approximated by [Mattiello 1994, Jones 2004]:

$$\text{TE}_{\text{min}} \approx \left(\frac{12 b_{\text{nominal}}}{\gamma^2} \right)^{1/3} . \tag{1.23}$$

In a spherical shell sampling strategy for the DSGs, we shall thus choose the sequence parameters so that the highest nominal b-value is kept relatively small.

Assuming that the diffusion process is well represented by the DT and adopting a spherical sampling of the DSGs for designing the appropriate acquisition scheme, we still need to decide how many DSGs directions and how many shells (b-values) are needed to accurately estimate the DT. Estimating the DT requires the estimation of six independent parameters since it is a 3×3 symmetric definite positive matrix. Theoretically, since the SID measured by the MRI scanner is orientationally dependent, if the measurements were not corrupted by noise, one could then use only six non-collinear DSG directions and a single b-value. Unfortunately, the measured SIDs do not match the theoretical one due to the presence of measurement noise [Gudbjartsson 1995] and thus [Basser 1994a] proposed a first acquisition scheme that involves many more non-collinear DSG directions with a single b-value. This is known as DT-MRI and this is the current way to analyse the diffusion in clinics thanks to the patent filed by [Basser 1996b]. Many studies have been carried out since the work of [Basser 1994a, Basser 1994b] to try to determine the optimal number of non-collinear DSGs directions and of b-values to obtain the *best* estimate of the DT [Correia 2009].

Optimal number of non-collinear DSG encoding directions. The dMRI community first focused on understanding how many DSG directions were needed. The work of [Papadakis 2000] pioneered in this direction. They concluded that, using a single nominal b-value of 1570s/mm^2, the optimal number of DSGs directions is between 18 and 21 to obtain the lowest variance on the estimate of the anisotropy indices FA, RA and VR. In 2004, a study performed by [Jones 2004] confirmed the conclusions of Papadakis: a robust estimation of anisotropy measures requires at least 20 non-collinear DSG directions and a robust estimation of the DT orientation requires at least 30 non-collinear DSG directions.

Optimal number of encoding b-values. At the infancy of DT-MRI, [Papadakis 1999] evaluated the total variance of the DT estimated by (i) repeating twice a dMRI

sequence with a single b-value at 1570s/mm^2 or (ii) using 2 distinct b-values of 1570s/mm^2 and 785s/mm^2 (half the first one). The comparison was performed for two sets of 6 DSG encoding directions differently distributed over the north hemisphere. Conclusions favored the multi-shell acquisition. However, a recent study [Jones 2007] just concluded the contrary. He showed that any scalar maps derived from the DT present a lower bias and a better precision in presence of noise when a single-shell acquisition is repeated M times rather than when a multi-shell acquisition with M different shells is used. Ten years after the work of [Papadakis 1999], [Correia 2009] published a study that is in accordance with Papadakis' findings: increasing the number of b-values seems to decrease the bias in ADC and FA estimates and allows a better discrimination of fascicles with different diffusivities by homogenizing their estimated FAs.

Using non-unit-norm DSGs to get more b-values. Tetrahedrally arranged $\sqrt{3}$-norm DSGs have been used within a single-shell acquisition at a given nominal b-value [Conturo 1996]. They showed that this strategy enables the estimation of the MD given by eq. (1.19) using only 4 DSG directions. The same year, [Pierpaoli 1996b] proposed to use hexahedrally arranged $\sqrt{2}$-norm DSGs instead and perform an accurate estimation of the DT using only 7 DSG directions. These two types of DSG arrangements served recently as the foundations for CUbe Rays to Vertices and Edges (CURVE)-Ball DTI proposed by [Peled 2009] that enables, from a single nominal b-value, to image the signal on three different shells at b_{nominal}, $2b_{\text{nominal}}$ and $3b_{\text{nominal}}$, while keeping the TE reasonably low since its minimum achievable value is proportional to $b_{\text{nominal}}^{1/3}$ form eq. (1.23). Using a single nominal b-value around 1000s/mm^2, [Hasan 2001, Batchelor 2002] showed that six icosahedrally arranged DSGs, which have $\sqrt{5}$-norm and thus result in a b-value at around 5000s/mm^2 without increasing the duration of the acquisition or the effects of T_2 relaxation, are optimal for the estimation of the DT. Recently, [Scherrer 2012] introduced and showed the great promises of the CUbe and SPhere (CUSP) acquisition scheme that makes use of the DSGs that are in-between a single shell at $b = 1000$s/mm^2 and its enclosing cube, achieving many different b-values in the range 1000-3000.

Conclusions regarding the benefits of using multi-shell acquisitions over single-shell acquisitions are rather contradictory in the literature. Even if the right range and/or number of b-values is not yet clear, recent literature tends to show that we should move towards hybrid acquisition schemes with a single nominal b-value but DSGs not necessarily distributed on this shell to gain more effective b-values. These strategies are however very recent and were not available at the time of Basser's patent and the integration of DTI in MRI scanners. Back then, however, it seemed that there was a consensus on using a single b-value and on the optimal number of non-collinear DSG encoding directions on this shell. The optimal sequence in clinics for DT-MRI therefore consists in a single encoding b-value usually in the range 800 – 1500s/mm^2 to avoid too low TEs and 30 non-collinear DSG encoding directions uniformly distributed on the north hemi-sphere. In a recent review of DTI computational methods, [Hasan 2010] report the parameters of a typical isotropic DTI protocol, which consists, on a 3T MRI scanner, in the application of 42 DSGs

in 42 non-collinear directions (which is slightly more than the optimal number of 30 proposed by [Jones 2004]) on a single spherical shell (typically $b = 1000\text{s/mm}^2$), with TR = 12s and TE = 65ms, achieving a full human brain coverage in 7 min with a spatial resolution of 2mm^3 and Signal-to-Noise Ratio (SNR) near 50.

1.4.2.2 Estimation framework

DT computation requires the estimation of only six independent parameters. One can see that the logarithm of the SID in eq. (1.17) is a linear function of the coefficients of the DT, which is very handy to encompass within a fast and robust estimation framework. From the original idea of DTI in [Basser 1994b], many researchers have been working on improving DT estimation. The advance in this particular topic are well summarized in [Fillard 2007] that proposes a complete log-Euclidean framework for the estimation of the DT. It consists in (i) reparametrizing the DT as the exponential of a matrix termed the *log-tensor*, on which usual operations (addition, subtraction and so on) can be employed at will without risking to lose the definite positivity of the original DT and (ii) minimizing a criterion that includes a data attachement term and an anisotropic regularization term. They reviewed three different data attachement terms, showing that one should build that term on the likelihood for Rician-corrupted data that corrects for the *shrinking effect* (underestimation of the true volume of the tensors). Very recently, [Liu 2013] introduced a robust variational framework that simultaneously performs the estimation and smoothing of the DTs.

1.4.3 Clinical applications of DT-MRI

Since most MRI scanner on the market are equipped with the analysis tools for DT-MRI [Basser 1996b], the literature carries an incredibly long list of publications on the use of DTI for the study of the human brain (mainly of its WM). DTI has indeed first been applied on healthy adult subjects (mean age 31 years old) [Pierpaoli 1996b] to gain knowledge about the values of the different scalar maps available from the eigenvalues of the estimated DT (section 1.4.1). After this study, DTI has been widely and successfully employed to study DT scalar metrics changes due to development (brain maturation, aging) and cerebral pathologies. In general, studies aimed at assessing changes in one diffusivity-based metric (often the MD given by eq. (1.19)) and one anisotropy metric (often the FA given by eq. (1.20)). In this section, we will give an insight into the wide range of successful clinical applications of DT-MRI to make the reader sensitive to (i) the importance of DTI to the clinicians and thus to why it is important that a candidate diffusion model for replacing DTI can reproduce all the usual DTI-derived scalar maps.

The dMRI community first focused on understanding the diffusivity and anisotropy patterns in the brain of adult healthy subjects through the work of [Pierpaoli 1996b]. In this founding paper, for each well-established fascicle of the WM (see section 1.2.3 for the list of these fascicles) as well as for the CSF and the GM, the authors have

reported the mean values and standard deviations of all the eigenvalues of the estimated DT. We refer the reader to the paper for a detailed knowledge on these values. The authors' findings reveal that (figure 3 of the paper):

- the WM is characterized by a high anisotropy and a low diffusivity,
- the GM is characterized by a low anisotropy and a low diffusivity,
- the CSF is characterized by a low anisotropy and a high diffusivity.

We report here only the summarizing table of this work (table 1.1) produced and published in [Le Bihan 2001], which gives the mean values and standard deviations of MD and of 1 – VR in some of the most important regions of the brain.

	MD ($\times 10^{-3}$mm^2/s)	1 – VR
CSF	3.19 ± 0.10	0.02 ± 0.01
GM (frontal cortex)	0.83 ± 0.05	0.08 ± 0.05
Caudate nucleus	0.67 ± 0.02	0.08 ± 0.03
WM		
CST	0.71 ± 0.04	0.93 ± 0.04
CC (splenium)	0.69 ± 0.05	0.86 ± 0.05
Internal capsule	0.64 ± 0.03	0.70 ± 0.08
Centrum semiovale	0.65 ± 0.02	0.27 ± 0.03

Table 1.1: **Mean value and standard deviation of MD and 1–VR for various regions of the brain.** Table reproduced from [Le Bihan 2001].

These values along with the mean values of the other scalar maps constitute the ground truth for the healthy adult brain. After this work, many publications arose on diffusivity and/or anisotropy changes in development and aging and in different brain pathologies by comparing a control group of healthy subjects with a group of patients matched in age and gender. In the following, we will list the different clinical applications of DT-MRI and summarize for each application the findings due to DTI. Cited publications are reviews of DTI findings on each subject. All individual studies are cited inside these works and we invite the interested reader to refer to these reviews for details on the individual analyses.

Developing brain, maturation and aging [Neil 2002, Sundgren 2004]. The study of changes in diffusivity and anisotropy measured via DTI in the development of normal and injured human brain revealed that, in normal brain development, diffusivity is higher in children than in adults, which is believed to be due to a decreasing total water content and anisotropy increases thanks to the myelination and re-organization of WM fascicles. In GM, modifications of the dendritic network seem to cause the variations of anisotropy. Brain injuries are shown to provoke a significant decrease of both diffusivity and anisotropy, the latter providing proof of WM alterations that none of the other MRI modalities had revealed.

Aging [Moseley 2002, Sundgren 2004]. Several studies evaluated the influence of aging using DTI. Conclusions were that the WM loses coherence from 20 years of age, which means that globally the anisotropy decreases with age. This phenomenon

is gender-independent. Possible given explanations are that axons progressively lose their myelin sheath with age and some axons may also disappear, which is corroborated by histological studies. Also, diffusivities are higher after 40 years of age. These findings must be kept in mind when studying brain pathologies with DTI, especially in the elderly.

Brain ischemia [Le Bihan 2001, Sotak 2002, Sundgren 2004]. Brain ischemia designates a decrease in cerebral blood flow that creates a flux from the extracellular space towards the intracellular space, resulting in a edema. There are two stages in brain ischemia: the acute phase in which the edema is cytotoxic (i.e., toxic to cells) and the chronic phase in which the edema is vasogenic (i.e., rapidly spreading edema). The MD decreases in the first phase (more severely in the WM) but increases to value higher than the normals in the second phase. During the first phase, there is also an acute elevation of FA in the WM only, which is followed in the second phase by a reduction. In the second phase, the architecture of the WM is indeed disrupted and loses coherence.

Multiple Sclerosis (MS) [Le Bihan 2001, Horsfield 2002, Sundgren 2004]. It is an inflammatory autoimmune disease that mainly affects the WM. There are several stages in MS. After the first symptoms (optic neuritis, pain in the limbs), a period of remission with occasional acute attacks starts (relapsing-remitting phase) and then either acute attacks accumulate leading to severe disability (secondary progessive phase) or acute attacks stabilize leading to limited disability. Lesions - alteration and/or destruction of the myelinated axons - are provoked in the WM during the attacks. The resulting MD has shown to be inflated and the resulting FA to be deflated. Increased diffusivity has also been observed away from the lesions in the WM, suggesting either that MS is a diffuse disease or that WM undergoes Wallerian degeneration.

Wallerian degeneration [Le Bihan 2001, Sundgren 2004]. It is a process of both axon and myelin degeneration that occurs when axons are injured or cells die after a cerebral infarct. Most often, it involves the CST. DTI reveals a reduction of FA both in the lesions provoked by the infarct and wherever the Wallerian degeneration started. Differently, the MD is almost not increased by Wallerian degeneration but significantly increased in the lesions, which makes MD a good candidate marker to distinguish lesions from areas that undergo Wallerian degeneration.

Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) [Le Bihan 2001, Horsfield 2002, Sundgren 2004, Chua 2008]. MCI is diagnosed when cognitive decline is greater than normal aging decline but with little functional impairment. It is considered as an early phase of AD, which is the most common form of dementia. AD is a neurogenerative disorder that mainly affects the GM and, to a less extent, the WM axons and oligodendrocytes. The early symptom is the loss of short-term memory and progressively confusion, irritability, trouble with language and loss of long-term memory. It provokes general brain atrophy (decreased volume). Earliest studies on MCI and AD mainly focused on the changes in anisotropy induced by the disease. It has been shown that the genu and splenium of the CC presented a lower FA and that the integrity measure via the LI was preserved in the CST but altered

in fascicles involved with cognitive functions. This decay in anisotropy is thought to be due to axonal loss and/or demyelination. More recent studies consistently identified an increase in MD and a decrease in FA in the centrum semiovale and in posterior cingulum fascicles. A very recent study [Douaud 2011] shows that in fact the only consistent difference between MCI and/or AD patients with healthy subjects is the FA reduction in the centrum semiovale. Unfortunately, they point out that this region is also a location of crossing fascicles with a decreased FA as already observed with DTI in healthy subjects.

Brain tumors [Sundgren 2004]. The MD has proven to be very successful in differentiating tumors, lesions, edema, necrotic and/or cystic regions from normal brain. There are generally three regions surrounding a tumor: the necrotic and/or cystic part, the edema and a region identified as contrast-enhancing. The formation of a tumor translates in an increase in MD and a decrease of FA, which is particularly marked in the necrotic and edema parts. These changes are generally attributed to increased water content and tumor infiltration.

CADASIL (Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leucoenuphalopathy) [Le Bihan 2001, Horsfield 2002]. It is an hereditary disease, which generates subcortical ischemic episodes and progressive dementia. Significant increase in diffusivity were observed in regions identified by other MRI modalities but not only, suggesting that either CADASIL is a diffuse disease or it leads to Wallerian degeneration. The MD and the mean VR was assessed as a measure of diffusivity and anisotropy respectively and found to be higher (higher VR implies less anisotropy) in the WM, not only in T_2-hyperintense regions. The variations were however found to be much more severe in T_2-hyperintense regions than in the rest of the WM. Proposed explanations suggest that T_2-hyperintense regions experienced a loss of axons whereas the rest of the WM could have only experienced myelin loss. DTI parameters therefore seems to be good markers of the clinical severity.

Ischemic leukoaraiosis [Horsfield 2002, Sundgren 2004]. Ischemic leukoaraiosis designates diffuse changes in the periventricular WM due to the proliferation of glial cells within the WM. In this case, DTI shows an increased MD and a decreased FA, which is usually explained by the fact that unexpected glial cells hinder water diffusion.

Amyotrophic Lateral Sclerosis (ALS) [Horsfield 2002]. Also known as the motor neuron disease or Lou Gehrig disease, ALS is a fatal progressive disease that atrophies muscles by attacking the motor neurons, i.e., mainly the CST. The disease has two forms depending on its onset: *bulbar-onset* patients will experience speaking and swallowing difficulties whereas *limb-onset* patients will first suffer from muscle weakness. Significant rises of MD were found along the CST for the two forms of ALS while significant drops of FA were only observed for bulbar-onset ALS. MD has been shown to be an indicator of disease duration: the higher the MD, the longer the disease. FA was found to indicate disease severity: the lower the FA, the more severe the disease.

Metabolic disorders [Horsfield 2002, Sundgren 2004]. Metabolic disorders in-

clude Krabbe disease or adrenoleukodystrophy, which are rare and often fatal autosomal recessive disorders occuring most often in young children. They are known to mainly affect the myelin sheath in the WM, with adrenoleukodystrophy creating brain lesions. DTI studies on these pathologies detected an increase in MD and a decrease in FA, once again explained by an increasing proportion of free water and by an alteration of axons and their myelin sheath in the WM.

Schizophrenia and affective disorders [Lim 2002, Kanaan 2005, Buchsbaum 2006, Kubicki 2007, Sexton 2009]. DTI analysis for the study of psychopathological disorders has been widely employed since the advent of DTI. In particular, schizophrenia is a mental disorder characterized by perturbed thoughts and emotions, resulting in delusions, disorganized speech and so on. No study has evidenced that MD could help in studying this disease. However, all published studies showed that FA decreases in many different regions of the WM. Patients with hallucinations also seem to experience an abnormal increase of FA in the CC. The FA was shown to increase inter alia with neuroleptic dosage, which is encouraging if one sees an increase of FA as a step towards cure. Roughly similar conclusions were drawn when studying other affective disorders.

Epilepsy [Arfanakis 2002, Assaf 2003, Sundgren 2004]. Epilespy groups the set of chronic neurological disorders characterized by seizures. The most common form of epilepsy arises in the temporal lobe with hippocampal sclerosis. Here again, DTI analysis revealed an increased MD and reduced FA in sclerotic hippocampi, which is usually explained by a destructuration of the WM that favors the expansion of the extracellular space.

Traumatic Brain Injury (TBI) [Sundgren 2004, Niogi 2010, Shenton 2012]. It occurs when an external force traumatically injures the brain. There is a lot of variability in the conclusions drawn by the many DTI analyses performed on TBI. Often no differences in MD is reported but this is not consistent across the studies. Similarly, FA is reported to decrease in the WM, but still not consistently. This can be due to the cross-sectional character of the designed studies and thus to the lack of longitudinal studies. [Shenton 2012] also suggest that, in the end, neither MD nor FA are specific biomarkers.

Overall, DT-MRI has been proven in the past decade to be the only MRI modality to be sensitive to diffuse axonal injuries. The study of mainly MD and FA differences between controls and patients suffering from various neurological pathologies often resulted in striking breakthroughs missed by conventional MRI or CT. However, these markers are non-specific in that their variations can be due to different still indistinguishable phenomena such as alteration of axonal membranes or myelin sheath or increase of glial cells proportion or other factors.

1.4.4 DTI Tractography

In addition to the multiple DTI-derived scalar maps of diffusivity and/or anisotropy indices, DT-MRI also provides information on the orientational organization of the fascicles within the WM. Following the voxelwise local orientation of diffusion iden-

tified by the principal eigenvector of the DT (the one with the largest associated eigenvalue), it is possible to reconstruct the tracts that compose the different fascicles in the WM. This is known as **tractography**. This technique enables for the first time in-vivo virtual dissections of the WM fascicles. Before the advent of tractography, knowledge about the network formed by the fascicles was established by post-mortem histological studies. Tractography thus generated a lot of enthusiasm. Indeed, it appeared as a tool of choice to confirm in-vivo the knowledge acquired on post-mortem studies and maybe to discover new fascicles that post-mortem studies missed. Also, neurosurgeons saw in tractography the means to detect which and how fascicles are damaged by a given pathology.

One of the first in-vivo virtual dissections obtained by tractography was proposed in [Catani 2002] on a healthy 39 year-old male: they provided reconstructions of the arcuate fascicle, the inferior longitudinal fascicle, the superior and inferior fronto-occipital fascicles, the uncinate fascicle, the cingulum, the anterior commissure, the corpus callosum, the internal capsule and the fornix, that are consistent with post-mortem studies. These reconstructions however relied on an accurate anatomical knowledge to define their starting Region Of Interest (ROI)s. To circumvent this issue, a DTI tractography template is provided in [Catani 2008], which helps in delineating ROIs to reconstruct the major fascicles in the WM. In the same work, the authors propose an atlas of 3D reconstructed fascicles as anatomical reference, which is further improved in [Thiebaut de Schotten 2011] by atlasing location, asymmetry and inter-subject variability. They show that this in vivo atlas is fairly consistent with previous histological atlases and enables the correct identification of fascicles affected by lesions.

In the meantime, DTI tractography has been widely used to study the WM architecture of pathological brains [Ciccarelli 2008]:

Neurogenerative disorders: (e.g., ALS, Parkinsonian disorders) DTI tractography revealed correlations between disconnecting lesions and clinical symptoms [Catani 2006];

Schizophrenia: the application of DTI tractography in **schizophrenia** led to inconsistent results until the work of [Jones 2006] that innovated by measuring tract-specific diffusivity and anisotropy indices. In this work, they found that tract-specific indices vary significantly with age, which might explain previous contradictory findings in schizophrenia;

Strokes: a decreased integrity of the CST has been shown to be associated to poorer outcome;

Multiple Sclerosis: MS patients present an overall significantly lower FA in the WM with respect to healthy subjects with a particularly pronounced drop in lesional regions. In order to avoid to prematurely terminate the reconstruction of the tracts, probability map of a fascicle of interest is computed from control subjects and registered on MS patients, in which FA is measured in highest probability regions. The FA along the CST, inter alia, was found to be significantly decreased in patients that experienced motor symptoms but not in others;

Epilepsy: malformed cortical areas seem not to be connected to the spinal cord

in patient with focal cortical dysplasia who experience epilepsy. Fornix and cingulum near the hippocampus seem to be subject to Wallerian degeneration. Language-involved regions are reorganized with functional plasticity in patients suffering from temporal lobe epilepsy;

Neurosurgical planning: (e.g., tumor ressections) neurosurgeons begin to use DTI tractography in presurgical planning.

1.4.5 Available databases

The advent and appeal of DTI over the past decade has convinced the dMRI community to conduct many studies on both normal and pathological brains which required careful designs of experiments and led to the public or on-demand availability of an incredibly large number of databases. For example, in their review of DTI studies in schizophrenia [Kubicki 2007], the authors report 18 different databases all constructed with single-shell acquisitions with a number of DSG directions that does not exceed 25. These databases are typical DT-MRI data. Also, over 27 databases constructed to study affective disorders with DTI, [Sexton 2009] report only data acquired using a single b-value and a number of DSG encoding directions most often under 32. They found only 3 databases using 41 or 51 encoding directions for the acquisition, still on a single shell.

A number of projects including those introduced in section 1.1.2 have also made large databases available for the study of both normal and pathological brains. For instance,

- Two DTI databases are available from the laboratory of brain anatomical MRI of the John Hopkins Medical Institute, which aim at facilitating the development of DTI analysis tools by providing control data. The diffusion protocol includes one b-value of 700 s/mm^2 and 32 non-collinear DSG directions,
- For the study of normal brain development, the Pediatric MRI project made available two databases: one with a single shell b = 1000s/mm^2 and 6 DSG directions and the other with the same DSG directions but applied on two shells at b = 500 and b = 1000s/mm^2.
- The large multi-site pediatric MRI and genetics data has been freely shared by the Pediatric Imaging, Neurocognition and Genetics (PING) project. It includes DTI datasets of about 1400 children between 3 and 20 years of age that can help in understanding genetic variations and developing patterns of brain connectivity. The diffusion protocol is again single low shell with few DSG directions,
- The Laboratory Of NeuroImaging (LONI) hosts a number of DTI databases: (i) the International Consortium for Brain Mapping (ICBM) continuously extends DTI data from John Hopkins Medical Institute to "generate a large normal DTI database and a probabilistic WM atlas", (ii) the Alzheimer's Disease Neuroimaging Initiative (ADNI) acquired a large DTI database using a single low shell (b = 1000 s/mm^2) protocol with 21 non-collinear DSG directions, in-plane resolution of 2.96 mm^2 and slice thickness of 3.3 mm to study Alzheimer's disease and (iii) the Parkinson's Progression Markers Initiative (PPMI) acquired a similar DTI database

to study Parkinsonian disorders,

- For normal brain connectivity purposes, the NIH Blueprint Human Connectome Project plans to make available a database with 3 different shells at $b = 1000$, 2000 and 3000s/mm^2 with 90 different but isotropically distributed DSG directions over each shell, with spatial resolution 1.25 mm^3,
- In the scope of obtaining the better possible reconstruction of WM fascicles in the normal brain, the NIH Human Connectome Project plans to make available a database with multiple shells up to $b = 10000$s/mm^2 and 200 DSG directions with spatial resolution of 1.5 mm^3.

The current available datasets were all designed for DTI analyses and, as a result, were often conducted with acquisition schemes that contain few DSG directions most often distributed on a single shell. We will see in the next section 1.5, that DTI presents some major limitations that raise the need for a better modeling of the diffusion. Research in this direction should however keep in mind that it would be of great importance to be able to retrospectively analyse each of the previously published data to go deeper in our understanding of normal and pathological brain development and of neurological disorders. Although very rich datasets will soon be available to achieve a better understanding of the normal human brain, they propose dMRI acquisition schemes that are far from being applicable in routine clinical practice, which precludes the possibility of disposing of similar databases for patients and thus limits their use to the normal human brain.

1.5 Limitations of Diffusion Tensor Imaging

1.5.1 A model at the voxel level

DTI is based under the assumption that the DIMD follow a Gaussian distribution described in eq. (1.15), resulting in a mono-exponential SID given by eq. (1.17). While the DT has been proven to model correctly single fascicles in the WM when using single-shell acquisitions with $b \leq 3000$s/mm^2 [Assaf 2005a, Ozarslan 2006, Caan 2010], several studies (including those validating the use of DTI for single fascicle voxels) in the last decade however observed non-monoexponential SID at b-values little more than the usual 1000s/mm^2 [Alexander 2002, Ozarslan 2006, Cheung 2009, Caan 2010], faulting the DTI Gaussian assumption. Investigations to understand the causes of this non-monoexponential behaviour have been carried out and it turns out that, at such low b-values, it is mainly due to two factors: CSF contamination (we recall that the entire brain stews in the CSF) [Kwong 1991] and presence of multiple non-collinear fascicles within single voxels [Alexander 2001].

The reason why DTI is not able to capture these effects lies into the definition of the true EAP eq. (1.7). In details, for each spatial location $\boldsymbol{x}_1$ in the voxel, a diffusion propagator $p_{\mathbf{x}_1+\mathbf{x}|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_1+\boldsymbol{x};\Delta)$, giving the probability of a molecule being at location $\boldsymbol{x}_1$ to move to $\boldsymbol{x}_1+\boldsymbol{x}$ under the effect of diffusion after time Δ, is defined. The true EAP then performs a weighted average of these different diffusion propagators with weights determined by the probability of presence of a molecule at

location $\boldsymbol{x}_1$. In DTI, the estimated EAP is a 3-dimensional Gaussian PDF, which implies two important simplifications:

- before undergoing diffusive motion, molecules are uniformly distributed over the voxel,
- no matter its location in the voxel, a molecule will always undergo the same diffusion motion governed by the 3-dimensional Gaussian PDF.

These two simplifications prevent DTI from assessing partial volume effects.

Also, despite the sensitivity of MD and FA to tissue microstructure [Bodini 2009, Salat 2009] for which we also report the reader to the previous section, these markers are non-specific [Pierpaoli 1996a]. For instance, a drop in FA can be attributed either to some brain pathology or to an increased orientational dispersion (due to the presence of crossing, diverging, kissing or fanning fascicles) or to other microstructural changes [Beaulieu 2009]. Also, as pointed out by [Shenton 2012], FA and MD often highly correlate (see section 1.4.3), making only useful one the two measures. These unfortunate consequences on DTI-derived scalar metrics are also the results of summarizing within a single PDF multiple diffusion processes arising from multiple different microstructural environments: any DTI-derived scalar will not be tissue-specific.

More specific measures have been proposed [Assaf 2008b, Niogi 2010]. In particular, the **mode of anisotropy** [Ennis 2006] is a more precise definition of anisotropy that better characterizes the shape of the diffusion and helps in distinguishing crossing fascicles from pathologies for which classic FA is low in both cases. The **inter-voxel coherence** [Pfefferbaum 2000] is a similarity measure between anisotropic DTs in a neighborhood. Finally, **parallel** (respectively, **perpendicular**) diffusivity given by eq. (1.18) has been shown to be a better marker for the study of axonal- (respectively, myelin-) related pathologies [Song 2002, Song 2003, Budde 2007, Budde 2011]. Yet, they are still derived from the DT which, by definition, approximates the EAP by a single PDF.

DTI tractography is of course also affected by DTI limitations [Ciccarelli 2008]. The low spatial resolution of DTI data (typically around 2.5 mm^3) yields voxel sizes that are far larger than the diameter of the axons composing the fascicles that tractography proposes to reconstruct. Noise and image artifacts, in addition to CSF contamination and orientational heterogeneity might yield reconstructed fascicles that do not truly exist (false positives) or true fascicles that are not properly reconstructed (false negatives).

Despite the plethora of DTI databases that are available upon request online (see section 1.4.5), dMRI is not a standard modality in clinical MRI examination protocols because doctors lack evidence of clear benefits of dMRI for the patient, over other MRI modalities. Indeed, the short acquisition time affordable in clinics with real patients who cannot stay long in the scanner provides low spatial and angular resolution diffusion data that, currently, can only be analyzed through DTI, which unfortunately has been shown to yield non-specific biomarkers and erroneous reconstructed fascicles. These facts raise the urgent demand for new diffusion analyzing tools that can overcome these limitations.

1.5.2 Voxelwise alternative diffusion models

Many alternative ways of analyzing the diffusion data that better represent multiple fascicles in a given voxel have been devised in the recent dMRI literature. Some of them focus on estimating the entire EAP, others on estimating the angular profile of the EAP known as the Diffusion Orientation Distribution Function (dODF), which is the integration of the EAP over its radial spherical coordinate, and finally others on estimating the Fiber Orientation Distribution Function (fODF), which is the distribution of the fascicles in the voxel. The maxima of the fODF can then be extracted [Ghosh 2013b] and used for tractography purposes.

A first way of analyzing the diffusion data is to resort to non-parametric techniques, in the sense that they estimate one of the three previous functions without assuming that it belongs to a parametric class of functions. Non-parametric estimation of the EAP is the basis of DSI [Wedeen 2000, Wedeen 2005], which tries to estimate the EAP by numerically computing the inverse Fourier transform of eq. (1.10). However, in order to satisfy the Nyquist conditions, a very large number of DSGs needs to be sampled on a Cartesian lattice as explained in section 1.4.2. Multiple shell Diffusion Propagator Imaging (DPI) [Descoteaux 2011] is another way to obtain an estimation of the EAP from multiple spherical samplings rather than a Cartesian sampling but still requires a large number of DSGs. As a result, many researchers focused on estimating the dODF out of the diffusion data. First, Q-Ball Imaging (QBI) [Tuch 2004] proposes a way to estimate an appoximation of the dODF, with no need to compute the EAP. However, at moderate b-values, it turns out that the local maxima of this approximated dODF do not match those of the true dODF [Zhan 2006]. This problem was corrected by [Canales-Rodríguez 2009] who proposed exact QBI, finding a relationship between the true dODF and the SID under the assumption of Gaussian diffusion. As one of the first diffusion models to account for multiple fascicles in a voxel, QBI gained a great popularity and its estimation has been improved over the past decade [Descoteaux 2007, Deriche 2009]. Also, estimating the angular profile of the EAP is just one step towards finding the orientations of the fascicles. It then requires to compute the fODF [Descoteaux 2009], with a possible accumumation of errors. Indeed, in practice, we are limited in the number of measured SIDs by the acquisition scan time, which decreases the contrast and angular resolution of these methods and providing *blurred* versions of the functions being estimated [Canales-Rodríguez 2010].

A second way of analyzing the diffusion data is to resort to parametric techniques, i.e. looking for the EAP or dODF or fODF into a class of parametric functions, which requires a less fine sampling of the DSGs. For example, generalized DTI [Ozarslan 2003, Liu 2004] captures multiple fascicles using higher-order tensors in addition to the usual second-order DT. In particular, [Ghosh 2010] approximate the EAP as the usual exponential multiplied by a correction polynomial term depending on the 4-th order diffusion tensor components. Spherical deconvolution approaches [Frank 2002, Tournier 2004, Tournier 2007, Jian 2007] directly estimate the fODF assuming a **fixed** usually Gaussian diffusion profile

for all the fascicles in the entire brain that have the same orientation. Diffusion Orientation Transform (DOT) [Ozarslan 2006] proposes to estimate isoradius surfaces of the EAP by expanding the Fourier-Bessel transform of the SID by means of the Spherical Harmonics (SH) basis under the assumption of monoexponential SIDs. Expressing the SID as a truncated spherical polar Fourier series also yield fast analytical EAP estimation [Cheng 2010]. These methods originally required long acquisition times as well but the recent introduction of compressed sensing [Merlet 2010, Merlet 2011, Merlet 2012, Deriche 2013] and dictionary learning [Merlet 2013] in dMRI analysis brings the methods closer to a possible clinical application.

All these analysis tools for dMRI (see the excellent review of [Assemlal 2011]) suffer from two major drawbacks:

- While they enable to some extent the characterization of multiple fascicles, they assume that all the fascicles share the same given microstructure throughout the brain therefore allowing only the accurate estimation of their orientations,
- They focus on describing the diffusion profile at the voxel level which results in poor representations of the true EAP given by eq. (1.7) and recalled here:

$$p_{\mathbf{x}}(\boldsymbol{x};\Delta) = \int_V p_{\mathbf{x}_1}(\boldsymbol{x}_1) p_{\mathbf{x}_1+\mathbf{x}|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_1+\boldsymbol{x};\Delta)\mathrm{d}\boldsymbol{x}_1\,,$$

since molecules at different locations in the voxel ($\mathbf{x_1}$) could undergo different diffusion processes $p_{\mathbf{x}_1+\mathbf{x}|\mathbf{x}_1=\boldsymbol{x}_1}(\boldsymbol{x}_1+\boldsymbol{x};\Delta)$ according to the tissue which they move into. For instance, if a voxel is composed of M fascicles, the molecules in the voxel will be at least distributed either in one of the M fascicles or outside the fascicles. Molecules in each of these $M+1$ different media may undergo very different diffusion processes that DSI, DPI, QBI, exact QBI, generalized DTI, spherical deconvolution, DOT and other similar methods will average at the voxel level, without assessing the proportions of these different media in the voxel.

The WM tissue microstructure is thus not taken into consideration: for instance, the generalized FA of two crossing fascicles proposed in [Tuch 2003, Campbell 2004] presents variations on synthetic data [Ozarslan 2005] and cannot therefore be a specific marker for the study of brain diseases. Also, even if these diffusion models shall improve tractography, previously exposed limitations pose the problem of validation and experimental confirmation as well as the need to optimize the currently too long acquisition protocols required for these methods [Dell'Acqua 2012]. Also, tractography performances could benefit from the estimation of the fractions of occupancy of the different fascicles in the voxel.

1.5.3 Tissue-specific diffusion models

Diffusion models outlined in section 1.5.2 improve DTI but still provide estimates of the EAP (or dODF or fODF) at the voxel level, averaging the true different diffusion processes arising from multiple environments/tissues in the voxel. In order to circumvent this problem, MCMs are a particularly appealing alternative. They

are based on a parametric diffusion model for the EAP that reflects more closely eq. (1.7) in the sense that the EAP is assumed to be a finite mixture of other PDFs, each one of them characterizing the diffusion in a particular environment/tissue (e.g., a fascicle or CSF), termed **compartment**. The weights of the mixture represent the **fractions of occupancy** of the different compartments, i.e. their relative proportions in the voxel.

A straightforward generalization of DTI into a MCM can be easily achieved by assuming that the EAP is a finite mixture of M zero-mean 3-dimensional Gaussian PDFs with associated DTs D_i ($i = 1, \ldots, M$). This M compartments are assumed to be representative of M putative non-collinear fascicles in the voxel. It is useful to incorporate in this mixture an additional compartment to account for CSF contamination: indeed, characterizing the proportion of free water - originally known as the **free-water model** [Pasternak 2009] - provides corrected MD and FA that are less correlated and thus more tissue-specific [Metzler-Baddeley 2012]. Under such assumptions, the SID for a given DSG $(b, \boldsymbol{g})$ has the following expression:

$$A\left(b, \boldsymbol{g}; \{D_i, w_i\}_{i=1,\ldots,M}, D_{\text{iso}}\right) = \left(1 - \sum_{i=1}^{M} w_i\right) \exp\left\{-bD_{\text{iso}}\right\} + \sum_{i=1}^{M} w_i \exp\left\{-b < \boldsymbol{g}, D_i \boldsymbol{g} >\right\} , \tag{1.24}$$

where w_1 to w_M are the occupancies of each fascicle in the voxel. This is known as the **multi-tensor model**. A full and comprehensive framework for the estimation of this model has been recently published [Scherrer 2012], in which the authors propose the CUSP acquisition scheme for sampling the DSGs, that they found to be currently optimal for the subsequent estimation of the multi-tensor model. Using the CUSP acquisition scheme, they show that they are capable of accurately estimating the full multi-tensor model with the only constraint that $D_{\text{iso}} = 3.0 \times 10^{-3}$ mm^2/s, which is its typical value at 37°C.

However, recent literature on the characterization of diffusion within the fascicles shows evidence that molecules in the fascicles undergo two distinct types of diffusion processes whether they are inside the axons or outside but close to the axons [Assaf 2005a]. In particular, recent dMRI analyses at $b \approx 2000$s/mm^2 show that water experiences non-Gaussian diffusion in the CC and the CST [Kunz 2011], which are brain areas known to contain uni-oriented fascicles. These findings support the idea that a non-monoexponential SID is not only due to orientational heterogeneity. Designing appropriate MCMs to account for these evidences is therefore crucial.

The rest of this manuscript is outlined as follows. In chapter 2, we go into details regarding the current available MCMs for modeling diffusion data putting the emphasis on MCMs that represent fascicles as geometrically constrained environments, hereafter referred to as **geometry-based MCMs**. We propose a new geometry-based MCM that generalizes some of them accounting for possible non-Gaussian DIMD. Using Monte-Carlo simulations, we generate simulated SIDs that we use to validate the proposed MCM and compare it to the existing ones. In chapter 3, we address the major limitation of MCMs that is: "how do we select a priori the right number of compartments in the mixture?" We review the existing solutions

for this model selection problem and propose a new approach based on model averaging that does not answer the question but rather assign probabilities to each model of being the right model to fit a given data. Chapter 4 is dedicated to tractography in which we briefly summarize the existing tractography algorithms and propose new deterministic and probabilistic algorithms that makes the most out of the information provided by MCMs and manage to perform faithful reconstructions of major fascicles in the brain. In chapter 5, we validate both the MCM proposed in chapter 2 and the tractography algorithms proposed in chapter 4 by presenting the results we obtained on the reconstruction of the CST from clinical data at the 3 DTI Tractography Challenges organized during the Medical Image Computing and Computer-Assisted Intervention (MICCAI) conference in 2011, 2012 and 2013.

Chapter 2

Local modeling of the diffusion

Contents

The usual protocol in dMRI consists in applying a series of DSGs

$$\boldsymbol{q}_{jk} = \gamma\delta G_k \boldsymbol{g}_j \text{ , } j \in [\![1, n_d]\!] \text{ , , } k \in [\![1, n_b]\!] \text{ ,}$$

in n_d non-collinear directions $\boldsymbol{g}_j$, using n_b different magnitude G_k. Each applied DSG yields a measured DW image S_{jk} that contains, in each voxel, the magnitude of the magnetization of all contributing spins [Alexander 2006]. A series of measured unweighted images S_0 are also acquired and averaged to obtain the true unweighted image Ψ_0 (or, to be precise, an approximation thereof). The predicted DW images Ψ_{jk} from a specific diffusion model or in absence of measurement noise are then reconstructed as $\Psi_{jk} = \Psi_0 A_{jk}$, where A_{jk} is the SID given by eq. (1.10) with $\boldsymbol{q} = \boldsymbol{q}_{jk}$.

Former and present clinical dMRI protocols only set a single magnitude G for all the DSGs ($n_b = 1$) and few DSG directions (typically at most $n_d = 30$). These are optimal protocols for subsequent DTI analyses (see section 1.4.2). Multi-shell acquisitions and/or High Angular Resolution Diffusion Imaging (HARDI) are not used in clinical practice yet, mainly due to acquisition time constraints, but also because of the lack of scientific and medical evidences of their real benefits for the patients.

2.1 Diffusional Kurtosis Imaging reveals non-Gaussian diffusion

As described in section 1.4.1, DTI describes a monoexponential SID, which, for a given DSG $\boldsymbol{q}_{jk} = \gamma\delta G_k \boldsymbol{g}_j$, is given by:

$$A_{jk} = A(\boldsymbol{q}_{jk}) = A(b_k, \boldsymbol{g}_j) = \exp\left\{-b_k < \boldsymbol{g}_j, D\boldsymbol{g}_j >\right\}, \tag{2.1}$$

where b_k is the b-value given by eq. (1.12) with $G = G_k$. It is theoretically straightforward to evaluate the departure of the DIMD from Gaussian displacements. Indeed, eq. (2.1) can be viewed as an expansion of the characteristic function of the distribution of the DIMD in terms of its cumulants, truncated at the second-order cumulant. Quantifying simultaneously the second- and fourth-order cumulants can thus characterized non-Gaussian displacements and is the basis of Diffusional Kurtosis Imaging (DKI).

2.1.1 Diffusional Kurtosis Imaging

Some researchers focused on explaining the departure from a Gaussian diffusion profile in the framework of mono-fascicle models. In this category of models lie those that expand the logarithm of the signal decay as a series of cumulants of the distribution of molecular displacements. Under the assumption of narrow distribution of molecular displacements, [Yablonskiy 2003] propose to limit the series to the first two even cumulants and thus exhibit estimates of the second- and fourth-order tensors. [Liu 2004] show that the real (respectively, imaginary) part of the logarithm of the signal decay is solely governed by the even-order (respectively, odd-order) cumulants and exhibit estimates of the second-, third-, and fourth-order tensors. However, these models suffer from a lack of physical interpretation of the estimated parameters. In particular, the second-order tensor is related to the variance properties of the distribution of molecular displacements that have been widely studied in the literature, as this is the basis of DTI [Pierpaoli 1996b, Le Bihan 2001, Basser 2002, Hasan 2010], but the fourth-order tensor is related to the kurtosis properties of the distribution and none of these works does establish that link. A similar approach that links kurtosis with the fourth-order tensor is DKI.

Local diffusion modeling. DKI is a biophysical model-free approach that was described in [Jensen 2005] to quantify the excess kurtosis of the distribution of the DIMD due to diffusion in the brain. It relies on the expansion of the logarithm of the SID in terms of the cumulants of the distribution of the DIMD, truncated to the fourth-order cumulant, which reads:

$$\log\left(A_{jk}\right) = -b_k D_{\text{app}}(\boldsymbol{g}_j) + \frac{1}{6} b_k^2 D_{\text{app}}^2(\boldsymbol{g}_j) K_{\text{app}}(\boldsymbol{g}_j), \tag{2.2}$$

where $D_{\text{app}}(\boldsymbol{g})$ and $K_{\text{app}}(\boldsymbol{g})$ are respectively the apparent diffusivity and apparent kurtosis in orientation $\boldsymbol{g}$. Similarly to the usual DT, which is a fully symmetric 3×3

matrix D with 6 independent elements, it is possible to exploit eq. (2.2) to derive the Kurtosis Tensor (KT), which is a fully symmetric $3\times3\times3\times3$ tensor W with 15 independent elements. Indeed, the following relations hold for any $\boldsymbol{g} = (g_1, g_2, g_3) \in \mathbb{S}^2$:

$$D_{\text{app}}(\boldsymbol{g}) = \sum_{i=1}^{3}\sum_{j=1}^{3} g_i g_j D_{ij} \text{ and } K_{\text{app}}(\boldsymbol{g}) = \frac{\text{MD}^2}{D^2_{\text{app}}(\boldsymbol{g})} \cdot \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{\ell=1}^{3} g_i g_j g_k g_\ell W_{ijk\ell}\,. \quad (2.3)$$

Estimation framework. The DT and KT are estimated in two steps: (i) the apparent diffusivity and apparent kurtosis is estimated for each DSG direction using all data acquired along this direction with multiple different b-values and (ii) eq. (2.3) is fitted to the previously estimated apparent diffusivities and kurtoses. Both steps were first achieved by means of least squared error estimators [Jensen 2005]. However, these estimators rely on the assumption that the measurement noise in dMRI is normally distributed, which is not the case in practice: the SID measured by a single coil in an MRI scanner is known to be affected by Rician noise [Gudbjartsson 1995] due to the nonlinear operation needed to compute the magnitude of the measured complex signal. Recently, [Veraart 2011b] introduced a constrained maximum-likelihood framework for the estimation of the apparent diffusivities and kurtoses that assumes a Rician noise model. They noticed that two types of constraints on kurtosis needed to be set in addition to the definite positivity of the DT: (i) the kurtosis has a mathematical lower bound of -2 and the authors even increased this bound to 0 stating that MCM always predict positive kurtosis and referring to [Jensen 2005, Jensen 2010] and (ii) the kurtosis needs not exceed $\frac{3}{D_{\text{app}}b}$ in order for eq. (2.2) to be a decreasing function of the b-value, as proven in [Jensen 2010]. This maximum-likelihood framework has been recently improved by the introduction of ternary quartics in the parametrisation of the kurtosis tensor [Ghosh 2013a]. Differently, [Sperl 2012] propose the use of compressed sensing theory to perform DKI estimation and assert that the denoising properties of compressed sensing enable linear least squares fitting to be more robust and could avoid the necessicity to resort to complicated frameworks like maximum-likelihood approaches.

Derived diffusion metrics. The 3D kurtosis properties of the distribution of the DIMD are contained in K_{app}. It is important for clinicians to resort to scalar maps to facilitate interpretation. [Lu 2006] proposed the spherical harmonics expansion up to the 4-th order, which allows the capture of the directional distribution of K_{app} with only three parameters C_0, C_2 and C_4 related to the spherical harmonics theory. An easy interpretation of the C_2 and C_4 coefficients is not straightfoward. [Hui 2008] propose more intuitive scalar metrics by introducing a suitable orthogonal transformation of the KT that enables the estimation of the diffusion kurtoses along the three eigenvectors of the DT, which leads to more easily interpretable indices such as Mean Kurtosis (MK) and kurtosis FA. They also propose a definition of parallel ($K_{/\!/}$) and perpendicular ($K_\perp$) kurtoses: the first one is defined as the apparent kurtosis in the principal eigenvector of the DT and the second one is defined

as the average of the apparent kurtoses in the other two eigenvectors of the DT. A slightly different definition of perpendicular kurtosis is adopted in [Jensen 2010].

Data acquisition requisites. Most of the studies involved with DKI recommend the use of six b-values ranging from 0 to 2500s/mm^2 and 30 DSG directions per b-value. The upper bound of the range of b-values is confirmed by theory: since typical values in the human brain are roughly $D_{\text{app}} \simeq 1.0 \times 10^{-3}$mm^2/s and $K_{\text{app}} \simeq 1$, this leads to $b \leq 3000$s/mm^2. [Jensen 2010] states that quantifying the DT and the KT requires the estimation of 21 parameters, which, in turn, requires at least three distinct b-values and at least 15 DSG directions per b-value. Specifically, they recommend to utilize three b-values of 0, 1000 and 2000s/mm^2 only and 30 gradient directions per b-value, which is sufficient to obtain accurate estimates together with a reasonable acquisition time.

Main reported results. A positive MK is revealed from previously published data on both rat and human brains using relatively high b-values in [Jensen 2003], along with large variations of the value in WM structures, suggesting deep differences in tissue structure. This was confirmed later in [Jensen 2005], in which they estimated the apparent diffusivity (D_{app}) and apparent kurtosis (K_{app}) in three orthogonal orientations in cortical GM and frontal WM and came to the same conclusion. Estimation of DKI-related scalar metrics was performed successfully on rat brains [Hui 2008, Cheung 2009] and revealed that quantifying the kurtoses brings additional information with respect to the diffusivities. The same group of researchers also analyzed changes in MK in the human brain due to age [Falangola 2007] or to various pathologies such as schizophrenia [Ramani 2007] or Attention-Deficit/Hyperactivity Disorder (ADHD) [Helpern 2007]. A thorough study also documents reference values of MD and MK derived from DKI in normal human brains [Minati 2007]. Overall, DKI was shown to be more sensitive to tissue microstructural changes than DTI is.

Limitations. The precision of diffusional kurtosis estimates is very poor with DTI data in which a single b-value of 1000s/mm^2 is used [Jensen 2005]. In general, the kurtosis estimation is very sensitive to the range of b-values chosen for data acquisition [Jensen 2005]. Indeed, the upper bound needs to be large enough in order for the effect of the b^2 term to be non-negligible with respect to the noise but also small enough to prevent greater powers of b in the expansion from contaminating the kurtosis estimate. This is also related to the fact that the estimated kurtosis, termed *diffusional kurtosis*, is just an approximation of the apparent kurtosis of the distribution of the DIMD: indeed, it is likely that the distribution presents other non-zero higher-order cumulants and that using higher b-values could allow to quantify them and robustify the kurtosis estimate just like considering a non-zero kurtosis robustifies the diffusivity estimates [Veraart 2011a]. The use of multiple b-values is very time-consuming: even if [Lu 2006] claims that DKI is clinically feasible (they propose a 10 minute acquisition), their acquisition (only 13 slices with 2.5mm thickness) would not cover an entire human brain. Interpolating from their acquisition time, a full human brain coverage can be estimated at approximately 37 minutes, which is not feasible in clinics. Also, it is admitted as solid ground truth

that the kurtosis has to be positive [Veraart 2011b] for any MCM whereas earlier mathematical background only proves that multi-tensor models always predict positive kurtosis [Jensen 2005, Jensen 2010]. Furthermore, this constraint is violated more than 80% of the time and mainly in deep WM structures such as the genu and the splenium of the CC [Veraart 2011b], suggesting that the distribution of DIMD within a single fascicle (which is typically the case in the CC) might present physically plausible negative kurtoses. Additionally, it is misleading to constraint parameters that define the diffusion process itself with parameters related to the measurement process.

Despite its promising results, DKI-derived scalar metrics, in the same line as DTI, lack specificity as they do not directly quantify tissue microstructure. To this purpose, diffusion models based on the tissue microstructure geometry are emerging in the recent dMRI literature.

2.2 Geometry-based Multi-Compartment Models

Due to the discrete nature of the voxel-based measurements in dMRI, the SID may result from water molecules diffusing in different tissues with possibly different microstructures, especially in the WM. A particularly suitable class of models to account for such an heterogeneity within a voxel is the class of MCMs, which assumes that water molecules are divided into several compartments (with or without authorizing exchange between them) that can exhibit different diffusion processes. An appealing way to model an intra-compartment diffusion process is to assume that it will be constrained by the geometry of the compartment: this is the basis of geometry-based MCMs.

In the context of dMRI, it was pioneered by [Stanisz 1997], which proposed a model of WM mictrostructure comprising three different compartments that account for glial cells, axons and extra-cellular space. The DIMD in all the compartments are assumed to follow a 3-dimensional Gaussian distribution [Pfeuffer 1998]. The intra-cellular compartments are assumed to undergo restricted diffusion within a sphere for glial cells and within 1-dimensional infinite parallel membranes for axons, with intra-cellular apparent diffusion coefficients computed accordingly. The extra-cellular compartment is assumed to undergo hindered diffusion, with extra-cellular apparent diffusion coefficient computed according to a tortuosity model. Exchange between intra- and extra-cellular spaces is also modeled, with exchange ratios that also depend on the geometry of the intra-cellular compartments (permeability and radius of the sphere for glial cells, permeability and distance between the membranes for axons).

This is an alternative model to DTI that directly measures tissue microstructure and proposes an analytical non-mono-exponential SID based on a weighted sum of three different mono-exponential SIDs in each compartment. However, Stanisz et al. model (i) does not account for the effects of orientation heterogeneity, (ii) describes the intra-compartment SID as mono-exponential and (iii) relies on a highly simplified

geometric representation of the axons.

Several subsequent geometry-based MCMs have been proposed that represent the axon geometry in a more plausible way as rigid or flexible cylinders and will be discussed in the following, namely the ball-and-stick model [Behrens 2003] (Section 2.2.1), the Composite Hindered And Restricted ModEl of Diffusion (CHARMED) [Assaf 2004] (Section 2.2.2) and the NODDI model [Zhang 2012] (Section 2.2.3).

2.2.1 The ball-and-stick model

Local diffusion modeling. The ball-and-stick model [Behrens 2003] describes the tissue microstructure as an extra-cellular space consisting of free water in which diffusion is assumed isotropic and unrestricted and an intra-cellular space containing the axons in which diffusion is restricted within cylinders of zero radius. In the original ball-and-stick model [Behrens 2003], the DIMD in the extra-cellular (respectively, intra-cellular) compartment follows a 3-dimensional centered Gaussian distribution with DT equal to $D_e = dI_3$ (respectively, $D_i = d\boldsymbol{\mu}\boldsymbol{\mu}'$), where d is the diffusivity assumed identical in the intra- and extra-compartments and $\pm\boldsymbol{\mu}$ is the axons orientation. The fraction of occupancy of the intra-cellular compartment is $f \in [0,1]$. These assumptions lead to the following expression of the SID [Behrens 2003]:

$$A_{jk} = (1-f)\exp\{-b_k d\} + f\exp\left\{-b_k d < \boldsymbol{\mu}, \boldsymbol{g}_j >^2\right\}, \tag{2.4}$$

for voxels in which all axons are orientated along a single orientation $\pm\boldsymbol{\mu}$. Later, the authors extended eq. (2.4) to the case of multiple axons orientations [Behrens 2007]:

$$A_{jk} = \left(1 - \sum_{i=1}^{M} f_i\right)\exp\{-b_k d\} + \sum_{i=1}^{M} f_i \exp\left\{-b_k d < \boldsymbol{\mu}_i, \boldsymbol{g}_j >^2\right\}. \tag{2.5}$$

Limiting the number of axons orientations to two, [Hosey 2005] proposes a ball-and-zeppelin model as an alternative to the ball-and-stick model, in which axons are represented by cylinders with positive radius. To do so, they assume that the DT that represents a single axon orientation is given by $D = D_\perp I_3 + (D_{/\!/} - D_\perp)\boldsymbol{\mu}\boldsymbol{\mu}'$. Furthermore, they let compartments have their own diffusivity. These assumptions yield the following expression of the SID:

$$\begin{aligned} A_{jk} &= \left(1 - \sum_{i=1}^{2} f_i\right)\exp\{-b_k D_{\text{isotropic}}\} \\ &+ \sum_{i=1}^{2} f_i \exp\{-b_k D_{\perp,i}\}\exp\left\{-b_k (D_{/\!/,i} - D_{\perp,i}) < \boldsymbol{\mu}_i, \boldsymbol{g}_j >^2\right\}, \end{aligned} \tag{2.6}$$

where one can observe that the radius of the cylinders does not explicitly appear in the SID, but play an important role through the diffusivities $D_{\perp,i}$ that the ball-and-stick models set to zero. Both models assume no exchange between compartments.

Estimation framework. The measurement noise is assumed to be Gaussian $\mathrm{N}\left(\Psi_0 A_{jk}, \sigma^2\right)$ and samples are asssumed to be independent. In the original model, [Behrens 2003] fit a ball-and-stick model that includes six independent parameters:

the two angles θ and ϕ that defines the axons orientation, the true unweighted signal Ψ_0, the fraction of occupancy f of the axons compartment, the diffusivity d and the measurment noise variance σ^2. The estimation is performed within a Bayesian framework by means of a MCMC algorithm choosing adequate priors for the parameters. Differently, [Hosey 2005] assume Rician measurement noise $\mathrm{R}\left(\Psi_0 A_{jk}, \sigma\right)$ and estimate $\tau = 1/\sigma^2$ rather than σ^2 itself. In addition, they observe that the MCMC algorithm fails to consistently identify the two axons orientations and found that the following constraint stabilizes the Markov chain:

$$\frac{1}{\eta} > \frac{f_1 \exp\left\{-(\Delta - \delta/3) D_{\perp,1}\right\}}{f_2 \exp\left\{-(\Delta - \delta/3) D_{\perp,2}\right\}} < \eta\,, \tag{2.7}$$

where η was chosen to 0.3.

Derived diffusion metrics. The ball-and-stick model directly quantifies (i) the axons fractions of occupancy (f_j), which allows the computation of the fraction of occupancy of freely diffusing water $\left(1 - \sum_{j=1}^{M} f_j\right)$, (ii) the axons orientations ($\boldsymbol{\mu}_j$) and (iii) the MD (d). The ball-and-zeppelin model is richer as it includes non-zero diffusivities in the orientations perpendicular to axons orientations.

Limitations. Despite its low number of independent parameters, the ball-and-stick model over-simplifies the geometry of the axons by representing them as zero-radius cylinders, which prevents the model from describing other important tissue microstructure related features such as anisotropy or the distinction between intra- and extra-cellular spaces. It only provides information about the axons orientation. This problem is partially solved in [Hosey 2005], which estimate a non-zero diffusivity in orientations perpendicular to the axon orientation. However, estimation of these diffusivities require the authors to put a constraint on the parameters given by eq. (2.7), which has no physical meaning and depends on a thresold that is chosen rather arbitrarily. Also, both models still describe a mono-exponential SID for both restricted and unrestricted compartments, making its use inappropriate at least at $b > 2000\mathrm{s/mm}^2$, given the recent findings on non-mono-exponential intra-compartment SIDs at such b-values [Cheung 2009, Kunz 2011].

2.2.2 The Composite Hindered And Restricted ModEl of Diffusion

The non mono-exponential SID has been successfully explained at the voxel level by geometry-based MCMs that propose a compartmentalization of water molecules into freely diffusing water molecules and water molecules trapped into axons (Section 2.2.1). The DIMD in each compartment is assumed to follow a 3-dimensional Gaussian distribution. However, if some of the compartments are subject to non-Gaussian diffusion, these models result inappropriate [Boss 1965, Assaf 2000]. q-Space theory [Callaghan 1991] is an appealing alternative in such cases, since it does not make any assumption on the diffusion process itself. This theory assumes that the duration δ of the pulse, during which each DSG is applied, is very short and negligible with respect to the diffusion time Δ between two successively applied DSGs, referred to as the Narrow Pulse Approximation (NPA). Under the NPA,

each DSG $\boldsymbol{q}$ is considered as constant over its application pulse. In practice, q-space analysis then follows five steps:

1. Choose a DSG direction,
2. Acquire measured DW images for multiple different b-values along this direction,
3. Extract for each voxel the corresponding SIDs and zero-fill the array for b-values higher than the maximal b-value used for acquisition (in order to increase Fourier transform resolution),
4. Perform the Fourier transform on the SIDs to obtain the PDF of the DIMD along the chosen DSG direction according to eq. (1.9),
5. Compute (i) the **mean displacement** as the full width at maximum height (i.e., the absolute difference between the two locations at which the PDF amounts to 0.5) and (ii) the **probability for zero displacement** as the value of the obtained PDF at zero,
6. Repeat the previous steps for other choices of the DSG direction.

The Fourier transform requires a large range of b-values to be accurate, which is why this method of analysis is often termed *high b-value q-space dMRI*. An important finding of this method was to demonstrate that diffusion observed at high b-values is mainly restricted. q-Space anaysis has also been employed for the study of normal and pathological human brains [Tuch 2002, Assaf 2002b, Assaf 2002a, Assaf 2002c, Mayzel-Oreg 2007]. However, it is very demanding in terms of acquisition time. In an application to multiple sclerosis, [Assaf 2002a] reported an entire MRI examination of 70 min with an included diffusion protocol of 28 min. In an application to vascular dementia and Alzeihmer's disease, the same group reported an entire MRI examination of 50 min with an included diffusion protocol of 20 min [Mayzel-Oreg 2007]. In both applications, clinical routine cannot afford such acquisition times. Also, dMRI is not a conventional MRI modality in clinics, which means that spending 40% of the examination time with the diffusion protocol is not yet conceivable in clinical practice. Finally, high b-value q-space dMRI does not provide microstructural parameters, which are the physical parameters of interest to ease the interpretation.

Local diffusion modeling. The CHARMED [Assaf 2004] is a biophysical diffusion model that combines ideas from DTI and q-space theory to provide easily interpretable microstructural parameters. The original model describes the tissue microstructure as a set of axons with possibly different orientations represented as impermeable cylinders with a fixed biologically plausible distribution. For each distinct axons orientation, the CHARMED assumes that the intra-axonal space is subject to restricted diffusion whereas the extra-axonal space is subject to hindered diffusion. Extra-axonal space is further reduced to a single compartment based on considerations about diffusion at low-q regime. The resulting SID reads:

$$A_{jk} = \left(1 - \sum_{i=1}^{M} f_i^r\right) A^h(\boldsymbol{q}_{jk}) + \sum_{i=1}^{M} f_i^r A_i^r(\boldsymbol{q}_{jk})\,. \tag{2.8}$$

The *hindered compartment* is modeled exactly as in DTI, which assumes that the

DIMD follow a 3-dimensional Gaussian distribution, leading to:

$$A^h(\boldsymbol{q}_{jk}) = \exp\left\{-b_k \boldsymbol{g}_j' D \boldsymbol{g}_j\right\}, \tag{2.9}$$

where D is 3×3 fully symmetric tensor. Each *restricted compartment* along a given axons orientation is modeled by decoupling the diffusion processes parallel and perpendicular to the axons orientation. Using q-space theory, it is shown that the resulting SID in a restricted compartment then reads:

$$A^r(\boldsymbol{q}_{jk}) = A^r_{/\!/}(\boldsymbol{q}_{jk}) A^r_{\perp}(\boldsymbol{q}_{jk}). \tag{2.10}$$

For each axons orientation, the DSG can be decomposed into the sum of two vectors $\boldsymbol{q} = \boldsymbol{q}_{/\!/} + \boldsymbol{q}_{\perp}$, where the first (respectively, the second) vector is parallel (respectively, perpendicular) to the axons orientation $\pm\boldsymbol{\mu}$. The DIMD parallel to the axons orientation are assumed to follow a 1-dimensional Gaussian distribution with zero mean and variance $D_{/\!/}$, leading to:

$$A^r_{/\!/}(\boldsymbol{q}_{jk}) = \exp\left\{-b_k D_{/\!/} < \boldsymbol{\mu}, \boldsymbol{g}_j >^2\right\}. \tag{2.11}$$

The contribution of the DIMD perpendicular to the axons orientation to the SID is given by Neumann's model for restricted diffusion within a cylinder of radius R [Neuman 1974]:

$$A^r_{\perp}(\boldsymbol{q}) = \exp\left\{-\frac{7R^4 b_k \left(1 - < \boldsymbol{\mu}, \boldsymbol{g}_j >^2\right)}{48\left(\Delta - \delta/3\right) D_{\perp} \mathrm{TE}} \left(2 - \frac{99R^2}{56 D_{\perp} \mathrm{TE}}\right)\right\}, \tag{2.12}$$

where TE is the echo time of the acquisition. Equation (2.12) is valid only for $\Delta \sim \delta$ that rarely matches regular clinical practice. As a consequence, the modeling of the restricted diffusion in the CHARMED has been improved [Assaf 2005b] by using a modified formula for restricted diffusion in cylinders proposed in [**?**], valid for $\delta < \Delta$. The authors of the CHARMED later showed how to estimate the distribution of axons radii, assuming that it follows a Gamma distribution [Assaf 2008a]. An improvement of the CHARMED has also been proposed by adding a free-water compartment necessary for *in vivo* dMRI analysis [Barazany 2009]. The representation of the axons used by the CHARMED has also been simplified to cylinders with a single radius and extra-cellular space characterized by a cylindrically symmetric DT whose parallel diffusivity was constrained to match the intra-axonal parallel diffusivity [Alexander 2008]. In the same vein as in [Stanisz 1997], tortuosity models that set the perpendicular diffusivity in the hindered compartment as a function of the parallel diffusivity have been subsequently introduced to obtain the Minimal Model of White Matter Diffusion (MMWMD) [Alexander 2010]. A taxonomy of MCMs constructed from the ball-and-stick and ball-and-zeppelin models and the CHARMED and the MMWMD, can be found in [Panagiotaki 2012].

Estimation framework. The estimation of the parameters of the CHARMED is performed approximating the Rician measurement noise $\mathrm{R}(S(\boldsymbol{q}), \sigma)$ by a Gaussian noise $\mathrm{N}\left(\sqrt{S^2(\boldsymbol{q}) + \sigma^2}, \sigma^2\right)$ [Assaf 2004, Assaf 2005a], which is valid for high SNR.

The non-linear least squares cost function is then minimized using the Levenberg-Marquardt minimization algorithm. The axons radius and the perpendicular diffusivity of the restricted compartments are kept fixed. The latter are also identical for all axons orientations. The estimation of the parameters of the MMWMD is performed by means of an MCMC algorithm described in [Alexander 2008], keeping the parallel and isotropic diffusivities fixed ($D_{/\!/} = 1.7 \times 10^{-3}\text{mm}^2/\text{s}$ and $D_{\text{isotropic}} = 3.0 \times 10^{-3}\text{mm}^2/\text{s}$). The tortuosity is not a free parameter but computed as the extra-axonal fraction of occupancy within the WM, i.e., without considering the CSF isotropic compartment. To initialize the MCMC procedure with a good starting point, a gradient descent scheme on a Rician-likelihood based cost function is previously performed [Alexander 2010].

Derived diffusion metrics. The CHARMED directly quantifies the axons density $(1 - f^h)$, the axons orientations (one can see $\boldsymbol{q}_{/\!/}$ as $< \boldsymbol{q}, \boldsymbol{\mu} > \boldsymbol{\mu}$, where $\boldsymbol{\mu}$ is an axons orientation) and the diffusivities parallel and perpendicular to each axons orientation. In addition, [Assaf 2008a] also provide an estimate of the axons radii and [Barazany 2009] subsequently quantify the CSF fraction of occupancy. [Alexander 2010] propose to define an index of axon density as the ratio of the intra-axonal fraction of occupancy within the WM over the cross-sectional axon area. Also, the PDF of the DIMD involved in Neuman's model has an analytical expression [Neuman 1974, formula 15], which allows one to express the moments of the distribution of the DIMD as 3-dimensional integrals, which can be numerically computed, giving access to the kurtosis properties of the CHARMED in the case of restricted diffusion within a cylinder characterized by Neuman's model [Assaf 2004, Assaf 2005a].

Data acquisition requisites. To perform the analysis of the CHARMED on human brain, [Assaf 2005a] propose to acquire 169 measured DW images along 169 different DSG (10 different b-values up to $10000\text{s}/\text{mm}^2$) with an increasing number of DSG directions as the b-value increases, for a total acquisition time of 17 min for 10 slices of 3mm thickness, which leads to more than an hour to cover an entire human brain with a 3mm^3 spatial resolution. This protocol was further improved in [De Santis 2011] using sampling scheme optimization and model parsimony testing. By further constraining the parameters of the CHARMED (estimating only one intra-axonal volume fraction and associated diffusivity and the volume fraction of a CSF compartment with fixed diffusivity), [Kunz 2011] report an aquisition time of 8 min for 44 slices of 2mm thickness, achieving a full brain coverage in 11 min with a spatial resolution of 2mm^3. An *in vivo* human brain imaging protocol has also been optimized for the estimation of the MMWMD [Alexander 2010]: it consists in the measurement of 360 measured DW images and 4 measured unweighted images distributed on 4 different shells with b = 530, 700, 2720 and $2780\text{s}/\text{mm}^2$. The acquisition time is about one hour for an in-plane spatial resolution of 1.8mm^2 and a total of 10 slices of 3.9mm thickness. This would lead to an acquisition time of about 3 hours to cover an entire human brain.

Main reported results. The hindered compartment provides diffusivities that are very similar to the one obtained by DTI. The restricted compartments are much

more discriminative between GM, WM and CSF [Assaf 2005a]. The latter is mainly dominant in the extra-axonal space (hindered compartment) and is not found in intra-axonal space where the diffusion is restricted. The CHARMED also demonstrates a great ability to estimate crossing fascicles with a better robustness to noise with respect to a two-tensor model [Assaf 2005a]. Recent dMRI analyses using the CHARMED at $b > 2000\text{s/mm}^2$ also show that water experiences non-Gaussian diffusion in the CC and the CST [Kunz 2011]. These regions are orientationally homogeneous, which supports that non-mono-exponential SID is not only due to orientation heterogeneity. Results on the kurtosis of the distribution of the DIMD estimated with the CHARMED are consistent with those presented in [Fieremans 2011], where a two-tensor model was used to capture the kurtosis and derive other microstructural parameters. The model shows good agreement with previous reported results using DKI [De Santis 2012].

Limitations. As mentioned in [Assaf 2005a], the CHARMED presents some drawbacks that, at the moment, shall preclude its use in clinical practice. It is indeed time-consuming for acquisition: the standard protocol for fitting the CHARMED requires one hour for a full human brain coverage. This is due to several facts [Assaf 2005a]: (i) the more b-values, the better the estimates of the restriced compartment parameters, (ii) the higher the b-value, the lower the SNR, which demands then more samples to re-increase the SNR and (iii) accurate elucidation of crossing fascicles requires at least 20 DSG directions. Also, the moments of the distribution of the DIMD as given by the CHARMED cannot be analytically related to the model parameters, which limits their use. Moreoever, kurtosis properties of the distribution of the DIMD can only be assessed under Neuman's model for restricted diffusion within a cylinder, whereas [Assaf 2005b] pointed out that it is preferable to use van Gelderen's model. Lastly, neither the CHARMED nor the MMWMD quantify the axonal-orientation dispersion which is useful for an accurate elucidation of bending, fanning or crossing axons that appear in the human brain [Bürgel 2006].

2.2.3 Neurite Orientation Dispersion and Density Imaging (NODDI)

All the models presented so far aimed at modeling the diffusion within the WM by representing axons with simplified restricted geometries (mostly, cylinders). Each model addresses one or more problems of the DTI, which is the current clinical reference for the analysis of dMRI, but none of them addresses them all. The non monoexponential SID is explained by all MCMs at the voxel level, but only the CHARMED and the MMWMD quantify non-Gaussian diffusion within compartments. Associated imaging protocols also still remain impractical for clinical translation and none of these models account for axonal-orientation dispersion.

The latter point has been addressed by new emerging models that introduce orientation dispersion by means of spherical deconvolution. To this end, they use appropriate directional distributions that enable analytical expressions of the deconvolution: [Kaden 2007] and [Sotiropoulos 2012] use the Bingham distribution [Mardia 1972] to deconvolve a Gaussian kernel for each axon-representative com-

partment yielding the *ball-and-racket* model and [Zhang 2011] model an intracellular compartment with Gaussian kernel deconvolved by a Watson distribution and an extracellular compartment with Gaussian kernel too, the DT of which is in turn deconvolved by the same Watson distribution. Despite the fact that these models are able to estimate the axonal-orientation dispersion, spherical deconvolution makes it difficult to turn back to the PDF of the DIMD from the resulting analytical expression of the SID, preventing one from understanding the underlying physical diffusion process. Morever, performing spherical deconvolution rather than full 3-dimensional $\mathbb{R}^3$-deconvolution requires to assume that the diffusion within an axon only depends on the axon orientation, which somehow deviates from geometry-based MCMs. In the same vein as the ball-and-stick and ball-and-zeppelin models, the use of the ball-and-racket model is thus limited to connectivity purposes.

In order to extend the use of such models, [Jespersen 2007] propose a *neurite* model that accounts for both moderately dispersed axons in the WM and highly dispersed dendrites at the frontier between WM and GM with a geometry-based MCM. The parameters of the neurite model have been shown to correlate more strongly with previous independent measures from histology than DTI-derived parameters [Jespersen 2010] and extracted neurite orientation distributions match those determined by Golgi analysis [Jespersen 2012]. Yet, analyses were conducted on *ex vivo* imaging of a baboon brain and the associated protocol was still prohibitively time-consuming for an *in vivo* translation. A very recent diffusion imaging analysis, termed NODDI, aims at proposing a geometry-based diffusion model that represents the neurites organization under an acquisition time constraint of 30 min.

Local diffusion modeling. The NODDI model assumes that the SID results from three impermeable compartments, namely the intracellular, extracellular and CSF compartments, which yields [Zhang 2012]:

$$A_{jk} = (1-\nu_{\text{iso}})\left(\nu_{\text{ic}}A_{\text{ic}}(\boldsymbol{q}_{jk}) + (1-\nu_{\text{ic}})A_{\text{ec}}(\boldsymbol{q}_{jk})\right) + \nu_{\text{iso}}A_{\text{iso}}(\boldsymbol{q}_{jk})\,, \tag{2.13}$$

where $A_{\text{ic}}(\boldsymbol{q}_{jk})$, $A_{\text{ec}}(\boldsymbol{q}_{jk})$ and $A_{\text{iso}}(\boldsymbol{q}_{jk})$ are the SIDs produced by diffusion in the intracellular, extracellular and CSF compartments, respectively, and $(1-\nu_{\text{iso}})\nu_{\text{ic}}$, $(1-\nu_{\text{iso}})(1-\nu_{\text{ic}})$ and ν_{iso} are their respective fractions of occupancy.
Differently from [Jespersen 2007], which approximate the orientation distributions by means of a truncated spherical harmonics series and thus, fails to accurately estimate lowly dispersed orientation distributions [Zhang 2011], the NODDI model expresses the intracellular SID as a Gaussian kernel spherically deconvolved by a Watson distribution [Zhang 2012]:

$$A_{\text{ic}}(\boldsymbol{q}_{jk}) = M\left(\frac{1}{2},\frac{3}{2},\kappa\right)^{-1}\int_{\mathbb{S}^2}\exp\left\{\left(\kappa < \boldsymbol{\mu},\boldsymbol{n}>^2 - b_k D_{/\!/} < \boldsymbol{g}_j,\boldsymbol{n}>^2\right)\right\}\mathrm{d}\boldsymbol{n}\,, \tag{2.14}$$

where M is a confluent hypergeometric function, $\boldsymbol{\mu}$ is the axons orientation, κ measures the orientation dispersion and $D_{/\!/}$ is the parallel diffusivity along the axons orientation. Note that the NODDI currently estimates only one orientation per voxel with its associated dispersion. Even if the geometry of the axons is not explicitly

specified in eq. (2.14), this intracellular compartment models cylindrically restricted diffusion and thus implicitly assumes that the axons are cylindrically-shaped. We refer to this cylindrical shape in the following as a **pseudo-cylinder**. The difference with the classic cylinder is that its radius is not explicitly given or estimated as such, but is rather proportional to $\sqrt{D_\perp \Delta}$, with $D_\perp$ measuring the apparent diffusivity in the space orthogonal to the axon orientation $\pm\boldsymbol{\mu}$.
The DIMD in the extracellular compartment are assumed to follow an anisotropic 3-dimensional centered Gaussian distribution with covariance matrix proportional to the following DT:

$$D_{\boldsymbol{\mu},\kappa,D_{/\!/},D_\perp} = M\left(\frac{1}{2},\frac{3}{2},\kappa\right)^{-1}\int_{\mathbb{S}^2}\exp\left\{\kappa < \boldsymbol{\mu},\boldsymbol{n} >^2\right\}\left(D_\perp I_3 + (D_{/\!/} - D_\perp)\boldsymbol{n}\boldsymbol{n}'\right)\mathrm{d}\boldsymbol{n}\,, \tag{2.15}$$

which yields the following extracellular SID:

$$A_{\mathrm{ec}}(\boldsymbol{q}_{jk}) = \exp\left\{-b_k \boldsymbol{g}_j' D_{\boldsymbol{\mu},\kappa,D_{/\!/},D_\perp}\boldsymbol{g}_j\right\}\,, \tag{2.16}$$

where $D_{\boldsymbol{\mu},\kappa,D_{/\!/},D_\perp}$ is given by eq. (2.15). Note that the parallel diffusivity is the same in intra- and extracellular compartments. The perpendicular diffusivity $D_\perp$ is assumed to be related to the parallel diffusivity $D_{/\!/}$ via the following tortuosity model:

$$D_\perp = D_{/\!/}(1-\nu_{\mathrm{ic}})\,, \tag{2.17}$$

which relies on the assumption that diffusing water molecules within the extracellular space have to circumnavigate between several obstacles (the axons). A few axons (small ν_{ic}) imply that extracellular water is almost free and eq. (2.17) leads to an isotropic extracellular DT in that case.
The DIMD in the CSF compartment are unrestricted and are assumed to follow an isotropic 3-dimensional centered Gaussian distribution, which leads to:

$$A_{\mathrm{iso}}(\boldsymbol{q}_{jk}) = \exp\left\{-b_k D_{\mathrm{iso}}\right\}\,, \tag{2.18}$$

where D_{iso} is free-water diffusivity.

Estimation framework. The NODDI model requires the estimation of 7 independent parameters, namely the two angles (θ,ϕ) of the axons orientation $\boldsymbol{\mu}$, the diffusivity parallel to the axons orientation $D_{/\!/}$, the concentration parameter κ of the Watson distribution, intracellular fraction of occupancy ν_{ic}, the free-water fraction of occupancy ν_{iso} and the free-water diffusivity D_{iso}. As in [Alexander 2008, Alexander 2010, Zhang 2011], the diffusivities are then kept fixed to typical values *in vivo*: $D_{/\!/} = 1.7\times10^{-3}$ mm^2/s and $D_{\mathrm{iso}} = 3.0\times10^{-3}$ mm^2/s. The 5 remaining parameters are estimated [Zhang 2012]. The estimation is performed by maximizing the likelihood of the parameters given the measured SIDs assumed to be corrupted with Rician noise [Gudbjartsson 1995], using a Gauss-Newton non-linear programming algorithm. The MCMC procedure employed in [Alexander 2010] to fit a similar model is not necessary for the NODDI model that does not explicitly incorporate the axons radius.

Derived diffusion metrics. The NODDI model directly estimates microstructure-related parameters: the intracellular fraction of occupancy ν_{ic} that gives the axons density, the mean axons orientation $\pm\boldsymbol{\mu}$ and the free-water fraction of occupancy ν_{iso}. In addition, an index of orientation dispersion can be defined from the estimated κ parameter as:

$$\text{OD} = \frac{2}{\pi}\arctan\left(\frac{1}{\kappa}\right) \in [0, 1]\,. \tag{2.19}$$

This index is more suited than κ itself to measure the orientation dispersion to which it is directly proportional. Also, the extracellular compartment has apparent parallel ($D'_{/\!/}$) and perpendicular ($D'_{\perp}$) diffusivities that not only depend on the parallel diffusivity $D_{/\!/}$ but also on the intracellular fraction of occupancy and orientation dispersion as follows:

$$D'_{/\!/} = D_{/\!/}\left(1 - \nu_{\text{ic}}(1 - \tau_1)\right) \text{ and } D'_{\perp} = D_{/\!/}\left(1 - \nu_{\text{ic}}\frac{1 + \tau_1}{2}\right), \tag{2.20}$$

where

$$\tau_1 = -\frac{1}{2\kappa} + \frac{1}{2F(\sqrt{\kappa})\sqrt{\kappa}}, \text{ with } F(x) = \frac{\sqrt{\pi}}{2}\exp\left\{-x^2\right\}\text{erfi}\,(x)\,, \tag{2.21}$$

measures the effect of orientation dispersion on the apparent diffusivities.

Data acquisition requisites. The optimized NODDI protocol [Zhang 2012, Tariq 2012] for the estimation of the NODDI model consists in the acquisition of 90 measured DW images and 9 measured unweighted images as follows: 30 DSG directions on a $b = 711$ s/mm^2 spherical shell and 60 DSG directions on a $b = 2855$ s/mm^2 spherical shell, TR of 12.5 s and TE of 78 ms with a spatial resolution of 2 mm^3. The acquisition of 50 slices takes 25 min, thus a full brain coverage (120 mm) can be scanned in 30 min.

Limitations. Despite the effort to decrease substantially the acquisition time towards a clinically feasible one, a dMRI protocol of 30 min remains impractical in clinical daily routine. Also, even if [Zhang 2012] discuss that the NODDI model is easily extendable to capture multi-modal orientation distributions, the current implementation and validation do not exhibit such properties. Given the computational time of three hours to recover mono-modal orientation distributions [Zhang 2012], it is likely that also computational time will be an issue for the estimation of multi-modal orientation distributions. Finally, the model strongly relies on spherical deconvolution and thus presents the same drawback as the models proposed in [Kaden 2007, Sotiropoulos 2012, Zhang 2011]: it lacks an expression of the PDF of the DIMD in terms of the NODDI microstructure-related parameters that would give an insight into the nature of the diffusion process itself.

2.3 Diffusion Directions Imaging

2.3.1 Motivations

Wherever the Gaussianity assumption of the diffusion process is met, the measured DW signal decay is monoexponential. Conversely, if one observes a non-

monoexponential measured DW signal decay in a voxel, then it would be misleading to maintain the Gaussianity assumption. Several works in the literature observed that the measured DW signal decay is actually non-monoexponential in some parts of the brain, both at low b-values [Alexander 2002, Ozarslan 2006, Cheung 2009, Caan 2010] and at high b-values [Mulkern 2000, Norris 2001, Bennett 2003, Yablonskiy 2003, Maier 2004, Schwarcz 2004, Sehy 2004, Assaf 2005a, Jensen 2005, Mulkern 2009], suggesting that the Gaussianity assumption is not supported by the data. The underlying responsible biophysical mechanisms are complex.

Clinical dMRI currently involves a single low b-value. At such large diffusion scales (low b-values), it is likely that this observed non-monoexponential decay is due to (i) intra-voxel orientation heterogeneities and/or (ii) free water contamination. Solutions to point (i) have been proposed in [Jbabdi 2012, Scherrer 2012] by resorting to multi-tensor models. However, they require multi-shell acquisitions, which are not yet available in clinics. To cirvumvent the issue, [Behrens 2007] earlier proposed a constrained version of the model of [Jbabdi 2012] but the constraints are somewhat arbitrary. Conversely, [Scherrer 2012] showed that the full multi-tensor model can be accurately estimated using a new sequence called CUSP imaging that they designed, which enables to perform multi-shell acquisitions within a clinically acceptable time. However, this strategy prevents one from carrying out a retrospective analysis of former dMRI data sets. Solutions to point (ii) have also been proposed and consist in including an isotropic compartment in addition to the fascicle compartment(s) to account for free water [Latour 2002, Metzler-Baddeley 2012, Scherrer 2012]. However, recently, NODDI experiments [Zhang 2012] suggested that the fraction of occupancy of such an isotropic compartment cannot be accurately estimated using clinical dMRI that involves only one non-zero b-value.

While the focus of this work is clinical dMRI, which can only afford to measure the diffusion at a large scale (low b-value), research dMRI at smaller diffusion scales (high b-values) reveals a great deal of how the diffusion **should not** be modeled. In particular, using high b-values, non-monoexponential measured DW signal decays have been observed within a single fascicle due to the presence of intra- and extra-cellular compartments [Assaf 2005a], which makes monoexpentional compartmentalization quite ineffective. Whether or not such a non-monoexponential behavior in individual fascicles has to be considered at low b-values is still a debate: on the one hand, [Assaf 2005a, Ozarslan 2006, Caan 2010] suggest that it is negligible for clinically acceptable b-values ($b \leq 3000\text{s/mm}^2$) and, on the other hand, [Cheung 2009], for example, have measured a significantly non-zero kurtosis term considering a single fascicle with $b = 2500$ s/mm^2: "does the non-zero kurtosis term reflect the lack of other fascicle and/or free water compartments in the model or is it simply that there is indeed a single fascicle with non-Gaussian diffusion?" In our opinion, since it has been observed that, even in a single fascicle, the Gaussianity assumption does not hold, we cannot pretend it does just because we are working with clinical acquisitions that are not able to detect the non-monoexponential behavior.

Instead of focusing on such an unclear b-value threshold under which the Gaussianity assumption could possibly be valid to model the diffusion in a single fascicle,

we rather propose in the following an MCM in which each compartment features a non-monoexponential SID that converges to monoexponential as the b-value decreases. In addition, the non-monoexponential behavior in a single fascicle stems from the use of a non-Gaussian probability distribution that we came up with on purpose and of which we can analytically compute all the moments. This allows us to apply clinically plausible constraints to our MCM to make it tractable in clinical practice. We coined this diffusion modeling Diffusion Directions Imaging (DDI), since we design it to accurately identify the different non-collinear fascicle orientations in a voxel.

In a single compartment (which can represent a fascicle or free water), the usual mono-exponential SID is described by the Stejskal & Tanner equation [Stejskal 1965b] given in eq. (2.1) and recalled here for sake of clarity:

$$A_{jk} = A(\boldsymbol{q}_{jk}) = A(b_k, \boldsymbol{g}_j) = \exp\left\{-b_k < \boldsymbol{g}_j, D\boldsymbol{g}_j >\right\} .$$

It boils down to assuming that the DIMD $\mathbf{x}$ are given by $\mathbf{x} = \sqrt{2(\Delta - \delta/3)}\mathbf{y}$, where $\mathbf{y}$ follows a zero-mean trivariate Gaussian distribution with covariance matrix D. In order to describe a non mono-exponential behavior in a compartment, we propose a new distribution for the random vector $\mathbf{y}$ that departs from the Gaussian distribution.

A fascicle is characterized, inter alia, by its orientation $\pm\boldsymbol{\mu}$. The diffusion within a fascicle is antipodally symmetric, which means that water molecules diffuse in the forward direction $+\boldsymbol{\mu}$ and in the backward direction $-\boldsymbol{\mu}$ in equal proportions, which reads:

$$p_{\mathbf{y}}(\boldsymbol{y}|\pm\boldsymbol{\mu}, \Theta) = \frac{1}{2}p_{\mathbf{w}}(\boldsymbol{y}|+\boldsymbol{\mu}, \Theta) + \frac{1}{2}p_{\mathbf{w}}(\boldsymbol{y}|-\boldsymbol{\mu}, \Theta) , \tag{2.22}$$

where $p_{\mathbf{y}}$ is the PDF of $\mathbf{y}$, Θ is a set of parameters that characterizes the fascicle out of its orientation and $\mathbf{w}$ is a random vector that describes a 3-dimensional random displacement along a specific direction.

Note that this formalism generalizes any parametric diffusion model that has been proposed so far in the literature: indeed only antipodally symmetric distributions have been used, for which $p_{\mathbf{w}}(\boldsymbol{y}|+\boldsymbol{\mu}, \Theta) = p_{\mathbf{w}}(\boldsymbol{y}|-\boldsymbol{\mu}, \Theta) = p_{\mathbf{y}}(\boldsymbol{y}|\pm\boldsymbol{\mu}, \Theta)$. In addition, it offers more flexibility since asymmetric distributions may be used as well to model the diffusion while preserving the antipodal symmetry of the latter.

In the following subsections, we describe DDI, which involves (i) the introduction of a new distribution for characterizing non-Gaussian 3-dimensional displacements in a specific direction (Section 2.3.2), (ii) the use of such a distribution to model the diffusion within a given fascicle (Section 2.3.3) and (iii) the design of an appropriate MCM that can be estimated from clinical diffusion data and that accounts for the contribution of the different fascicles and the CSF environment to the SID (Section 2.3.4).

2.3.2 Non-Gaussian 3D displacements in a specific direction

In this section, we introduce a novel distribution for characterizing non-Gaussian 3-dimensional displacements in a specific direction (Section 2.3.2.1). We derive its

PDF (Section 2.3.2.2) that allows us to compute all the moments of the introduced distribution (Section 2.3.2.3).

2.3.2.1 Definition

A 3-dimensional random displacement $\mathbf{w}$ along a specific direction $\boldsymbol{\mu}$ results from the combination of a random 2-dimensional direction (*directional variability*) and a random 1-dimensional displacement along that direction (*variability in amplitude*) that the Gaussian distribution alone tries to capture altogether in its covariance matrix. Conversely, we aim at better separating these two sources of variability. Ideally, one could think of expressing the random variable $\mathbf{w}$ in spherical coordinates as $\mathbf{w} = r\mathbf{v}$ so that both the radial component (r) and the angular component ($\mathbf{v}$) are explicitly modeled. However, to the best of our knowledge, there is no pair of known parametric probability distributions (a univariate one on $\mathbb{R}^+$ and a directional one on the 2-dimensional unit sphere to model r and $\mathbf{v}$ respectively) that leads to an analytic PDF for modeling the diffusion in the fascicle.

To circumvent this issue, we express the random variable $\mathbf{w}$ as the sum of two independent random variables, $\mathbf{w} = \mathbf{v} + \mathbf{z}$, where [5, 6]:

- $\mathbf{v}$ follows a **Von-Mises & Fisher (vMF)** probability distribution (see appendix A.2) on the sphere of radius $\sqrt{\nu d}$ ($\nu \in [0,1]$, $d > 0$) with mean direction $+\boldsymbol{\mu}$ ($\|\boldsymbol{\mu}\| = 1$) and concentration parameter $\kappa \geq 0$; it has a constant norm: $\|\mathbf{v}\|^2 = \nu d$. The vMF probability distribution admits a PDF on the 2-sphere of radius $\sqrt{\nu d}$. Its expression is given in [Jupp 1989] for $\nu d = 1$ and can easily be extended on the sphere of radius $\sqrt{\nu d} > 0$ by affine transformation:

$$p_{\mathbf{v}}(\boldsymbol{v}; +\boldsymbol{\mu}, \kappa, d, \nu) = \frac{1}{(\nu d)^{3/2}} \frac{\kappa}{4\pi \sinh \kappa} \exp\left\{ \frac{\kappa}{\sqrt{\nu d}} \boldsymbol{\mu}' \boldsymbol{v} \right\}, \tag{2.23}$$

for any $\boldsymbol{v} \in \mathbb{R}^3$ such that $\|\boldsymbol{v}\|^2 = \nu d$.

- $\mathbf{z}$ follows a zero-mean trivariate Gaussian distribution with covariance matrix $\Sigma = \frac{(1-\nu)d}{\kappa+1}(I_3 + \kappa \boldsymbol{\mu}\boldsymbol{\mu}')$; it has a constant orientation $\pm\boldsymbol{\mu}$ for $\kappa \to \infty$. The trivariate Gaussian probability distribution admits a PDF on $\mathbb{R}^3$. Using the Sherman-Morrison-Woodbury identity [Hager 1989] to invert D, it reads:

$$p_{\mathbf{z}}(\boldsymbol{z}; \pm\boldsymbol{\mu}, \kappa, d, \nu) = \frac{\kappa + 1}{\left(2\pi(1-\nu)d\right)^{3/2}} \exp\left\{ -\frac{(\kappa+1)\|\boldsymbol{z}\|^2 - \kappa(\boldsymbol{\mu}'\boldsymbol{z})^2}{2(1-\nu)d} \right\}, \tag{2.24}$$

for any $\boldsymbol{z} \in \mathbb{R}^3$.

- $\mathbf{v}$ and $\mathbf{z}$ are statistically independent.

For large values of κ, this definition of the random vector $\mathbf{w}$ completely separates the directional variability (captured by $\mathbf{v}$) and the variability in amplitude (captured by $\mathbf{z}$). The *mean direction* of the displacements is identified by $+\boldsymbol{\mu}$. The parameter $\kappa \geq 0$ measures the *directional dispersion* (the higher κ, the lower the dispersion), the parameter $d > 0$ measures the *displacement scale* and the parameter $\nu \in [0,1]$ is a *non-Gaussianity index* of the compartment (the closer to zero ν, the more Gaussian the diffusion).

2.3.2.2 Probability Density Function

The PDF of the random displacements $\mathbf{w}$ then amounts to the convolution of a vMF PDF defined on a 2-dimensional manifold of $\mathbb{R}^3$ given by eq. (2.23) and a Gaussian PDF defined on $\mathbb{R}^3$ given by eq. (2.24). The proof of its existence and the details of its derivation are given in appendix A.3.3. It reads:

$$\begin{aligned} p_{\mathbf{w}}(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu) &= C(\kappa,d,\nu)\exp\left\{-\frac{(\kappa+1)w_\perp^2+w_{/\!/}^2}{2(1-\nu)d}\right\} \\ &\times \int_{-1}^{1}\exp\left\{\frac{r_\nu\kappa}{2}t^2+\left(\kappa+\frac{r_\nu w_{/\!/}}{\sqrt{\nu d}}\right)t\right\} I_0\left(\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}w_\perp\sqrt{1-t^2}\right)\mathrm{d}t\,, \end{aligned} \tag{2.25}$$

for any $\boldsymbol{w}\in\mathbb{R}^3$, where:

$$\begin{aligned} r_\nu &:= \frac{\nu}{1-\nu}\,, \\ C(\kappa,d,\nu) &:= \frac{\kappa(\kappa+1)\exp\left\{-\frac{r_\nu(\kappa+1)}{2}\right\}}{2\left(2\pi(1-\nu)d\right)^{3/2}\sinh\kappa}\,, \\ (w_{/\!/},w_\perp) &:= \left(<\boldsymbol{\mu},\boldsymbol{w}>,\sqrt{\|\boldsymbol{w}\|^2-<\boldsymbol{\mu},\boldsymbol{w}>^2}\right)\,, \end{aligned}$$

and I_0 is the 0-th order modified Bessel function [Abramowitz 1972], with the convention that $<\boldsymbol{\mu},\boldsymbol{w}>=\|\boldsymbol{w}\|$ for any $\boldsymbol{w}\in\mathbb{R}^3$ when $\kappa=0$.

The PDF of the distribution of $\mathbf{w}$ naturally involves the two following random variables:

$$\mathrm{w}_{/\!/} := <\boldsymbol{\mu},\mathbf{w}> \text{ and } \mathrm{w}_\perp := \sqrt{\|\mathbf{w}\|^2-<\boldsymbol{\mu},\mathbf{w}>^2}\,. \tag{2.26}$$

They have a straightforward interpretation: $\mathrm{w}_{/\!/}$ is the random displacement along direction $+\boldsymbol{\mu}$ and $\mathrm{w}_\perp$ is the modulus of the random displacement in the plane orthogonal to direction $+\boldsymbol{\mu}$. One can then show (proof given in appendix A.3.4) that the joint PDF of $(\mathrm{w}_{/\!/},\mathrm{w}_\perp)$ is given by:

$$p_{(\mathrm{w}_{/\!/},\mathrm{w}_\perp)}(w_{/\!/},w_\perp) = \frac{\kappa}{2\sinh\kappa}\int_{-1}^{1} e^{\kappa t}p_{\mathrm{n}_t}(w_{/\!/})p_{\mathrm{r}_t}(w_\perp)dt\,, \tag{2.27}$$

where n_t is a univariate Gaussian random vector with mean $\sqrt{\nu d}t$ and variance $(1-\nu)d$ and r_t is a univariate Rice random vector with distance parameter $\sqrt{\nu d(1-t^2)}$ and scale parameter $\sqrt{\frac{(1-\nu)d}{\kappa+1}}$.

2.3.2.3 Moments

The joint PDF allows us to analytically compute all the moments of $\mathrm{w}_{/\!/}$ and $\mathrm{w}_\perp$. We focus on (i) the mean of the distribution, (ii) the second-order non-central moments, which are related to the variance properties of the distribution and are the basis to define diffusivity indices and (iii) the fourth-order non-central moments, which are related to the kurtosis properties of the distribution and are a direct measure of the departure from Gaussianity.

The mean of the distribution of the random vector $\mathbf{w}$ can be easily computed using eq. (2.23). Let P be the unitary matrix that rotates $\boldsymbol{\mu}$ to $\boldsymbol{e}_3 = (0,0,1)'$, i.e., $P\boldsymbol{\mu} = \boldsymbol{e}_3$. We can write:

$$\begin{aligned}
\mathrm{E}[\mathbf{w}] &= \mathrm{E}[\mathbf{v}] + \mathrm{E}[\mathbf{z}] = \mathrm{E}[\mathbf{v}] \\
&= \sqrt{\nu d}\frac{\kappa}{4\pi\sinh\kappa}P'\int_{A_2}\exp\{\kappa\cos\theta\}\begin{pmatrix}\sin\theta\cos\phi\\ \sin\theta\sin\phi\\ \cos\theta\end{pmatrix}\sin\theta\mathrm{d}\theta\mathrm{d}\phi \\
&= \sqrt{\nu d}\frac{\kappa}{2\sinh\kappa}\left(\int_0^{\pi}\cos\theta\exp\{\kappa\cos\theta\}\sin\theta\mathrm{d}\theta\right)P'\boldsymbol{e}_3 \\
&= \sqrt{\nu d}\frac{\kappa}{2\sinh\kappa}\left(\int_{-1}^{1}t\exp\{\kappa t\}\,\mathrm{d}t\right)\boldsymbol{\mu} = \sqrt{\nu d}B_1(\kappa)\boldsymbol{\mu}\ ,
\end{aligned} \tag{2.28}$$

where $\mathrm{E}[\cdot]$ is the expected value operator, $A_2 = [0,\pi]\times[0,2\pi]$ and $B_1(\kappa)$ is defined by eq. (A.19) for $p = 1$.

The expressions of the second and fourth raw moments of the distribution of $\mathrm{w}_{/\!/}$ and $\mathrm{w}_{\perp}$ can be found using the equations of appendix A.3.4.2. Simple derivations yield:

$$\begin{aligned}
\mathrm{E}\left[\mathrm{w}_{/\!/}^2\right] &= d\left[1-2\nu\xi(\kappa)\right] \\
\mathrm{E}\left[\mathrm{w}_{\perp}^2\right] &= 2d\left[\frac{1-\nu}{\kappa+1}+\nu\xi(\kappa)\right] \\
\mathrm{E}\left[\mathrm{w}_{/\!/}^4\right] &= d^2\left[3-2\nu^2-4\nu(3-2\nu)\xi(\kappa)+8\nu^2\zeta(\kappa)\right] \\
\mathrm{E}\left[\mathrm{w}_{\perp}^4\right] &= 8d^2\left[\frac{(1-\nu)^2}{(\kappa+1)^2}+\frac{2\nu(1-\nu)\xi(\kappa)}{\kappa+1}+\nu^2\zeta(\kappa)\right]\ ,
\end{aligned} \tag{2.29}$$

where

$$\xi(\kappa) = \begin{cases}\dfrac{\coth\kappa}{\kappa}-\dfrac{1}{\kappa^2}\,, & \text{for } \kappa>0\,,\\ \dfrac{1}{3}-\dfrac{\kappa^2}{45}\,, & \text{for small } \kappa\,,\end{cases} \quad \text{and} \quad \zeta(\kappa) = \begin{cases}\dfrac{1-3\xi(\kappa)}{\kappa^2}\,, & \text{for } \kappa>0\,,\\ \dfrac{1}{15}-\dfrac{\kappa^2}{315}\,, & \text{for small } \kappa\,.\end{cases} \tag{2.30}$$

2.3.3 Fascicle diffusion model in DDI

Recent findings revealed that water within axons is subject to restriced diffusion where a Gaussian distribution appears inappropriate to model the DIMD [Assaf 2005a, Kunz 2011]. Accordingly, we characterize separately the intra-axonal and extra-axonal spaces. Previous geometry-based MCMs that distinguish intra-axonal and extra-axonal spaces [Assaf 2004, Alexander 2010, Zhang 2012] assume no exchange between these spaces. As a result, the extra-axonal space is confounded with the extra-cellular space, which is the sum of the extra-axonal spaces of each axon. It allows one to simplify the MCM to only one extra-cellular compartment, in which diffusing water has to circumnavigate between the axons, which can be nicely captured by tortuosity models [Stanisz 1997, Alexander 2010, Zhang 2012]. The diffusion within each (intra-)axonal compartment needs then to be modeled according to a non-Gaussian distribution.

We attempt to describe the axonal space a bit differently. We assume that the axonal space is a pseudo-cylinder (i.e., a cylinder whose geometry is defined through the shape of the diffusion) at the heart of which lies the intra-axonal space and in which the extra-axonal space is a fine layer over the intra-axonal space that closely fits the geometry of the latter. Hence, water diffusing in the extra-axonal space mainly follows the orientation of the axon (depending on the orientational dispersion) given by the intra-axonal space so that the extra-axonal diffusion process evolves in a dimensionally reduced space with respect to the intra-axonal diffusion. With such a description of the axonal space, the diffusion within the entire axonal space (intra- and extra-axonal spaces) needs to be modeled by a non-Gaussian distribution to meet the recent findings in the literature. In addition, the assumed geometry of the axonal space is not strictly a cylinder but a pseudo-cylinder, which shall precludes the use of Neuman's model [Neuman 1974] or van Gelderen's model [**?**], valid only for restricted diffusion within a cylinder.

We thus designed a distribution that can characterize the diffusion within such an axonal space, authorizing yet not quantifying exchange between extra- and intra-axonal spaces. The distribution introduced in section 2.3.2 is particularly well suited for this purpose. Indeed, in general, the DIMD will be the sum of two independent displacements: one occurring in the 3-dimensional intra-axonal space (represented by $\mathbf{z}$) with associated PDF given in eq. (2.24) whose diffusivity is weighted by the fraction of occupancy of the corresponding space and one occurring in the 2-dimensional extra-axonal space (represented by $\mathbf{v}$) with associated PDF given in eq. (2.23) whose diffusivity is weighted by the fraction of occupancy of the corresponding space. This yields an analytic expression of the PDF of the DIMD in the axonal space given by eq. (2.25).

The 3-dimensional diffusion occurs in the intra-axonal space and is assumed to follow a Gaussian distribution given by eq. (2.24). The corresponding DT is assumed to be cylindrically symmetric as in [Alexander 2010, Zhang 2012]. The 2-dimensional diffusion in the extra-axonal space is however mismodeled by eq. (2.23): indeed, it is not an antipodally symmetric PDF and, consequently, neither is the resulting PDF given by eq. (2.25). In the following, we thus introduce an antipodally symmetric 3-dimensional distribution, based on the PDF given by eq. (2.25), which we use to model the diffusion within our proposed representation of the axonal space (Section 2.3.3.1). We derive its expected value (Section 2.3.3.2), its second-order moments which are related to the concept of diffusivity (Section 2.3.3.3) and its fourth-order moments which are related to the kurtosis properties of the distribution (Section 2.3.3.4). We propose a clinically plausible constraint for the diffusion within fascicle (Section 2.3.3.5). We show how the classic FA and MD and other scalar metrics vary with the parameters of our model (Section 2.3.3.6). Finally, we analytically derive the expression of the SID in terms of the parameters of the proposed distribution that characterizes the diffusion in the fascicle (Section 2.3.3.7).

2.3.3.1 Definition and Probability Density Function

Let $\pm\boldsymbol{\mu}$ be the orientation of the axon and let $\mathbf{w}^+$ and $\mathbf{w}^-$ be the 3-dimensional random displacements along the directions $+\boldsymbol{\mu}$ and $-\boldsymbol{\mu}$, respectively. We assume that the PDFs of the random vectors $\mathbf{w}^+$ and $\mathbf{w}^-$ are given by eq. (2.25) with mean directions $+\boldsymbol{\mu}$ and $-\boldsymbol{\mu}$ respectively, and equal remaining parameters (κ, d and ν). We define the PDF of the DIMD $\mathbf{y}$ along the axon orientation $\pm\boldsymbol{\mu}$ as an equally weighted mixture of the two PDFs of $\mathbf{w}^+$ and $\mathbf{w}^-$, as proposed by eq. (2.22). It reads:

$$\begin{aligned} p_{\mathbf{y}}(\boldsymbol{y};\pm\boldsymbol{\mu},\kappa,d,\nu) &= C(\kappa,d,\nu)\exp\left\{-\frac{(\kappa+1)y_\perp^2+y_{/\!/}^2}{2(1-\nu)d}\right\} \\ &\times \int_{-1}^{1}\exp\left\{\frac{r_\nu\kappa}{2}t^2+\kappa t\right\}\cosh\left(\frac{r_\nu y_{/\!/}t}{\sqrt{\nu d}}\right) I_0\left(\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}y_\perp\sqrt{1-t^2}\right)\mathrm{d}t\,, \end{aligned} \tag{2.31}$$

for any $\boldsymbol{y}\in\mathbb{R}^3$, where:

$$\begin{aligned} r_\nu &:= \frac{\nu}{1-\nu}\,, \\ C(\kappa,d,\nu) &:= \frac{\kappa(\kappa+1)\exp\left\{-\frac{r_\nu(\kappa+1)}{2}\right\}}{2\left(2\pi(1-\nu)d\right)^{3/2}\sinh\kappa}\,, \\ (y_{/\!/},y_\perp) &:= \left(<\boldsymbol{\mu},\boldsymbol{y}>,\sqrt{\|\boldsymbol{y}\|^2-<\boldsymbol{\mu},\boldsymbol{y}>^2}\right)\,, \end{aligned}$$

and I_0 is the 0-th order modified Bessel function [Abramowitz 1972], with the convention that $<\boldsymbol{\mu},\boldsymbol{y}>=\|\boldsymbol{y}\|$ for any $\boldsymbol{y}\in\mathbb{R}^3$ when $\kappa=0$.

In this context, the parameters of the resulting distribution have an intuitive interpretation:

- $\pm\boldsymbol{\mu}$ is the *orientation* of the axon,
- κ is an *orientation dispersion index*, which quantifies the concentration of water molecules around the orientation of the axon,
- d is the *free diffusivity* along the axon,
- ν is the *fraction of occupancy of the extra-axonal space*.

In a given voxel, there are multiple axons that are aligned with the same orientation, forming a bundle of axons hereafter referred to as a **fascicle**. If we assume:

- no exchange between the different axons composing the fascicle (which does not prevent water inside an axon from exchanging between intra- and extra-axonal spaces),

- equal contribution of each axon to the diffusion in the fascicle,

- identical distributions of the DIMD in each axon composing the fascicle,

the PDF of the DIMD within the fascicle is then simply an equally weighted sum of many PDFs given by eq. (2.31) that share the same parameters. The resulting PDF is thus given by eq. (2.31) and, in the following, we shall confound the distribution of the DIMD within an axon with the distribution of the DIMD within a fascicle and only use the fascicle terminology.

The PDF given by eq. (2.31) naturally involves the two following random variables:

$$\mathrm{y}_{/\!/} := < \boldsymbol{\mu}, \mathbf{y} > \text{ and } \mathrm{y}_{\perp} := \sqrt{\|\mathbf{y}\|^2 - < \boldsymbol{\mu}, \mathbf{y} >^2}\,, \tag{2.32}$$

which are of primary importance since they provide an insight into the diffusion parallel ($\mathrm{y}_{/\!/}$) and perpendicular ($\mathrm{y}_{\perp}$) to the orientation of the fascicle.

2.3.3.2 Fascicle apparent mean displacements

The means of the distribution of $\mathbf{w}^+$ and $\mathbf{w}^-$ can be easily derived from eq. (2.28) and amount to $\mathrm{E}[\mathbf{w}^+] = +\sqrt{\nu d} B_1(\kappa)\boldsymbol{\mu}$ and $\mathrm{E}[\mathbf{w}^-] = -\sqrt{\nu d} B_1(\kappa)\boldsymbol{\mu}$, respectively. Consequently, the mean of the distribution of the DIMD $\mathbf{y}$ in the fascicle reads:

$$\mathrm{E}[\mathbf{y}] = \frac{\mathrm{E}[\mathbf{w}^+] + \mathrm{E}[\mathbf{w}^-]}{2} = \mathbf{0}\,, \tag{2.33}$$

which implies, in particular, that

$$\mathrm{E}[< \boldsymbol{c}, \mathbf{y} >] = < \boldsymbol{c}, \mathrm{E}[\mathbf{y}] > = 0\,,\ \forall \boldsymbol{c} \in \mathbb{R}^3\,. \tag{2.34}$$

2.3.3.3 Fascicle apparent diffusivities

The apparent diffusivity, also commonly termed ADC, in a given direction $\boldsymbol{c} \in \mathbb{R}^3$, $\|\boldsymbol{c}\| = 1$, is defined as

$$D(\boldsymbol{c}) = \mathrm{V}[< \boldsymbol{c}, \mathbf{y} >] \overset{eq.\ (2.34)}{=} \mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^2\right], \tag{2.35}$$

where $\mathrm{V}[\cdot]$ is the variance operator. Generally, three diffusivities are of particular interest, namely the MD, the parallel diffusivity $D_{/\!/}$ and the perpendicular diffusivity $D_{\perp}$. They are defined as follows [Jensen 2010]:

$$\begin{aligned} D_{/\!/} &= D(\pm\boldsymbol{\mu})\,, \\ D_{\perp} &= \frac{1}{2\pi} \int_{\boldsymbol{c} \in (\mathrm{span}\{\pm\boldsymbol{\mu}\})^{\perp} \cap \mathbb{S}^2} D(\boldsymbol{c}) dS(\boldsymbol{c})\,, \\ \mathrm{MD} &= \frac{1}{4\pi} \int_{\boldsymbol{c} \in \mathbb{R}^3 \cap \mathbb{S}^2} D(\boldsymbol{c}) dS(\boldsymbol{c})\,, \end{aligned} \tag{2.36}$$

where dS denotes integration with respect to the spherical measure and $\mathbb{S}^2$ is the 2-sphere of $\mathbb{R}^3$. In other words, the MD is the spherical mean of $D(\boldsymbol{c})$ over the entire 3-dimensional space $\mathbb{R}^3$, $D_{/\!/}$ is the diffusivity along orientation $\pm\boldsymbol{\mu}$ and $D_{\perp}$ is the spherical mean of $D(\boldsymbol{c})$ over the plane $(\mathrm{span}\{\pm\boldsymbol{\mu}\})^{\perp}$, orthogonal to the orientation $\pm\boldsymbol{\mu}$. It can be shown (see the proof in appendix B.2) that these three diffusivities are related by the following equation:

$$\mathrm{MD} = \frac{D_{/\!/} + 2D_{\perp}}{3}\,. \tag{2.37}$$

In order to compute $D_{/\!/}$ and $D_{\perp}$, the random vector $\mathbf{y}$ can be conveniently rewritten as follows:

$$\mathbf{y} = \mathrm{y}_{/\!/}\boldsymbol{\mu} + \mathrm{y}_{\perp}\mathbf{u}\,,\ \text{with } \mathbf{u} = \frac{\mathbf{y} - \mathrm{y}_{/\!/}\boldsymbol{\mu}}{\mathrm{y}_{\perp}} \in (\mathrm{span}\{\boldsymbol{\mu}\})^{\perp} \cap \mathbb{S}^2\,, \tag{2.38}$$

where $y_{/\!/}$ and $y_{\perp}$ are given in eq. (2.32). The parallel diffusivity then reads:

$$D_{/\!/} = \mathrm{E}\left[< \boldsymbol{\mu}, \mathbf{y} >^2\right] = \mathrm{E}\left[y_{/\!/}^2\right] ,$$

and the following equation holds for any $\boldsymbol{c} \in (\mathrm{span}\{\pm\boldsymbol{\mu}\})^{\perp} \cap \mathbb{S}^2$:

$$D(\boldsymbol{c}) = \mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^2\right] = \mathrm{E}\left[y_{\perp}^2 < \boldsymbol{c}, \mathbf{u} >^2\right] . \tag{2.39}$$

Since the two smallest eigenvalues of the covariance matrix of the distribution of $\mathbf{z}$ are equal (see eq. (2.24)), the distribution of $\mathbf{y}$ satisfies the cylindrical symmetry property, i.e., the distribution of $\mathbf{u}$ is uniform on the unit 1-sphere, and thus, $y_{\perp}$ and $\mathbf{u}$ are statistically independent. Consequently, eq. (2.39) simplifies to:

$$D(\boldsymbol{c}) = \mathrm{E}\left[y_{\perp}^2\right]\mathrm{E}\left[< \boldsymbol{c}, \mathbf{u} >^2\right] .$$

Furthermore, since $\mathbf{u}$ is uniform on the unit 1-sphere, the equality $< \boldsymbol{c}, \mathbf{u} > \overset{d}{=} \cos\theta$, with $\theta \sim \mathcal{U}[0, 2\pi]$, holds in distribution and we can thus write:

$$\mathrm{E}\left[< \boldsymbol{c}, \mathbf{u} >^2\right] = \frac{1}{2\pi}\int_0^{2\pi} \cos^2\theta \mathrm{d}\theta = \frac{1}{2} . \tag{2.40}$$

For any $\boldsymbol{c} \in (\mathrm{span}\{\boldsymbol{\mu}\})^{\perp} \cap \mathbb{S}^2$, we thus obtain that $D(\boldsymbol{c}) = \frac{1}{2}\mathrm{E}\left[y_{\perp}^2\right]$ which does not depend on $\boldsymbol{c}$. Therefore, the perpendicular diffusivity reads:

$$D_{\perp} = \frac{1}{2}\mathrm{E}\left[y_{\perp}^2\right] .$$

Since the distribution of $\mathbf{y}$ is a mixture distribution, we have, for any function $h(\cdot)$ such that both $\mathrm{E}\left[h(\mathbf{w}^+)\right]$ and $\mathrm{E}\left[h(\mathbf{w}^-)\right]$ exist:

$$2\mathrm{E}\left[h(\mathbf{y})\right] = \mathrm{E}\left[h(\mathbf{w}^+)\right] + \mathrm{E}\left[h(\mathbf{w}^-)\right] . \tag{2.41}$$

Applying eq. (2.41) with the functions $\boldsymbol{x} \mapsto < \boldsymbol{\mu}, \boldsymbol{x} >$ and $\boldsymbol{x} \mapsto \sqrt{\|\boldsymbol{x}\|^2 - < \boldsymbol{\mu}, \boldsymbol{x} >^2}$ and observing that the non-central moments of $w_{/\!/}^+$ and $w_{/\!/}^-$ (respectively, $w_{\perp}^+$ and $w_{\perp}^-$) are equal, we obtain the following expressions for the diffusivities:

$$D_{/\!/} = \mathrm{E}\left[w_{/\!/}^2\right] \text{ and } D_{\perp} = \frac{1}{2}\mathrm{E}\left[w_{\perp}^2\right] ,$$

where $\mathrm{E}\left[w_{/\!/}^2\right] = \mathrm{E}\left[(w_{/\!/}^+)^2\right] = \mathrm{E}\left[(w_{/\!/}^-)^2\right]$ and $\mathrm{E}\left[w_{\perp}^2\right] = \mathrm{E}\left[(w_{\perp}^+)^2\right] = \mathrm{E}\left[(w_{\perp}^-)^2\right]$ are given by eq. (2.29). The final expression of the parallel and perpendicular diffusivities are then:

$$\begin{aligned} D_{/\!/}^{\text{fascicle}} &= d\left[1 - 2\nu\xi(\kappa)\right] , \\ D_{\perp}^{\text{fascicle}} &= d\left[\frac{1-\nu}{\kappa+1} + \nu\xi(\kappa)\right] , \end{aligned} \tag{2.42}$$

where the ξ function is defined in eq. (2.30), and thus the MD reads:

$$\mathrm{MD}^{\text{fascicle}} = \frac{d}{3}\left(1 + 2\frac{1-\nu}{\kappa+1}\right) . \tag{2.43}$$

In the same vein as the NODDI model [Zhang 2012], eq. (2.42) expresses the fascicle apparent axial and radial diffusivities as functions of the free diffusivity d along the fascicle orientation, the orientation dispersion index κ and the extra-axonal fraction of occupancy ν in a physically plausible way and not as independent free parameters as in the multi-tensor [Tuch 2002, Scherrer 2012], ball-and-stick [Behrens 2003, Jbabdi 2012] and CHARMED [Assaf 2005a] models. In particular, the FA of the fascicle [Basser 1996a] does not simply depend on the radial/axial diffusivity ratio but is rather determined by the orientation dispersion and the extra-axonal fraction of occupancy as follows:

$$\mathrm{FA}^{\,\mathrm{fascicle}} = \frac{1 - R(\nu,\kappa)}{\sqrt{1 + 2R^2(\nu,\kappa)}}\,, \text{ with } R(\nu,\kappa) := \frac{\frac{1-\nu}{\kappa+1} + \nu\xi(\kappa)}{1 - 2\nu\xi(\kappa)}\,. \tag{2.44}$$

It has been pointed out in [Zhang 2012] that κ is not an intuitive orientation dispersion index for two reasons: (i) it does not admit an upper bound, which makes difficult to set a good contrast for an optimal visualization of this index and (ii) it is in fact inversely proportional to the orientation dispersion. Consequently, we adopt the Orientation Dispersion (OD) index proposed in the same work and reported in this manuscript by eq. (2.19).

2.3.3.4 Fascicle apparent kurtoses

Similarly to the apparent diffusivities, we can define the apparent kurtosis in a given orientation $\boldsymbol{c} \in \mathbb{R}^3 \bigcap \mathbb{S}^2$ as the kurtosis of the distribution of the random variable $< \boldsymbol{c}, \mathbf{y} >$:

$$K(\boldsymbol{c}) = \mathrm{K}\left[< \boldsymbol{c}, \mathbf{y} >\right] \overset{eq.\ (2.34)}{=} \frac{\mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^4\right]}{\mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^2\right]^2} - 3\,, \tag{2.45}$$

where $\mathrm{K}[\cdot]$ is the kurtosis operator. In particular, the MK and the axial ($K_{/\!/}$) and radial ($K_\perp$) kurtoses can be specified. In the context of dMRI, these quantities were first defined (and studied) in [Hui 2008, Wu 2010] and later generalized into DKI as follows [Jensen 2010]:

$$\begin{aligned} K_{/\!/} &:= K(\pm\boldsymbol{\mu})\,, \\ K_\perp &:= \frac{1}{2\pi} \int_{\boldsymbol{c} \in (\mathrm{span}\{\pm\boldsymbol{\mu}\})^\perp \bigcap \mathbb{S}^2} K(\boldsymbol{c}) dS(\boldsymbol{c})\,, \\ \mathrm{MK} &:= \frac{1}{4\pi} \int_{\boldsymbol{c} \in \mathbb{R}^3 \bigcap \mathbb{S}^2} K(\boldsymbol{c}) dS(\boldsymbol{c})\,, \end{aligned} \tag{2.46}$$

where dS denotes integration with respect to the spherical measure and $\mathbb{S}^2$ is the 2-sphere of $\mathbb{R}^3$. In other words, the MK is the spherical mean of $K(\boldsymbol{c})$ over $\mathbb{R}^3$, $K_{/\!/}$ is the kurtosis along orientation $\pm\boldsymbol{\mu}$ and $K_\perp$ is the spherical mean of $K(\boldsymbol{c})$ over the plane $(\mathrm{span}\{\pm\boldsymbol{\mu}\})^\perp$, orthogonal to the orientation $\pm\boldsymbol{\mu}$. Note that, originally, the radial kurtosis was defined in a different way in [Hui 2008]. By analogy with the diffusivities, they extract an orthonormal basis of $(\mathrm{span}\{\pm\boldsymbol{\mu}\})^\perp$, compute the apparent kurtoses in the two extracted orientations and average them to obtain

a radial kurtosis. Even if it is less general than the one in [Jensen 2010] since it does not average over all the orientations in $(\text{span}\{\pm\boldsymbol{\mu}\})^{\perp}$ but only over the basis orientations, the two definitions are in fact equivalent under the assumption of cylindrical symmetry that we adopted. In particular, the definition introduced by [Hui 2008] allows the definition of a kurtosis FA (FA_K), in a very similar way to the classic FA:

$$\text{FA}_K = \frac{|K_{/\!/} - K_{\perp}|}{\sqrt{K_{/\!/}^2 + 2K_{\perp}^2}}, \tag{2.47}$$

which **is not defined** for Gaussian diffusion profiles for which $K_{/\!/} = K_{\perp} = 0$.

In order to compute $K_{/\!/}$ and $K_{\perp}$, we use the decomposition of the random vector $\mathbf{y}$ proposed in eq. (2.38). The parallel kurtosis then reads:

$$K_{/\!/} = \frac{\mathrm{E}\left[< \boldsymbol{\mu}, \mathbf{y} >^4\right]}{\mathrm{E}\left[< \boldsymbol{\mu}, \mathbf{y} >^2\right]^2} - 3 = \frac{\mathrm{E}\left[\mathrm{y}_{/\!/}^4\right]}{\mathrm{E}\left[\mathrm{y}_{/\!/}^2\right]^2} - 3,$$

and the following equation holds for any $\boldsymbol{c} \in (\text{span}\{\pm\boldsymbol{\mu}\})^{\perp} \cap \mathbb{S}^2$:

$$K(\boldsymbol{c}) = \frac{\mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^4\right]}{\mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^2\right]^2} - 3 = \frac{\mathrm{E}\left[\mathrm{y}_{\perp}^4 < \boldsymbol{c}, \mathbf{u} >^4\right]}{\mathrm{E}\left[\mathrm{y}_{\perp}^2 < \boldsymbol{c}, \mathbf{u} >^2\right]^2} - 3. \tag{2.48}$$

Since the distribution of $\mathbf{y}$ satisfies the cylindrical symmetry property, the distribution of $\mathbf{u}$ is uniform on the unit 1-sphere, and thus, $\mathrm{y}_{\perp}$ and $\mathbf{u}$ are statistically independent. Consequently, eq. (2.39) simplifies to:

$$K(\boldsymbol{c}) = \frac{\mathrm{E}\left[\mathrm{y}_{\perp}^4\right]\mathrm{E}\left[< \boldsymbol{c}, \mathbf{u} >^4\right]}{\mathrm{E}\left[\mathrm{y}_{\perp}^2\right]^2\mathrm{E}\left[< \boldsymbol{c}, \mathbf{u} >^2\right]^2} - 3.$$

Furthermore, since $\mathbf{u}$ is uniform on the unit 1-sphere, the equality $< \boldsymbol{c}, \mathbf{u} > \overset{d}{=} \cos\theta$, with $\theta \sim \mathcal{U}[0, 2\pi]$, holds in distribution and we can thus write:

$$\mathrm{E}\left[< \boldsymbol{c}, \mathbf{u} >^4\right] = \frac{1}{2\pi}\int_0^{2\pi} \cos^4\theta \mathrm{d}\theta = \frac{3}{8}.$$

Using this equation in addition with eq. (2.40), we obtain that:

$$K(\boldsymbol{c}) = \frac{3}{2}\frac{\mathrm{E}\left[\mathrm{y}_{\perp}^4\right]}{\mathrm{E}\left[\mathrm{y}_{\perp}^2\right]^2} - 3, \forall \boldsymbol{c} \in (\text{span}\{\pm\boldsymbol{\mu}\})^{\perp} \cap \mathbb{S}^2,$$

which does not depend on $\boldsymbol{c}$. Therefore, the perpendicular kurtosis reads:

$$K_{\perp} = \frac{3}{2}\frac{\mathrm{E}\left[\mathrm{y}_{\perp}^4\right]}{\mathrm{E}\left[\mathrm{y}_{\perp}^2\right]^2} - 3.$$

Applying eq. (2.41) with the functions $\boldsymbol{x} \mapsto < \boldsymbol{\mu}, \boldsymbol{x} >$ and $\boldsymbol{x} \mapsto \sqrt{\|\boldsymbol{x}\|^2 - < \boldsymbol{\mu}, \boldsymbol{x} >^2}$ and observing that the non-central moments of $\mathrm{w}_{/\!/}^+$ and $\mathrm{w}_{/\!/}^-$ (respectively, $\mathrm{w}_{\perp}^+$ and $\mathrm{w}_{\perp}^-$) are equal, we obtain the following expressions for the kurtoses:

$$K_{/\!/} = \frac{\mathrm{E}\left[\mathrm{w}_{/\!/}^4\right]}{\mathrm{E}\left[\mathrm{w}_{/\!/}^2\right]^2} - 3 \text{ and } K_{\perp} = \frac{3}{2}\frac{\mathrm{E}\left[\mathrm{w}_{\perp}^4\right]}{\mathrm{E}\left[\mathrm{w}_{\perp}^2\right]^2} - 3,$$

where $\mathrm{E}\left[\mathrm{w}_{/\!/}^{j}\right] = \mathrm{E}\left[(\mathrm{w}_{/\!/}^{+})^{j}\right] = \mathrm{E}\left[(\mathrm{w}_{/\!/}^{-})^{j}\right]$ and $\mathrm{E}\left[\mathrm{w}_{\perp}^{j}\right] = \mathrm{E}\left[(\mathrm{w}_{\perp}^{+})^{j}\right] = \mathrm{E}\left[(\mathrm{w}_{\perp}^{-})^{j}\right]$, $j = \{1,2\}$, are given by eq. (2.29). The final expression of the parallel and perpendicular kurtoses are then:

$$\begin{aligned} K_{/\!/}^{\text{fascicle}} &= -\frac{2\nu^2\left(1-4\xi(\kappa)-4\zeta(\kappa)+6\xi^2(\kappa)\right)}{\left(1-2\nu\xi(\kappa)\right)^2} \leq 0\,, \\ K_{\perp}^{\text{fascicle}} &= -\frac{3\nu^2\left(\xi^2(\kappa)-\zeta(\kappa)\right)}{\left(\frac{1-\nu}{\kappa+1}+\nu\xi(\kappa)\right)^2} \leq 0\,, \end{aligned} \tag{2.49}$$

where the ξ and ζ functions are defined in eq. (2.30). In particular, for isotropic diffusion ($\kappa = 0$), we have that:

$$K_{/\!/}^{\text{fascicle}} = K_{\perp}^{\text{fascicle}} = -\frac{6}{5}\left(\frac{\nu}{3-2\nu}\right)^2 .$$

2.3.3.5 Physically plausible constraint

One advantage of our proposed fascicle diffusion model is that analytical expressions of apparent parallel and perpendicular diffusivities and kurtoses are available, enabling to set physically plausible constraints on some parameters that will help to fit the model with clinical data. Any diffusion model, in the limit of its assumptions, is able to accurately quantify the fascicle apparent diffusivities in the Corpus Callosum (CC) and at the basis of the Cortico-Spinal Tract (CST), which are highly orientationally homogeneous regions that contain a single fascicle. A physically plausible constraint is a constraint that has been validated experimentally by several different diffusion models. In particular, the value of the fascicle apparent perpendicular diffusivity cannot be set to some previously estimated value of the literature as different models yield different values (for instance, DTI and DKI do not agree on its value). On the other hand, considering a non-null kurtosis in the fascicle diffusion model does not affect the estimated value of the fascicle apparent parallel diffusivity: [Pierpaoli 1996b] report $D_{/\!/} = 1.7 \times 10^{-3}$ mm^2/s using DTI and [Veraart 2011b] report $D_{/\!/} = 1.71 \times 10^{-3}$ mm^2/s using DKI. We thus set the following constraint on the parameters of our fascicle diffusion model:

$$D_{/\!/}^{\text{fascicle}} = d\left[1-2\nu\xi(\kappa)\right] = 1.71 \times 10^{-3}\ \text{mm}^2/\text{s}\,. \tag{2.50}$$

2.3.3.6 Illustrations

Figures 2.1 and 2.2 illustrate how the FA given by eq. (2.44) and MD given by eq. (2.43) respectively vary with the orientation dispersion index OD and the extra-axonal fraction of occupancy ν. In fig. 2.2, we also applied the constraint given by eq. (2.50). No matter the value of the extra-axonal fraction of occupancy ν, we observe that our model will inflate the MD and deflate the FA in orientationally heterogeneous regions, i.e. in regions where the OD is important. Interestingly, non-Gaussian diffusion will tend to accelerate the FA deflation and MD inflation. In other words, our model is able to explain an abnormal drop in FA (or, similarly,

an abnormal rise in MD) by a number of situations that lie in-between the two following scenari:

- a non-Gaussian diffusion profile resulting from heterogeneities in orientation, with Gaussian diffusion profiles along each orientation of diffusion,
- a non-Gaussian diffusion profile resulting from a single orientation of diffusion along which the diffusion profile is actually non-Gaussian.

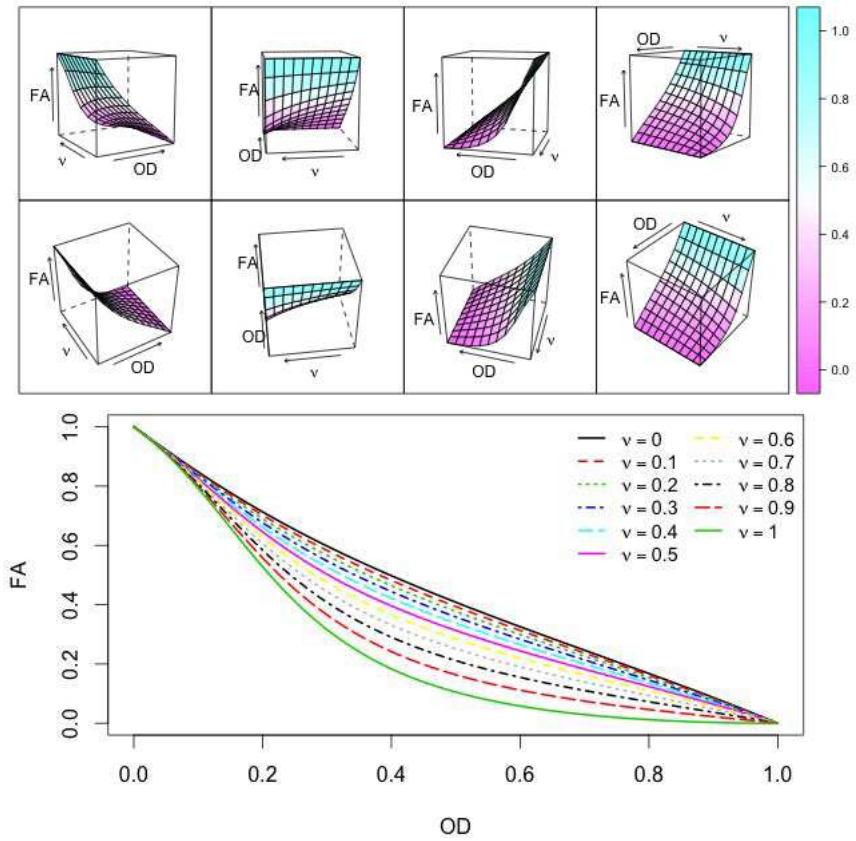


Figure 2.1: **Fascicle Fractional Anisotropy.** Variations of the FA with the OD and the extra-axonal fraction ν from eq. (2.44) according to the proposed DDI model.

Figures 2.3 to 2.5 illustrate how $K_{/\!/}$ and $K_{\perp}$ given by eq. (2.49) and FA_K given by eq. (2.47) respectively vary with the orientation dispersion index OD and the extra-axonal fraction of occupancy ν. No matter the ν value, we observe that our model will inflate $K_{/\!/}$ in orientationally heterogeneous regions, i.e. in regions where the OD is important. Interestingly, the more non-Gaussian the diffusion, the lower $K_{/\!/}$, independently from the OD. Also, except for the extreme values of the extra-axonal fraction of occupancy ν (when $\nu = 0$, the diffusion profile is Gaussian whereas when $\nu = 1$, the diffusion profile is bimodal), $K_{\perp}$ is not a monotonic function of the OD. For low OD, it undergoes a decrease which is accelerated and more important as ν increases. Then, for a specific OD value, which is proportional to ν, it reaches a

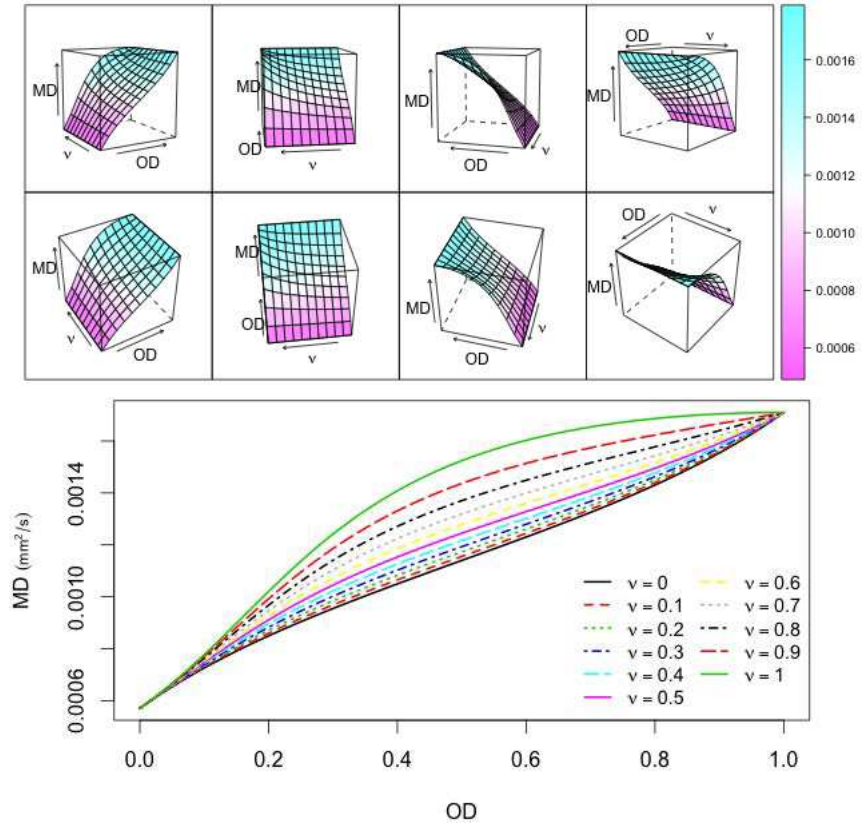


Figure 2.2: **Fascicle Mean Diffusivity.** Variations of the MD with the OD and the extra-axonal fraction ν from eqs. (2.43) and (2.50) according to the proposed DDI model.

minimum and for high OD, it increases at approximately constant speed no matter the ν value. As a consequence, FA_K undergoes three different phases of variation with OD. For low OD, it decreases from 1 to 0 rapidly, with a speed that is inversely proportional to the ν value. For moderate OD, it increases (at constant speed for all ν values) and then it decreases again to 0.

2.3.3.7 Associated SID

Using q-space theory, the SID measured by dMRI experiments is related to the distribution of the DIMD $\mathbf{y}$ via eqs. (1.9) and (1.10). Assuming that the PDF of the distribution of the DIMD $\mathbf{y}$ is given by eq. (2.31) yields the following expression of the SID (see appendix B.1):

$$\begin{aligned} A_{jk} &= \exp\left\{-\frac{b_k(1-\nu)d}{\kappa+1}\left(1+\kappa<\boldsymbol{\mu},\boldsymbol{g}_j>^2\right)\right\} \\ &\times \frac{\kappa}{\sinh\kappa}\begin{cases} \left|\operatorname{sinc}\left(\sqrt{2b_k\nu d-\kappa^2}\right)\right| & , \text{ if } 2b_k\nu d \geq \kappa^2 \text{ and } \boldsymbol{g}_j \perp \boldsymbol{\mu}\,, \\ \dfrac{|\alpha\sinh\alpha\cos\beta+\beta\cosh\alpha\sin\beta|}{\alpha^2+\beta^2} & , \text{ otherwise,} \end{cases} \end{aligned} \tag{2.51}$$

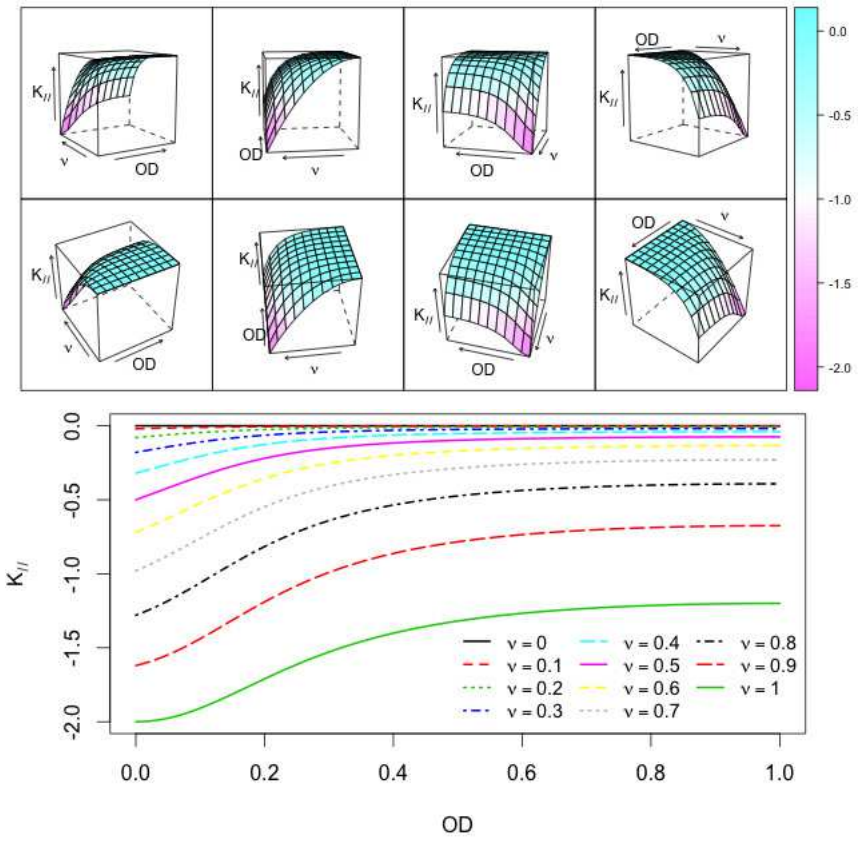


Figure 2.3: **Fascicle axial kurtosis.** Variations of the axial kurtosis with the OD and the extra-axonal fraction ν from eq. (2.49) according to the proposed DDI model.

where:

$$\alpha = \sqrt{\frac{\mathrm{Re}\,(z) + |z|}{2}}\,,\ \beta = \frac{\mathrm{Im}\,(z)}{\sqrt{2(\mathrm{Re}\,(z) + |z|)}} \text{ and } z = \kappa^2 - 2b_k\nu d + 2i\kappa\sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j > \,.$$

In order to measure exclusively the diffusion along the orientation $\pm\boldsymbol{\mu}$ of the fascicle, one can apply a DSG with $\boldsymbol{g}_j = \boldsymbol{\mu}$. This yields the following expression of the SID:

$$A_{jk} = \exp\{-b_k(1-\nu)d\}\left|\cos\sqrt{2b_k\nu d}\right| \frac{\left|1 + \sqrt{2b_k\nu d}\frac{\coth\kappa}{\kappa}\tan\sqrt{2b_k\nu d}\right|}{1 + \frac{2b_k\nu d}{\kappa^2}}\,. \tag{2.52}$$

In order to measure exclusively the diffusion perpendicular to the orientation $\pm\boldsymbol{\mu}$ of the fascicle, one can apply a DSG in a direction perpendicular to $\pm\boldsymbol{\mu}$. For low b-values ($b_k < \kappa^2/(2\nu d)$), this yields the following expression of the SID:

$$A_{jk} = \exp\left\{-\frac{b_k(1-\nu)d}{\kappa+1}\right\}\left(1 - \frac{2b_k\nu d}{\kappa^2}\right)^{-1/2} \frac{\sinh\left(\kappa\sqrt{1 - \frac{2b_k\nu d}{\kappa^2}}\right)}{\sinh\kappa}\,, \tag{2.53}$$

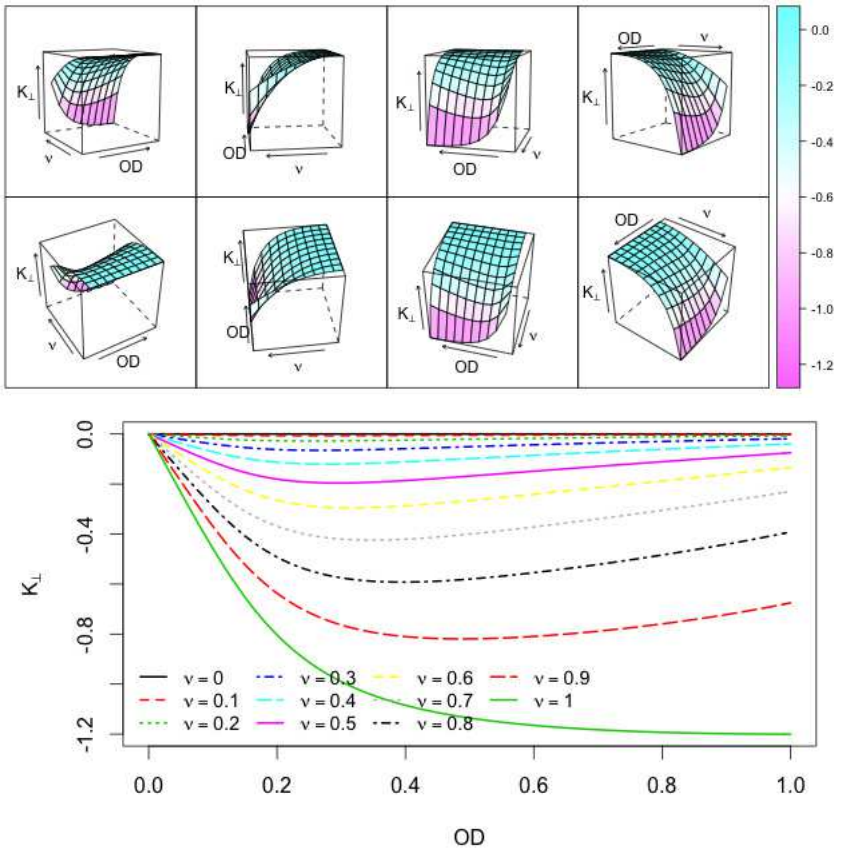


Figure 2.4: **Fascicle radial kurtosis.** Variations of the radial kurtosis with the OD and the extra-axonal fraction ν from eq. (2.49) according to the proposed DDI model.

and, for high b-values ($b_k \geq \kappa^2/(2\nu d)$):

$$A_{jk} = \exp\left\{-\frac{b_k(1-\nu)d}{\kappa+1}\right\}\frac{\kappa}{\sinh\kappa}\left|\operatorname{sinc}\left(\sqrt{2b_k\nu d-\kappa^2}\right)\right|. \tag{2.54}$$

The detailed derivation of the three previous expressions are given in appendix B.1.1.

As shown in appendix B.1.2, eq. (2.51) takes a simple form under the assumption of low orientational dispersion ($\kappa \to \infty$), which leads to an anisotropically restricted compartment whose SID reads:

$$A_{jk} = \exp\left\{-b_k(1-\nu)d < \boldsymbol{\mu}, \boldsymbol{g}_j >^2\right\}\left|\cos\left(\sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j >\right)\right|. \tag{2.55}$$

It also takes a simple form under the assumption of high orientational dispersion ($\kappa \to 0$), as detailed in appendix B.1.3, which leads to an isotropically restricted compartment whose SID reads:

$$A_{jk} = \exp\left\{-b_k(1-\nu)d\right\}\left|\operatorname{sinc}\left(\sqrt{2b_k\nu d}\right)\right|. \tag{2.56}$$

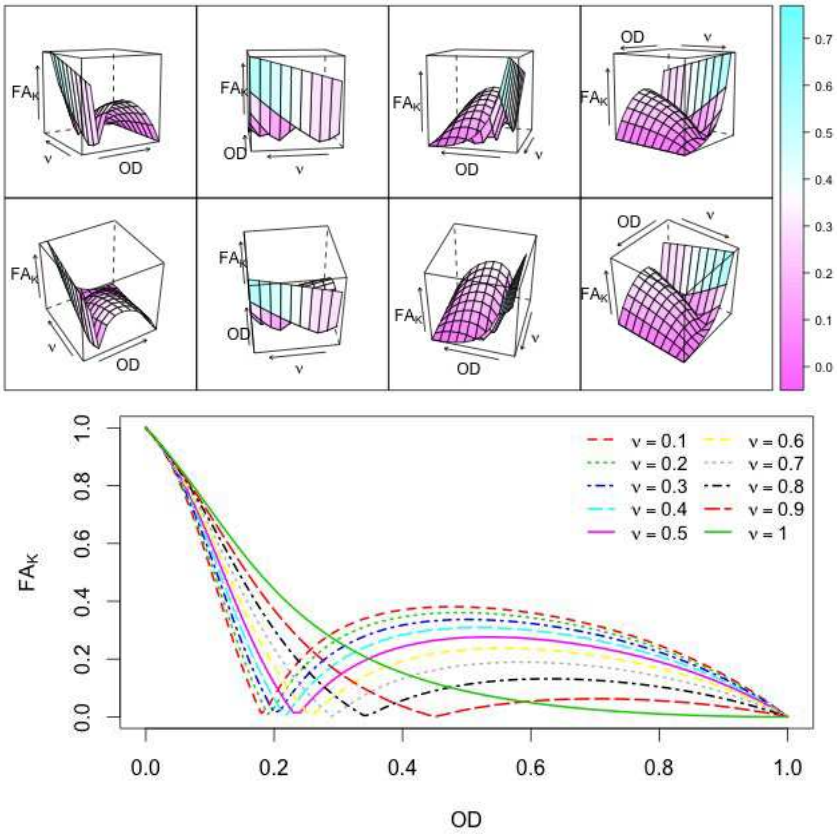


Figure 2.5: **Fascicle kurtosis Fractional Anisotropy.** Variations of the kurtosis FA with the OD and the extra-axonal fraction ν from eq. (2.47) according to the proposed DDI model.

2.3.4 Voxel diffusion model in DDI

The model presented in section 2.3.3 allows one to characterize anisotropic diffusion within a single fascicle or isotropically restricted diffusion (for instance, within glial cells) based on a simplified yet physically plausible geometry of the tissue (axons or glial cells). In order to account for the presence of multiple fascicles with different orientations, which was first brought to light by [Tuch 1999, Tuch 2002], MCMs are particularly well suited. They assume that the voxel contains several populations of water molecules: some of them are within fascicles, others within glial cells and the remainders are free, outside of these restricted areas. The DIMD of each population is characterized by a PDF and the global PDF of the distribution of the DIMD in the voxel is given as a mixture of all the individual PDFs, weighted according to the fraction of occupancy of the corresponding population in the voxel.

In section 2.3.4.1, we emcompass our fascicle diffusion model into an unconstrained MCM, referred to as the full DDI MCM. In section 2.3.4.2, we show that it is possible to derive any scalar maps at the fascicle or voxel level, leading to corrected maps of, for example, FA or MD. In section 2.3.4.3, we apply a number of physi-

cally plausible constraints to the full DDI MCM to obtain a constrained DDI MCM feasible in clinics. In section 2.3.4.4, we express the SID in terms of the independent parameters of the constrained DDI MCM. Finally, we show in section 2.3.4.5 that our model boils down to some well known models under some assumptions.

2.3.4.1 The full DDI MCM

We assume that in a given voxel, water molecules are divided into $M+2$ populations. A first population, characterized by its DIMD $\mathbf{y}_0$, contains water in the CSF, subject to isotropically unrestricted diffusion, with fraction of occupancy w_0. A second population, characterized by its DIMD $\mathbf{y}_s$, contains water trapped into glial cells, subject to isotropically restricted diffusion, with fraction of occupancy w_s. The remaining M populations, characterized by the DIMD $\mathbf{y}_i$, $i \in [\![1, M]\!]$, contain water trapped within axons distributed in M distinct fascicles, subject to anisotropic diffusion, with associated fractions of occupancy w_i, $i \in [\![1, M]\!]$.

In order to describe the diffusion in such an heterogenous medium, the following model is adopted:

- The DIMD $\mathbf{y}_0$ are assumed to follow an isotropically centered Gaussian distribution with DT $D_0 = d_0 I_3$, whose PDF reads:

$$p_{\mathbf{y}_0}(\boldsymbol{y}|d_0) = (2\pi d_0)^{-3/2} \exp\left\{-\frac{\|\boldsymbol{y}_0\|^2}{2d_0}\right\} . \tag{2.57}$$

- The DIMD $\mathbf{y}_s$ are assumed to follow a distribution whose PDF is given by eq. (2.31) with $\kappa = 0$ and thus reads:

$$p_{\mathbf{y}_s}(\boldsymbol{y}|d_s, \nu_s) = (2\pi(1-\nu_s)d_s)^{-3/2} \exp\left\{-\frac{1}{2}\left(\frac{\|\boldsymbol{y}\|^2}{(1-\nu_s)d_s} + r_{\nu_s}\right)\right\} \frac{\sinh\left(\frac{r_{\nu_s}\|\boldsymbol{y}\|}{\sqrt{\nu_s d_s}}\right)}{\frac{r_{\nu_s}\|\boldsymbol{y}\|}{\sqrt{\nu_s d_s}}} . \tag{2.58}$$

- The DIMD $\mathbf{y}_i$ in the i-th fascicle are assumed to follow a distribution whose PDF is given by eq. (2.31) with orientation $\boldsymbol{\mu}_i$ and geometry (κ_i, d_i, ν_i), $i \in [\![1, M]\!]$.
- The fractions of occupancy of the $M+2$ compartments that are assume to compose the voxel sum to one:

$$w_0 + w_s + \sum_{i=1}^{M} w_i = 1 . \tag{2.59}$$

Such a composition in the voxel yields the following PDF that defines the distribution of the DIMD $\mathbf{y}$ at the voxel level (and has thus zero mean):

$$p_{\mathbf{y}}(\boldsymbol{y}) = w_0 p_{\mathbf{y}_0}(\boldsymbol{y}|d_0) + w_s p_{\mathbf{y}_s}(\boldsymbol{y}|d_s, \nu_s) + \sum_{i=1}^{M} w_i p_{\mathbf{y}_i}(\boldsymbol{y}\,|\boldsymbol{\mu}_i, \kappa_i, d_i, \nu_i) . \tag{2.60}$$

2.3.4.2 Voxel-based scalar metrics

Section 2.3.3.3 was dedicated at showing how one can analytically derive the parallel and perpendicular apparent diffusivities from the PDF given in eq. (2.31). The MCM offers the possibility to compute these quantities at the voxel level as follows:

$$\begin{aligned} D_{/\!/}^{\text{voxel}} &= w_0 d_0 + w_s D^{\text{glial cells}} + \sum_{i=1}^{M} w_i D_{/\!/}^{\text{fascicle,i}}, \\ D_{\perp}^{\text{voxel}} &= w_0 d_0 + w_s D^{\text{glial cells}} + \sum_{i=1}^{M} w_i D_{\perp}^{\text{fascicle,i}}, \end{aligned} \tag{2.61}$$

where the quantities $D_{/\!/}^{\text{fascicle,i}}$ and $D_{\perp}^{\text{fascicle,i}}$ are given by eq. (2.42) and the quantity $D^{\text{glial cells}}$ is obtained by letting $\kappa \to 0$ in eq. (2.42), which leads to:

$$D^{\text{glial cells}} = d_s \left(1 - \frac{2\nu_s}{3}\right). \tag{2.62}$$

Equations (2.42) and (2.62) can be inserted into eq. (2.61) to obtain the complete expressions of the parallel and perpendicular apparent diffusivities at the voxel level. Once the parallel and perpendicular diffusivities are known, one can analytically compute all the usual scalar maps proposed by DTI such as the MD given by eq. (1.19), the FA and RA given by eq. (1.20) or the shape coefficients c_ℓ, c_p and c_s given by eq. (1.21), by taking $d_1 = D_{/\!/}$ and $d_2 = d_3 = D_{\perp}$. Note that in our model, we will always have $c_p = 0$ since we assume cylindrical symmetry for the diffusion perpendicular to the fascicle orientation. Our model, like any MCM, is very handy as each of these scalar maps can be computed either at the compartment level using the apparent diffusivities given by eqs. (2.42) and (2.50) or at the voxel level using the apparent diffusivities given by eqs. (2.42), (2.61) and (2.62). In particular, the physically plausible constraint described by eq. (2.50) that we adopted sets the parallel apparent diffusivity at the compartment level $D_{/\!/}^{\text{fascicle}}$ to a fixed value for any fascicle but this does not prevent the parallel apparent diffusivity at the voxel level $D_{/\!/}^{\text{voxel}}$ to be estimated: for instance, $D_{/\!/}^{\text{voxel}}$ will increase towards d_0 as the fraction of occupancy of the free-water compartment w_0 increases. In a word, voxel-based versions of the usual scalar maps obtained by means of an MCM are corrected values that account for tissue environment heterogeneity. Voxel-based kurtoses can be analytically derived as well. Their expressions are however complex and not reported here.

2.3.4.3 The constrained DDI MCM

In order to easily translate the DDI MCM to clinics, we adopt the following four assumptions:

- The diffusivity in the CSF compartment is fixed and set to its value at 37°C: $d_0 = D^{\text{CSF}} = 3.0 \times 10^{-3}$ mm^2/s;
- The work of [Stanisz 1997] suggests that part of the departure from Gaussian DIMD can be explained by the inclusion of a compartment to model diffusion within glial cells, in which water is subject to isotropically restricted diffusion. However, [Alexander 2010] explain that, while postmortem data show this effect important, it is in fact negligible *in vivo*. We thus set $w_s = 0$;
- The parallel diffusivity in all the fascicle compartments is known and set according to prior knowledge using eq. (2.50);
- The OD indices defined through the κ_i's and the extra-axonal fractions of occupancy ν_i are identical for all the fascicle compartments; in other words, we view the OD and extra-axonal occupancy as global quantities at the voxel level; in conjunction with the previous assumption, it boils down to assuming that, in a voxel, the different fascicles can have different orientations but all share the same geometry described by (κ, d, ν).

Under such assumptions, the PDF of the DIMD at the voxel level simplifies to:

$$p_{\mathbf{y}}(\boldsymbol{y}) = w_0 p_{\mathbf{y}_0}(\boldsymbol{y}|D^{\text{CSF}}) + \sum_{i=1}^{M} w_i p_{\mathbf{y}_i}\left(\boldsymbol{y}\left|\boldsymbol{\mu}_i, \kappa, \frac{D_{/\!/}^{\text{fascicle}}}{1-2\nu\xi(\kappa)}, \nu\right.\right), \tag{2.63}$$

where $\sum_{i=0}^{M} w_i = 1$, $D^{\text{CSF}} = 3.0 \times 10^{-3}$ mm^2/s and $D_{/\!/}^{\text{fascicle}} = 1.71 \times 10^{-3}$ mm^2/s. Subsequently, the voxel-based corrected diffusivities read:

$$\begin{aligned} D_{/\!/}^{\text{voxel}} &= w_0 D^{\text{CSF}} + (1-w_0) D_{/\!/}^{\text{fascicle}}, \\ D_{\perp}^{\text{voxel}} &= w_0 D^{\text{CSF}} + \frac{(1-w_0) D_{/\!/}^{\text{fascicle}}}{1-2\nu\xi(\kappa)} \left(\frac{1-\nu}{\kappa+1} + \nu\xi(\kappa)\right). \end{aligned} \tag{2.64}$$

The PDF given by eq. (2.63) can be visualized as multiple iso-radius surfaces, for different values of radius. In detail, for a fixed radius r, each point at coordinates $(\theta, \phi) \in [0, \pi] \times [0, 2\pi]$ on the 2-sphere is dilated according to the value of the PDF at (r, θ, ϕ). The surface at a point with a high PDF value will be stretched out whereas the surface at a point with a low PDF value will grow hollow. Figure 2.6 gives an example where three fascicles in equal proportions oriented along the principal axes cross with a concentration $\kappa = 10$ and a free water fraction of occupancy $w_0 = 0.15$. Figures 2.6(a,c,e) were produced assuming no extra-axonal space ($\nu = 0$) (which can be seen as a ball-and-zeppelin diffusion model) whereas figs. 2.6(b,d,f) were produced with a proportion $\nu = 0.9$ of extra-axonal space. Figures 2.6(a,b) were generated with $r = 0.02$, (c,d) with $r = 0.04$ and (e,f) with $r = 0.08$. Regardless of the radius, dividing the axonal space within an intra- and an extra-axonal space seems to better delineate each fascicle.

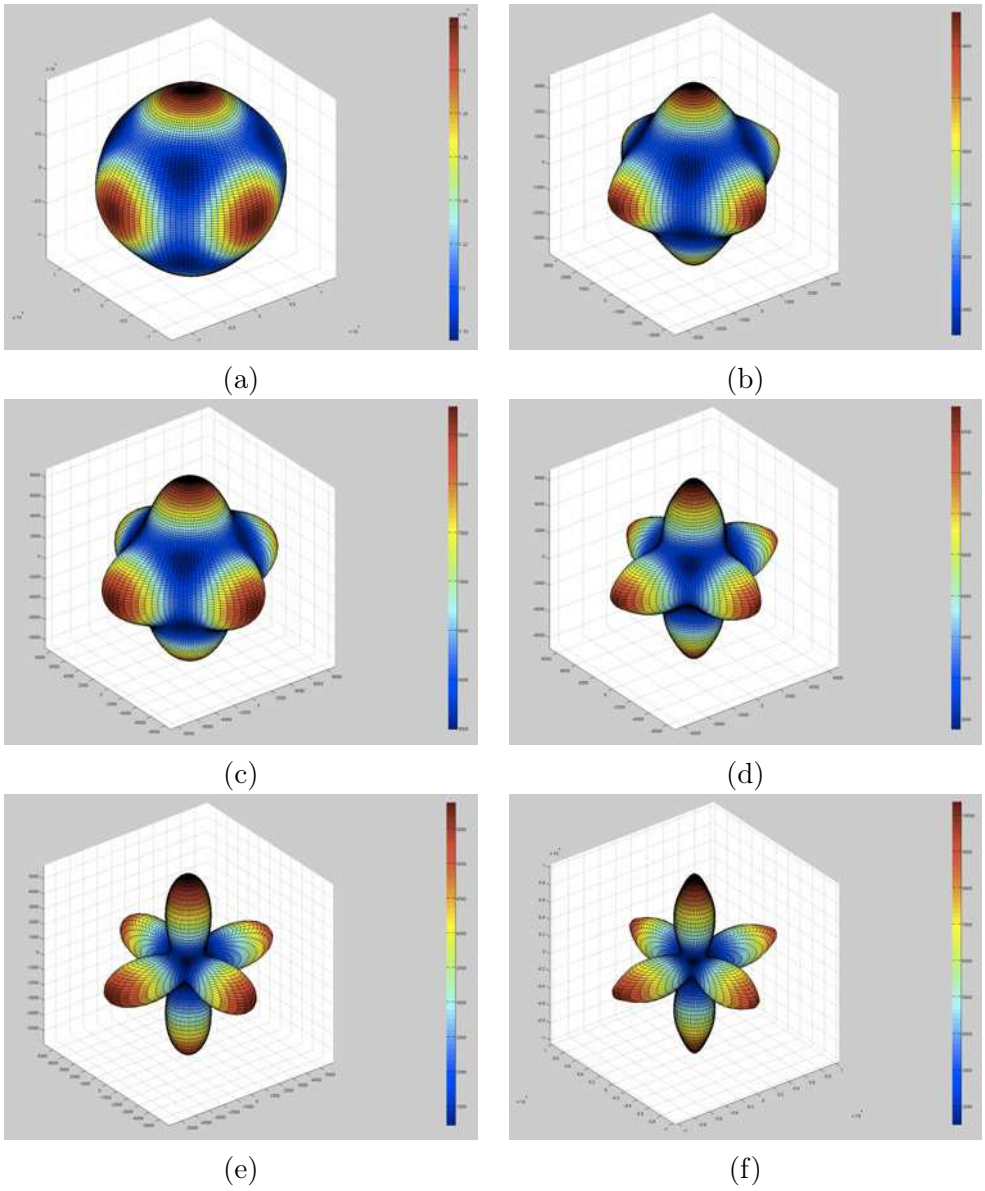


Figure 2.6: **Representation of iso-radius surfaces of the PDF of the DDI MCM.** Common parameters were the orientations set to the three reference axes, κ set to 10, w_0 set to 0.15 and fascicle occupancies set to 0.85/3. On the left column, no extra-axonal space ($\nu = 0$) and, on the right column, strong extra-axonal space ($\nu = 0.9$). Increasing radii from top to bottom: $r = 0.02$, 0.04 and 0.08.

2.3.4.4 Associated SID

Using q-space theory, the SID measured by dMRI experiments is related to the distribution of the DIMD $\mathbf{y}$ via eqs. (1.9) and (1.10). Assuming that the PDF of the distribution of the DIMD $\mathbf{y}$ is given by eq. (2.63) yields the following expression

of the SID:

$$A_{jk} = \left| \sum_{i=0}^{M} w_i A_{jk}^{(i)} \right| , \tag{2.65}$$

where $A_{jk}^{(0)}$ is given by eq. (2.56) with $d = D^{\text{CSF}} = 3.0 \times 10^{-3}$ mm^2/s and $\nu = 0$ and $A_{jk}^{(i)}$ $(i = 1, \ldots, M)$ is given by eq. (2.51) with d depending on κ and ν via eq. (2.50).

In summary, the proposed constrained DDI model comprises $3M+2$ independent parameters, where M is the assumed number of fascicles in the voxel. In details, the independent parameters of the DDI model are:

- the angles $(\theta_i, \phi_i) \in [0, \pi] \times [0, 2\pi]$ $(i = 1, \ldots, M)$ that define the orientations of the M fascicles,
- the fractions of occupancy $w_i \in [0, 1]$ $(i = 1, \ldots, M)$ of each fascicle,
- the concentration $\kappa \geq 0$ of water molecules around the orientation of the fascicle, which measures the orientational dispersion,
- the extra-axonal fraction of occupancy $\nu \in [0, 1]$.

Under such a model, the PDF of the DIMD in the voxel is given by eq. (2.63) and yields a SID given by eq. (2.65).

2.3.4.5 Relationship to other MCMs

The contribution to the SID of the diffusion within a fascicle is given by eq. (2.51). When the extra-axonal fraction of occupancy drops to zero ($\nu \to 0$), it reads:

$$A_{jk} = \exp\left\{ -\frac{b_k d}{\kappa + 1} \left(1 + \kappa < \boldsymbol{\mu}, \boldsymbol{g}_j >^2 \right) \right\} . \tag{2.66}$$

At this point, assuming the same free diffusivity in all the compartments (fascicles and CSF), the SID predicted by the DDI MCM boils down to the SID predicted by the ball-and-zeppelin MCM [Hosey 2005]. If we further assume low orientational dispersion ($\kappa \to \infty$), eq. (2.66) simplifies to:

$$A_{jk} = \exp\left\{ -b_k d < \boldsymbol{\mu}, \boldsymbol{g}_j >^2 \right\} , \tag{2.67}$$

and the SID predicted by the DDI MCM boils down to the SID predicted by the ball-and-stick Multi-Compartment Model (MCM) [Behrens 2003].

2.3.5 Summary of the geometry-based MCMs

Table 2.1 summarizes both the discussed state-of-the-art geometry-based MCMs and the the geometry-based MCMs introduced in this manuscript. The table regroups the independent parameters, constraints and model equations of each MCM and classifies them whether their estimation require single- or multi-shell acquisitions. In particular, the constrained DDI model introduced in section 2.3.4.3 is identified as DDID to emphasize that diffusivities of the different compartments are fixed.

In section 2.3.4.5, we showed that without eq. (2.50) and assuming that the diffusivities of the different compartments are identical, the DDI MCM can reduce to the

ball-and-zeppelin and ball-and-stick model by setting some parameters adequately. We thus introduced in table 2.1 a variant of the constrained DDI MCM identified as DDI, in which the diffusivities of the different compartments are assumed identical and estimated. This variant allows for a direct comparison with the ball-and-stick (BS) and ball-and-zeppelin (BZ) models. In turn, variants of the latter models with fixed diffusivities, identified as BSD and BZD, are introduced as well, in order to allow for a direct comparison with the fixed diffusivities DDI MCM (DDID) proposed in section 2.3.4.3.

2.4 Validation through simulations

2.4.1 Material & Methods

2.4.1.1 Objectives

The goals of this section, which is the last section of this chapter, is to compare the behaviour of the fixed diffusivities DDI model proposed in section 2.3 (DDID in table 2.1) with respect to other widely employed geometry-based MCMs. Using single-shell simulated diffusion data (which corresponds to what clinicians may afford), we focused on evaluating the influence of (i) the SNR, (ii) the b-value (iii) and (iii) the number of DSGs directions over the **fascicle orientation(s)**, the **free water fraction of occupancy** and the **extra-axonal fraction of occupancy**.

In the list presented in table 2.1, we chose to compare the fixed diffusivities DDI model mainly to its estimated diffusivity variant (DDI in table 2.1) and to the ball-and-stick and ball-and-zeppelin models (BS and BZ in table 2.1) and their fixed diffusivities variants (BSD and BZD in table 2.1), because these are all diffusion models adapted to single-shell diffusion data. However, neither the ball-and-stick nor the ball-and-zeppelin model exhibits an extra-axonal compartment. Consequently, even if it would require multi-shell diffusion data, we additionally fitted for single fascicle cases the MMWMD to allow for the comparison of its extra-axonal fraction of occupancy with the one obtained via DDI modeling. We chose the MMWMD over the CHARMED or NODDI because it is the only one currently implemented in the Camino software [Cook 2006]. It can thus be fitted with a very fast procedure similar to ours without resorting to a long MCMC procedure.

The experimental setup also allowed us, to some extent, to understand the effects of fixing diffusivities in the different compartments of an MCM.

2.4.1.2 Simulated data

Simulations were performed using the Monte-Carlo diffusion simulator [Hall 2009] in the Camino software [Cook 2006]. The simulator requires to set the number of spins N_s and the temporal resolution of the simulation through a number of timesteps N_t into which the simulation is divided. In accordance to the Camino guidelines, we set N_s = 100000 and N_t = 3200. The simulator assumes that each fascicle is made of a number of axons that can be represented as cylinders with a given radius and

a given distance between two cylinders: simulations were done with fixed radius cylinders with a radius set to 2×10^{-3}mm and a separation of 4.1×10^{-3}mm. The permeability of the membranes of the cylinders was set to zero. Finally, the free diffusivity along the orientation of each fascicle was set to $1.71 \times 10^{-3}\text{mm}^2/\text{s}$.

We synthetized 105 noise-free dMRI data sets without free water compartment ($w_0 = 0$) using a TE of 90ms, time Δ between two successively applied DSG of 40ms and pulse duration δ of 30ms, which reflects clinical conditions. We generated various configurations of fascicles and various diffusion sequences by varying the intensity of the DSGs (b-value) and the number of DSG directions. Specifically, we used:

- 1 fascicle always oriented along the z-axis with orientation $\pm(0, 0, 1)$,
- 7 different configurations for the second fascicle: we progressively separated it from the first one towards the x-axis with separation angle $a \in \{0, 15, 30, 45, 60, 75, 90\}$ (in degrees), such that the second fascicle is oriented along $\pm(-\sin a, 0, \cos a)$,
- 5 different b-values $b \in \{1000, 2000, 3000, 4000, 5000\}$ (in s/mm^2),
- 3 different sets of DSG directions uniformly distributed over the 2-sphere with $n \in \{16, 30, 61\}$ directions for each set.

For each data set we added five B0 images where no DSG were applied.

Finally, we corrupted each of the 105 noise-free generated data sets with 6 decreasing SNRs $\rho \in \{36, 30, 24, 18, 12, 6\}$ (in decibels), leading to the synthesis of a total of 630 dMRI datasets, each one containing 1000 replicates.

For a given SNR, a given b-value and a given number of DSG directions, we assessed the quality of the estimated parameters (fascicle orientation(s), free water occupancy, extra-axonal occupancy) visually by their median and inter-quartile range and we compared the estimates provided by DDID to the estimates provided by the other models by means of paired Wilcoxon tests. Specifically, we evaluated the quality of the estimated fascicle orientation(s) by computing the angular error between the estimated fascicle orientation and the true fascicle orientation. For two-fascicle cases, we first matched each estimated orientation to the corresponding ground truth and then compute the average of the angular errors commited on each fascicle.

2.4.1.3 Estimation framework

The estimation framework is identical for the models BS, BSD, BZ, BZD, DDI and DDID. It is based on a cost function, a derivative-free optimization algorithm and an initialization step. We shall abandon here the subscript $_k$ used to identify multiple b-values as all data sets were synthetized with a single shell.

Cost function. The measured DW signal S_j under the application of a DSG $\boldsymbol{q}_j = \gamma\delta G\boldsymbol{g}_j$ is the magnitude of Gaussian-corrupted data with standard deviation Ψ_0/ρ and is therefore Rician-corrupted with scale parameter Ψ_0/ρ. Ψ_0 is the true unweighted signal estimated as the average over all measured unweighted signals S_0.

[Gudbjartsson 1995] showed that, for moderate to high SNRs ρ, the noise can be well approximated by an offset Gaussian distribution with mean $\sqrt{\Psi_j^2 + \Psi_0^2/\rho^2}$ and standard deviation Ψ_0/ρ. Following [Panagiotaki 2012], we use this approximation and thus define our cost function as:

$$J = \sum_{j=1}^{n} \left(S_j - \sqrt{\Psi_j^2 + \frac{\Psi_0^2}{\rho^2}} \right)^2 , \tag{2.68}$$

where Ψ_j is the predicted DW signal from a given diffusion model.

For the estimation of the MMWMD, the same cost function was available in the Camino software and was therefore used.

Optimization algorithm The cost function given by eq. (2.68) needs to be minimized. Depending on the underlying diffusion model, this can be a very complex non-linear least squares problem and the derivative of the cost function may or may not be trivial. We thus resort to derivative-free optimization algorithms and use the ones implemented in the NLOpt optimization library [Johnson 2013]. Specifically, we use the Constrained Optimization BY Linear Approximations (COBYLA) described in [Powell 1994], which linearly approximates the cost function. At the cost of computational time, it outperforms the other derivative-free optimization algoritms when the cost function is not twice differentiable since all the other algorithms rely on quadratic approximations of the cost function. In order to reduce the computational time, we found more efficient to limit the number of function evaluations to a rather low number $N_{\text{eval}} = 200$ and to restart $N_r = 10$ times the optimization from the last found position. Lastly, the number of restarts was allowed to be less than N_r if the decrease of the cost function was less than 0.1.

The optimization of the cost function for estimating the MMWMD was however not performed with COBYLA since it is not an available option in the Camino software. Choices were the Levenberg-Marquardt algorithm [Levenberg 1944] or an MCMC procedure. We chose the Levenberg-Marquardt algorithm with N_r multiple restarts, which is closer to the procedure used to fit the other diffusion models and much faster.

Initialization. A derivative-free optimization algorithm like COBYLA requires a good starting point to improve its efficiency. We estimated the different models by increasing degrees of freedom given in table 2.1. Following the ideas in [Panagiotaki 2012], we initialized the orientation(s) of the fascicles for the different models as follows:

- the ball-and-stick model was initialized with a rough DTI estimation of the principal orientation of diffusion,
- the ball-and-zeppelin model was initialized with the previously estimated ball-and-stick fascicle orientation(s),
- the DDI model was initialized by the previously estimated ball-and-zeppelin fascicle orientation(s).

This sequential fitting was applied independently to the fixed diffusivities models and to the estimated diffusivity models. As suggested in [Scherrer 2012], for searching $M = 2$ fascicles, the initial DT principal eigenvector was rotated around its minor eigenvector by an angle of $\frac{d_\perp}{d_{/\!/}}\frac{\pi}{4}$, with $d_{/\!/}$ and $d_\perp$ being its parallel and perpendicular diffusivities. The other parameters, when estimated, were initialized to the following fixed values: $\kappa = 5$, $d = 1.71 \times 10^{-3}$ mm^2/s, $\nu = 0.5$ and $w_0 = 0.5$.

The initialization for the MMWMD was performed directly and in a blind way by the Camino software.

2.4.2 Results

We compared the angular error and estimated free water fraction of occupancy between models. The ground truth free water occupancy was 0 in the simulations. The first true fascicle orientation was $(0, 0, 1)$ and the second one was in the plane (x, z), moved away from the first one with increasing angles in the range $[0°, 90°]$ with a step of 15°.

We also qualitatively compared the estimated extra-axonal fraction of occupancy between DDID, DDI and MMWMD. The ground truth extra-axonal occupancy was 0.19 in the simulations.

2.4.2.1 Influence of the SNR

We evaluated the influence of the SNR on the angular error and the estimated free water fraction of occupancy for each model. In this scope, we fixed the b-value to 1000 s/mm^2 and the number of DSG directions to 30, which are clinically relevant values. Figure 2.7 qualitatively summarizes the results. Solid lines represent the median over the 1000 replicates for a specific SNR, while long and short dashed lines represent respectively the lower and upper quartile. The over-simplified representation of the fascicle in the ball-and-stick models (BSD and BS) forces the model to compensate by estimating a large free water occupancy. For all the other models, the estimated occupancy converges to the ground truth value and the angular error decreases as the SNR increases. Interestingly, except for the ball-and-stick model, it seems that estimating a common diffusivity for all the compartments yields free water occupancy estimates that are more robust to noise than those provided with fixed diffusivities. However, if the differences seem to be large for the ball-and-zeppelin model, DDID and DDI produce very similar estimates. Finally, the DDI models seem to provide lower angular errors.

We performed a quantitative analysis of these results for the 60°-crossing fascicle case. We aimed at comparing the estimated free water occupancies and angular errors of each model to the ones obtained by our proposed model (DDID). Tables 2.2 and 2.3 give the mean value over the 1000 replicates for each SNR of the estimated free water occupancies and angular errors, respectively. The results of paired Wilcoxon tests are also reported by adding the superscript * to values that are statistically different from the DDID value (orange column). The quantitative

analysis confirms the conclusions of the qualitative analysis regarding the free water occupancy (Table 2.2): BSD and BS provide strongly biased free water occupancies, BZ provides estimated free water occupancies that are significantly more robust to noise than those provided by BZD and the same trend can be observed for the DDI model but it is not statistically significant. In terms of mean values, BZ seems to be the best one for estimating the free water occupancy but DDID provides significantly better estimated fascicle orientations compared to all other models (Table 2.3).

2.4.2.2 Influence of the b-value

We evaluated the influence of the b-value on the angular error and the estimated free water fraction of occupancy for each model. In this scope, we fixed the SNR to 30 dB and the number of DSG directions to 30, which are clinically relevant values. Figure 2.8 qualitatively summarizes the results. Solid lines represent the median over the 1000 replicates for a specific b-value, while long and short dashed lines represent respectively the lower and upper quartile. The over-simplified representation of the fascicle in the ball-and-stick models (BSD and BS) forces the model to compensate by estimating a large free water occupancy that increases with the b-value. For all the other models, the estimated occupancy is very close to the ground truth value and the angular error decreases as the b-value increases. Interestingly, except for the ball-and-stick model, it seems that estimating a common diffusivity for all the compartments yields free water occupancy estimates that are more robust to noise than those provided with fixed diffusivities. However, the difference is larger between BZD and BZ than between DDID and DDI. Finally, for large separation angles between fascicles, fixing the diffusivities provides lower angular errors and, for small separation angles, it is the contrary: estimating a common diffusivity provides lower angular errors. DDID or DDI (according to the separation angle) outperforms the other models for estimating fascicle orientations.

We performed a quantitative analysis of these results for the 60°-crossing fascicle case. We aimed at comparing the estimated free water occupancies and angular errors of each model to the ones obtained by our proposed model (DDID). Tables 2.4 and 2.5 give the mean value over the 1000 replicates for each b-value of the estimated free water occupancies and angular errors, respectively. The results of paired Wilcoxon tests are also reported by adding the superscript * to values that are statistically different from the DDID value (orange column). The quantitative analysis confirms the conclusions of the qualitative analysis regarding the free water occupancy (Table 2.4): BSD and BS provide strongly biased free water occupancies, BZ (resp. DDI) provides estimated free water occupancies that are significantly more robust to noise than those provided by BZD (resp., DDID) except at the clinical b-value of 1000 s/mm^2 for which there is no statistical difference between DDI and DDID. In terms of mean values, BZ seems to be the best one for estimating the free water occupancy but DDID provides significantly better estimated fascicle orientations compared to all other models (Table 2.5).

2.4.2.3 Influence of the number of DSG directions

We evaluated the influence of the number of DSG directions on the angular error and the estimated free water fraction of occupancy for each model. In this scope, we fixed the b-value to 1000 s/mm^2 and the SNR to 30 dB, which are clinically relevant values. Figure 2.9 qualitatively summarizes the results. Solid lines represent the median over the 1000 replicates for a specific number of DSG directions, while long and short dashed lines represent respectively the lower and upper quartile. The number of DSG directions does not seem to affect the estimation of the free water occupancy. The over-simplified representation of the fascicle in the ball-and-stick models (BSD and BS) forces the model to compensate by estimating a large free water occupancy. For all the other models, the estimated occupancy is close to the ground truth value and the angular error decreases as the number of DSG directions increases. Interestingly, except for the ball-and-stick model, it seems that estimating a common diffusivity for all the compartments yields estimates that are more robust to noise than those provided with fixed diffusivities. However, if the differences seem to be large for the ball-and-zeppelin model, DDID and DDI produce very similar estimates. Finally, for large separation angles between fascicles, fixing the diffusivities provides lower angular errors and, for small separation angles, it is the contrary: estimating a common diffusivity provides lower angular errors. DDID or DDI (according to the separation angle) outperforms the other models for estimating fascicle orientations.

We performed a quantitative analysis of these results for the 60°-crossing fascicle case. We aimed at comparing the estimated free water occupancies and angular errors of each model to the ones obtained by our proposed model (DDID). Tables 2.6 and 2.7 give the mean value over the 1000 replicates for each number of DSG directions of the estimated free water occupancies and angular errors, respectively. The results of paired Wilcoxon tests are also reported by adding the superscript * to values that are statistically different from the DDID value (orange column). The quantitative analysis confirms the conclusions of the qualitative analysis regarding the free water occupancy (Table 2.6): BSD and BS provide strongly biased free water occupancies, BZ provides estimated free water occupancies that are significantly more robust to noise than those provided by BZD and the same trend can be observed for the DDI model but it is not statistically significant. In terms of mean values, BZ seems to be the best one for estimating the free water occupancy but DDID provides significantly better estimated fascicle orientations compared to all other models (Table 2.7).

2.4.2.4 Estimation of the extra-axonal fraction of occupancy

We evaluated the influence of the SNR, the b-value and the number of DSG directions on the estimated extra-axonal fraction of occupancy for DDID, DDI and MMWMD. Figure 2.10 summarizes the results. Solid lines represent the median over the 1000 replicates for a specific number of DSG directions, while long and short dashed lines represent respectively the lower and upper quartile. Figure 2.10(a) shows the

influence of the SNR at $b = 1000$ s/mm^2 with 30 DSG directions (clinical conditions). As the SNR increases, both DDID and MMWMD converge to the ground truth with a reduced variability in the estimates. The convergence is faster with DDID than with MMWMD. DDI provides a biased estimate of the extra-axonal occupancy even with very high SNRs. Figure 2.10(b) shows the influence of the b-value at SNR = 30 dB with 30 DSG directions (clinical conditions). The b-value does not influence the estimation of the extra-axonal occupancy. DDI provides a positively biased estimate, MMWMD provides a negatively biased estimate and DDID provides an unbiased estimate. Figure 2.10(c) shows the influence of the number of DSG directions at SNR = 30 dB with $b = 1000$ s/mm^2 (clinical conditions). For the DDI models, increasing the number of DSG directions slightly reduces the variance of the estimated extra-axonal occupancy but does not affect the mean estimate, while, for the MMWMD, it does not affect the variance of the estimate but the mean estimate increases and converges to the ground truth value. DDI provides a positively biased estimate, MMWMD provides a negatively biased estimate whose bias reduces as the number of DSG directions increases and DDID always provides an unbiased estimate.

2.4.2.5 Conclusions from the simulation study

The simulation study allows us to reach the following conclusions:

- BZ and DDI outperform the other models for estimating the free water fraction of occupancy,
- DDID still provides estimates of the free water occupancy that are very close to the ground truth value,
- When the separation angle between fascicles is small, DDI outperforms the other models for estimating fascicle orientations,
- When the separation angle between fascicles is large, DDID outperforms the other models for estimating fascicle orientations,
- DDID is the only model able to accurately (unbiasness) estimate the extra-axonal fraction of occupancy from clinical diffusion data ($b = 1000$ s/mm^2 and 30 DSG directions),
- It is recommended to achieve the best SNR possible to increase the estimates precision (inverse of variance), in particular the precision on the estimated extra-axonal occupancy.

In summary, we introduced a novel diffusion model coined Diffusion Directions Imaging (DDI) that outperforms existing geometry-based diffusion models for estimating the fascicle orientations in clinical conditions ($b = 1000$ s/mm^2 and 30 DSG directions). Moreover, if we set the diffusivity in the fascicle compartments with eq. (2.50) and the diffusivity of free water to its typical value at 37°C, we are able, under clinical conditions ($b = 1000$ s/mm^2 and 30 DSG directions), to achieve an unbiased estimate of the extra-axonal fraction of occupancy whose precision increases with the SNR and to accurately estimate the free water fraction of occupancy.

We also need to validate our proposed diffusion model on real clinical diffusion data. However, with real data, the exact number of fascicles per voxel is not known. We must first address the model selection problem inherent to any MCM. This will be the object of the next chapter, at the end of which we will provide a validation of the DDI model on real clinical diffusion data.

Model	Fascicle orientation	Concentration	Intra-axonal free diffusivity (10^{-3} mm^2/s)	Extra-axonal free diffusivity (10^{-3} mm^2/s)	Axon radius	Extra-axonal frac.	Intra-axonal frac.	Free water frac.	Free water diffusivity (10^{-3} mm^2/s)	Intra-axonal SID	Extra-axonal SID	Deg. of freedom
BSD	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	∞	d from eq. (2.50)	NA	NA	0	w_i	$1-\sum_{i=1}^{M} w_i$	3.0	eq. (2.67)	none	$3M$
BS [Behrens 2003]	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	∞	d	NA	NA	0	w_i	$1-\sum_{i=1}^{M} w_i$	d	eq. (2.67)	none	$3M$+1
BZD	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	κ	d from eq. (2.50)	NA	NA	0	w_i	$1-\sum_{i=1}^{M} w_i$	3.0	eq. (2.66)	none	$3M$+1
BZ [Hosey 2005]	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	κ	d	NA	NA	0	w_i	$1-\sum_{i=1}^{M} w_i$	d	eq. (2.66)	none	$3M$+2
DDID	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	κ	$(1-\nu)d$, with d from eq. (2.50)	νd, with d from eq. (2.50)	NA	νw_i	$(1-\nu)w_i$	$1-\sum_{i=1}^{M} w_i$	3.0	eq. (2.51)		$3M$+2
DDI	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	κ	$(1-\nu)d$	νd	NA	νw_i	$(1-\nu)w_i$	$1-\sum_{i=1}^{M} w_i$	d	eq. (2.51)		$3M$+3
CHARMED [Assaf 2004]	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	NA	d_i	d_i	fixed	$1-\sum_{i=1}^{M} f_i^r$	f_i^r	0	NA	eqs. (2.10) to (2.12)	eq. (2.9) with full DT D	$4M$+6
CHARMED [Assaf 2005a]	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	NA	d_i	d_i	fixed	$1-\sum_{i=1}^{M} f_i^r$	f_i^r	0	NA	[?]	eq. (2.9) with full DT D	$4M$+6
CHARMED [Alexander 2008]	$\boldsymbol{\mu}_i(\theta_i, \phi_i)$	NA	d	d	R	$1-\sum_{i=1}^{M} f_i^r$	f_i^r	0	NA	[?]	eq. (2.9) with $D = d_\perp^e I_3 + (d - d_\perp^e)\boldsymbol{\mu}\boldsymbol{\mu}'$	$3M$+3
MMWMD [Alexander 2010]	$\boldsymbol{\mu}_1(\theta_1, \phi_1)$	NA	$d = 1.7$	d	R	$\nu_{ic}(1-\nu_0)$	$(1-\nu_{ic}) \cdot (1-\nu_0)$	ν_0	3.0	[?]	eq. (2.9) with $D = d(1-\nu_{ic})I_3 + d\nu_{ic}\boldsymbol{\mu}\boldsymbol{\mu}'$	5
NODDI [Zhang 2012]	$\boldsymbol{\mu}_1(\theta_1, \phi_1)$	κ	$d = 1.7$	d	NA	$\nu_{ic}(1-\nu_0)$	$(1-\nu_{ic}) \cdot (1-\nu_0)$	ν_0	3.0	eq. (2.14)	eqs. (2.15) to (2.17)	5

Table 2.1: **Summary of the geometry-based MCMs.** Independent parameters, constraints and link to corresponding model equations. NA means that the parameter is not an output of the model. A parameter written solely with its letter is an independent parameter to be estimated. Models in the green part of the table require single shell acquisitions while those in the orange part require multi-shell acquisitions. BS: ball-and-stick, BZ: ball-and-zeppelin, BSD/BZD/DDID: BS/BZ/DDI with fixed diffuvities. Unreferenced models are introduced in this manuscript.

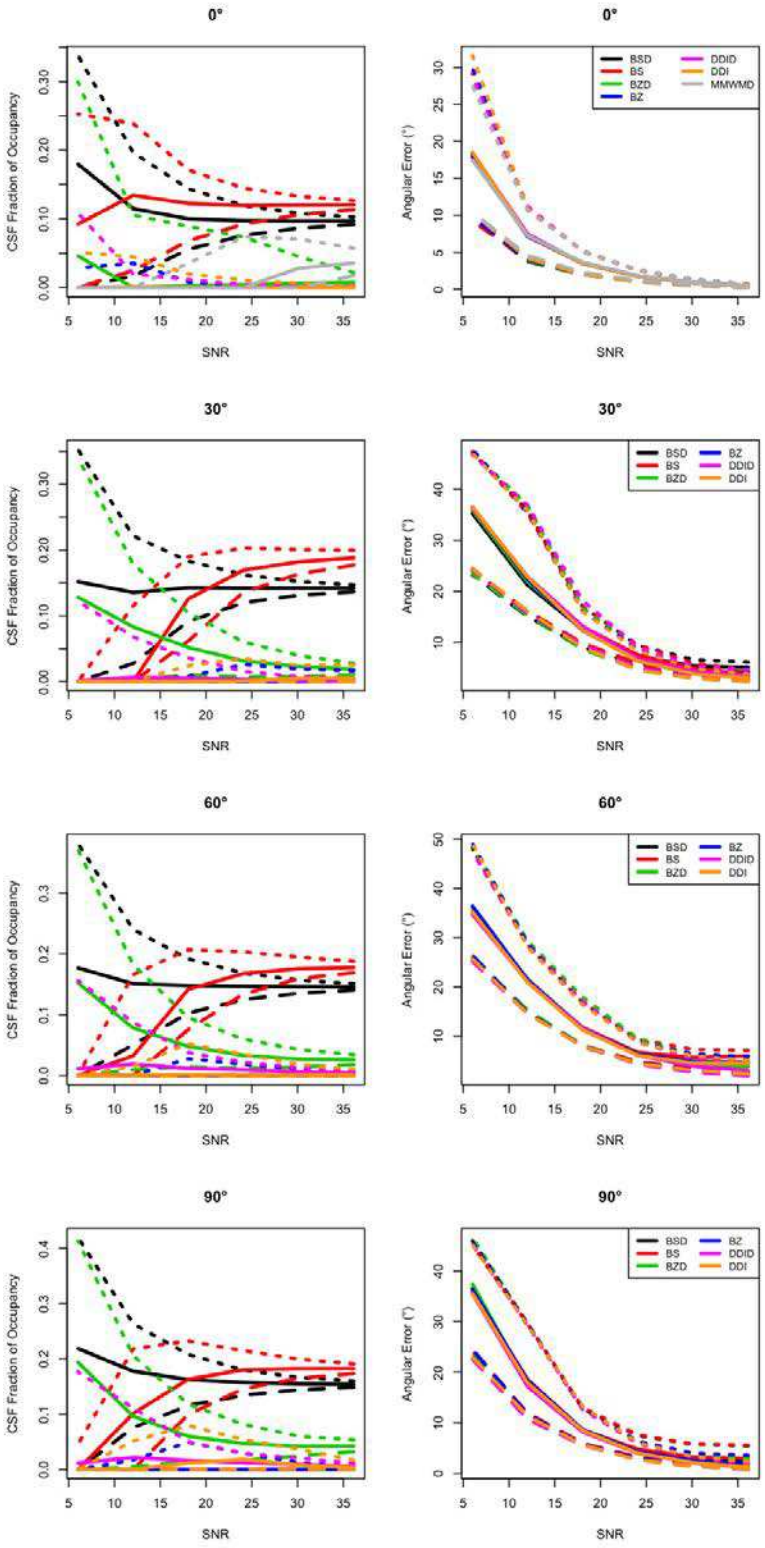


Figure 2.7: **Influence of the SNR (dB) for different separation angles between fascicles.** 0° (1st row), 30° (2nd row), 60° (3rd row), 90° (4th row). Left column: estimated free water occupancy. Right column: angular error on the estimated orientations. Solid, long and short dashed lines represent median, lower and upper quartile, respectively.

SNR (dB)	BSD	BS	BZD	BZ	DDID	DDI
36	0.15^*	0.18^*	0.026^*	$7.3e-5^*$	$3.8e-3$	$4.9e-4$
30	0.15^*	0.18^*	0.027^*	$6.3e-5^*$	$6.1e-3$	$3.9e-4$
24	0.15^*	0.17^*	0.033^*	$5.5e-5^*$	$9.5e-3$	$5.1e-4$
18	0.15^*	0.14^*	0.048^*	$6.6e-5^*$	0.012	$9.9e-4$
12	0.15^*	0.032^*	0.078^*	$4.1e-5^*$	0.019	$2.3e-4^*$
6	0.18^*	$1.3e-5^*$	0.15^*	$3.3e-5^*$	0.011	$7.3e-5^*$

Table 2.2: **Quantitative comparison of the estimated free water occupancies for the 60°-crossing fascicle case at different SNRs.** Values are the mean over the 1000 replicates for each SNR. Paired Wilcoxon tests have been performed between each model and our proposed model (DDID, identified by the orange column). Statistically significant ($p < 0.05$) differences are indicated by the superscript *.

SNR (dB)	BSD	BS	BZD	BZ	DDID	DDI
36	3.9^*	6.0^*	3.9^*	4.7^*	3.0	4.5^*
30	4.3^*	5.7^*	4.2^*	4.9^*	4.0	4.6^*
24	6.3^*	6.7^*	6.4^*	6.3^*	6.1	6.1
18	11.4	11.8^*	11.7^*	11.6^*	11.6	11.3
12	21.6^*	21.2^*	21.6^*	21.4^*	21.1	20.9
6	35.2^*	36.3^*	35.7^*	36.4^*	34.7	35.2^*

Table 2.3: **Quantitative comparison of the angular errors (°) for the 60°-crossing fascicle case at different SNRs.** Values are the mean over the 1000 replicates for each SNR. Paired Wilcoxon tests have been performed between each model and our proposed model (DDID, identified by the orange column). Statistically significant ($p < 0.05$) differences are indicated by the superscript *.

b-value (s/mm^2)	BSD	BS	BZD	BZ	DDID	DDI
1000	0.15^*	0.18^*	0.027^*	$6.3e-5^*$	$6.1e-3$	$3.9e-4$
2000	0.22^*	0.23^*	0.013^*	$2.9e-5^*$	0.025	$2.3e-4^*$
3000	0.25^*	0.26^*	0.029^*	$2.6e-5^*$	0.021	$1.2e-4^*$
4000	0.27^*	0.28^*	0.018^*	$2.7e-5^*$	$7.1e-3$	$9.9e-5^*$
5000	0.28^*	0.28^*	$7.8e-3^*$	$2.8e-5^*$	$9.8e-3$	$1.6e-4^*$

Table 2.4: **Quantitative comparison of the estimated free water occupancies for the 60°-crossing fascicle case at different b-values.** Values are the mean over the 1000 replicates for each SNR. Paired Wilcoxon tests have been performed between each model and our proposed model (DDID, identified by the orange column). Statistically significant ($p < 0.05$) differences are indicated by the superscript *.

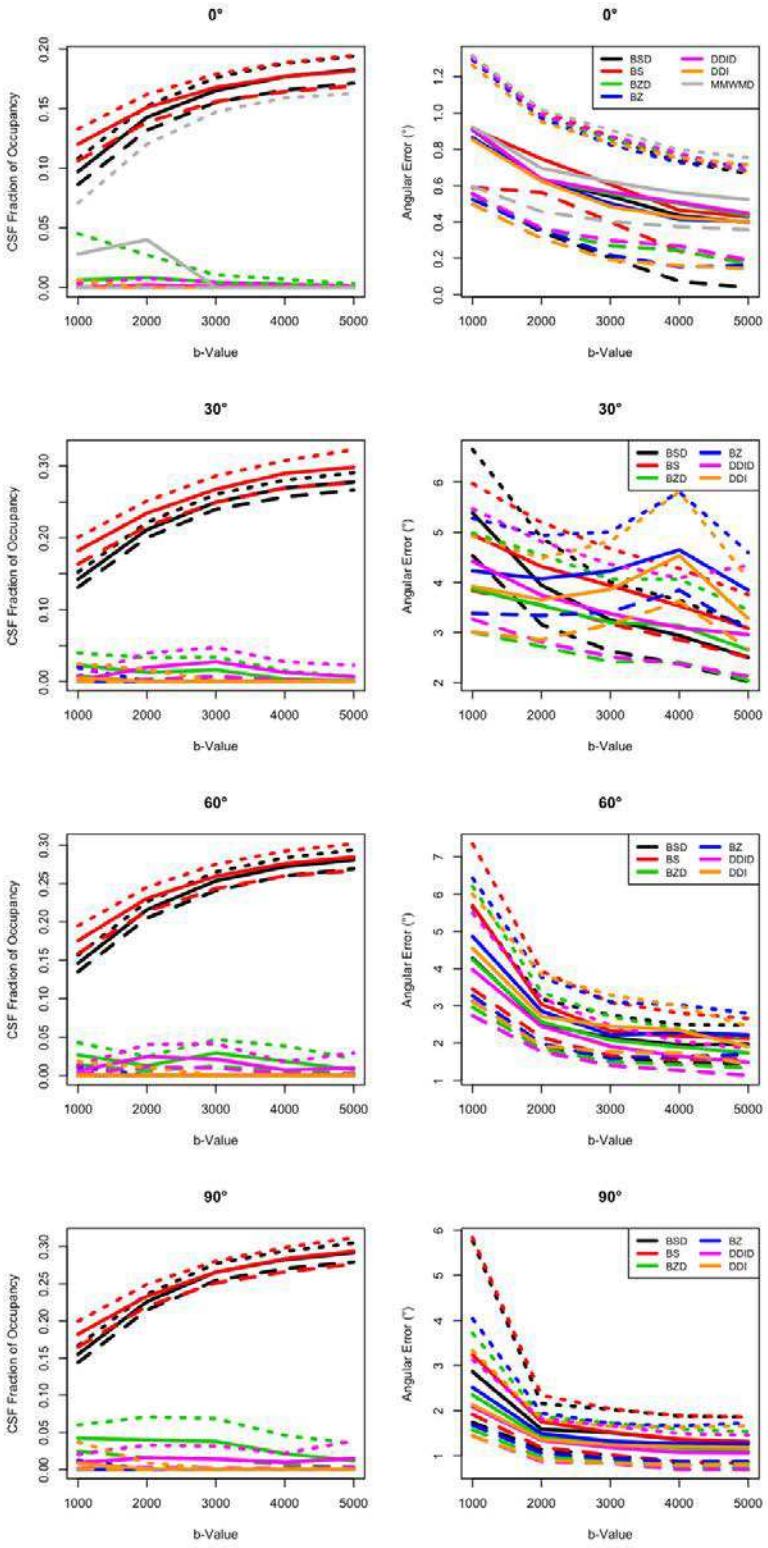


Figure 2.8: **Influence of the b-value for different separation angles between fascicles.** 0° (1st row), 30° (2nd row), 60° (3rd row), 90° (4th row). Left column: estimated free water occupancy. Right column: angular error on the estimated orientations. Solid, long and short dashed lines represent median, lower and upper quartile, respectively.

b-value (s/mm^2)	BSD	BS	BZD	BZ	DDID	DDI
1000	4.3*	5.7*	4.2*	4.9*	4.0	4.6*
2000	2.5	3.0*	2.6*	2.9*	2.5	2.7*
3000	2.1*	2.3*	2.1*	2.2*	1.9	2.5*
4000	2.0*	2.2*	1.9*	2.3*	1.6	2.3*
5000	2.0*	2.1*	1.7*	2.2*	1.5	1.9*

Table 2.5: **Quantitative comparison of the angular errors (°) for the 60°-crossing fascicle case at different b-values.** Values are the mean over the 1000 replicates for each SNR. Paired Wilcoxon tests have been performed between each model and our proposed model (DDID, identified by the orange column). Statistically significant ($p < 0.05$) differences are indicated by the superscript *.

Nb. DSG grads	BSD	BS	BZD	BZ	DDID	DDI
16	0.15*	0.16*	0.035*	$3.6e-5$*	$8.5e-3$	$1.7e-4$*
30	0.15*	0.18*	0.027*	$6.3e-5$*	$6.1e-3$	$3.9e-4$
61	0.15*	0.18*	0.026*	$6.4e-5$*	$5.7e-3$	$6.9e-4$

Table 2.6: **Quantitative comparison of the estimated free water occupancies for the 60°-crossing fascicle case at different numbers of DSG directions.** Values are the mean over the 1000 replicates for each SNR. Paired Wilcoxon tests have been performed between each model and our proposed model (DDID, identified by the orange column). Statistically significant ($p < 0.05$) differences are indicated by the superscript *.

Nb. DSG grads	BSD	BS	BZD	BZ	DDID	DDI
16	6.2*	6.4*	4.9*	5.1*	4.5	4.8*
30	4.3*	5.7*	4.2*	4.9*	4.0	4.6*
61	3.5*	4.7*	3.4*	4.2*	3.1	4.1*

Table 2.7: **Quantitative comparison of the angular errors (°) for the 60°-crossing fascicle case at different number of DSG directions.** Values are the mean over the 1000 replicates for each SNR. Paired Wilcoxon tests have been performed between each model and our proposed model (DDID, identified by the orange column). Statistically significant ($p < 0.05$) differences are indicated by the superscript *.

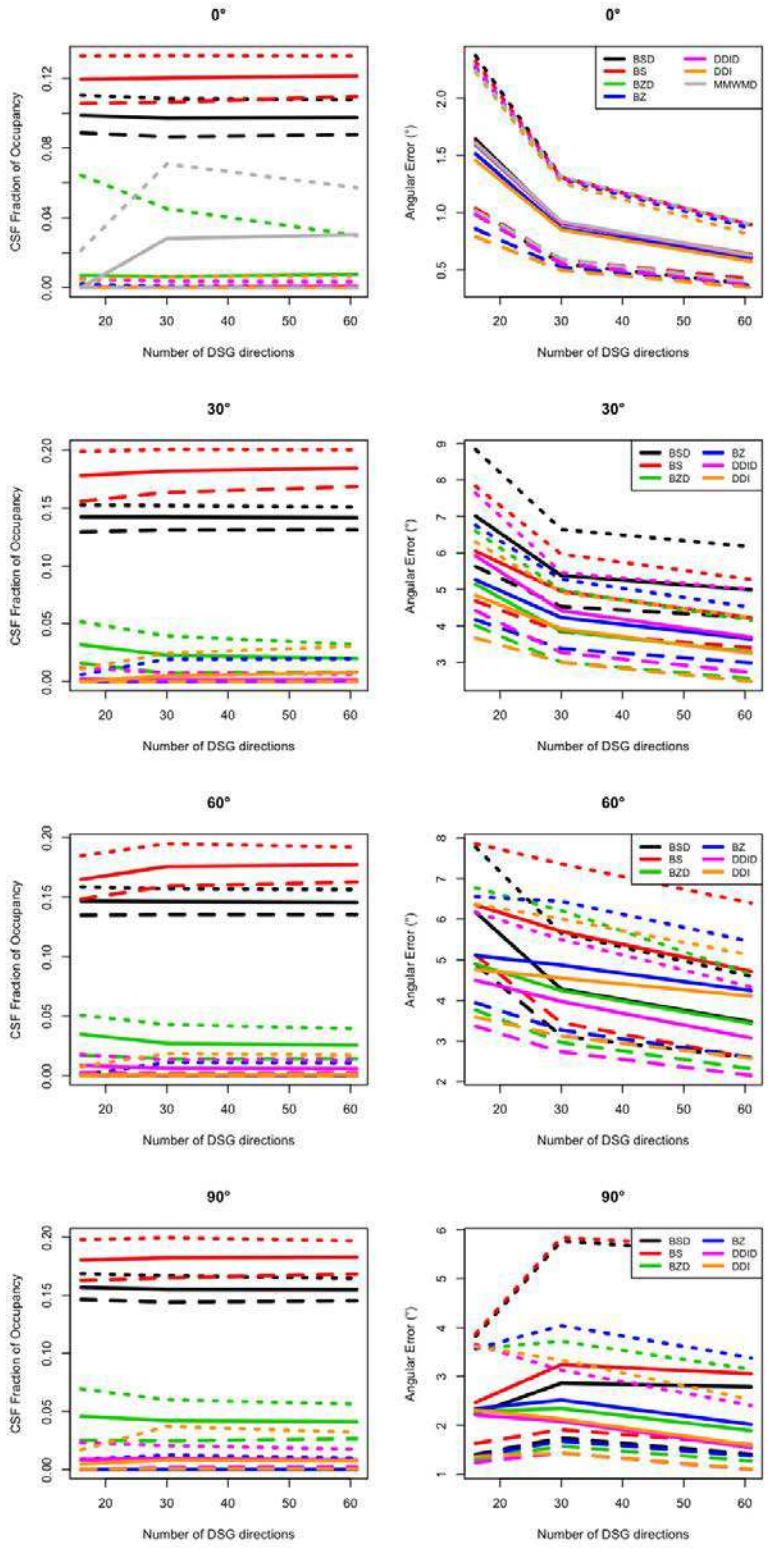


Figure 2.9: **Influence of the number of DSG directions for different separation angles between fascicles.** 0° (1st row), 30° (2nd row), 60° (3rd row), 90° (4th row). Left column: estimated free water occupancy. Right column: angular error on the estimated orientations. Solid, long and short dashed lines represent median, lower and upper quartile, respectively.

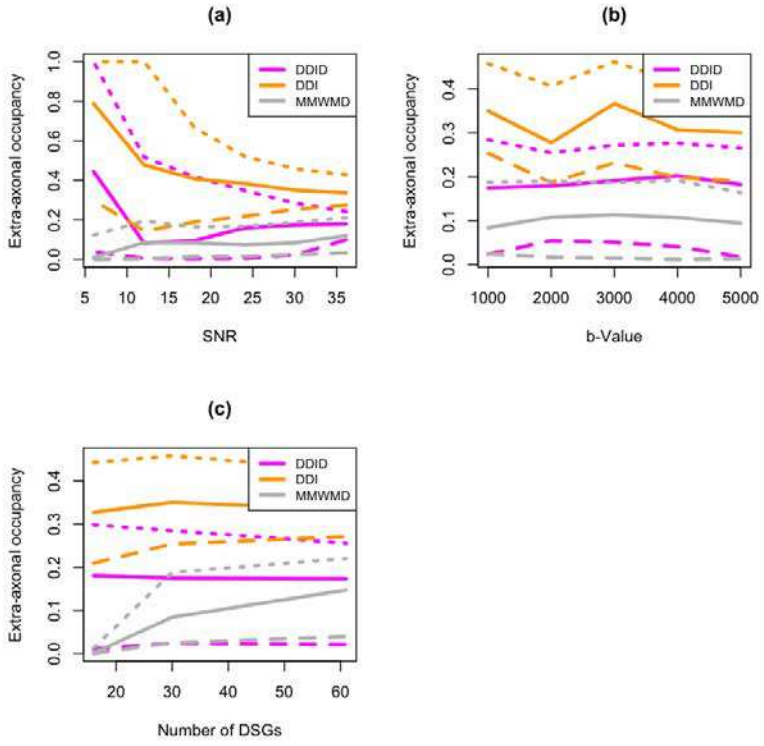


Figure 2.10: **Estimated extra-axonal fraction of occupancy for single fascicle voxel.** Influence of the SNR (top left), *b*-value (top right) and number of DSG directions (bottom left). Compared models are DDID (purple), DDI (orange) and MMWMD (gray). Solid, long and short dashed lines represent median, lower and upper quartile, respectively.

Chapter 3

Determination of the number of fascicles

Contents

In chapter 2, we showed that MCMs can be inferred from the measured DW images and are of great interest to study the microstructure of the WM. These models assume that, in a given voxel of an imaged brain, water molecules can be divided into several **compartments**, each of them contributing to the predicted DW signal in its own way. We reviewed the geometry-based MCMs proposed in the recent dMRI literature, namely the ball-and-stick model, the ball-and-zeppelin model, the CHARMED, the MMWMD and NODDI. We analyzed their strengths and weaknesses and proposed a new geometry-based MCM, coined DDI, that shows promising results on single-shell low angular resolution measured DW images. The main difference between these models stems from how diffusion is modeled in the fascicle(s).

Yet, they share a number of commonalities: (i) they include a specific compartment to model free isotropic diffusion, (ii) the other compartments always aim at describing the diffusion in the fascicle(s), (iii) the *fraction of occupancy* of each compartment weights the contribution of the compartment to the overall predicted DW signal and (iv) they require the number of compartments to be fixed a priori. The last point is the major drawback of MCMs and often precludes their use in practice. Determining how many fascicles populate a given voxel constitutes a model selection problem, which can be tackled in two very different ways. **Brute force** approaches, detailed in section 3.1, attribute to each voxel an optimal number

of fascicles. Differently, **averaging** approaches evaluate the relative likelihood of a set of models given the data and output an averaged model in each voxel. These approaches are the object of section 3.2, in which we review the averaging methods used in dMRI for the estimation of the number of fascicles and we propose a novel method. Finally, we apply in section 3.3 the new proposed approach to the DDI model introduced in section 2.3, which enables its validation on real clinically acquired diffusion data.

For the remainder of this chapter, as usual, we denote by S_{jk} the measured DW image/signal in response to a b-value b_k ($k = 1, \ldots, n_b$) and a DSG direction $\boldsymbol{g}_j$ ($j = 1, \ldots, n_g$) and by Ψ_{jk} the corresponding predicted DW image/signal under a given generative MCM. The true unweighted image/signal Ψ_0 is taken to be the average of all the measured unweighted images/signals S_0 in all our fitting procedures. Finally, let $N = n_g n_b$ and $\mathbf{S}_N = \{S_j\}_{j=1,\ldots,n_g,n_g+1,\ldots,N}$ be the set of all measured DW images.

3.1 Brute force approaches

Brute force approaches select the *optimal* diffusion model among a set of models as the diffusion model that best fits the data. The goodness of fit of each model is generally assessed by the fitting error, which reads:

$$\widehat{E}_g^{\text{fit}} = \frac{1}{N} \sum_{j=1}^{N} (S_j - \Psi_j)^2 \ .$$

Related methods usually rely on an F-test to determine whether the fitting error of a model is significantly lower than the fitting error of another one [Alexander 2002, Kreher 2005, Scherrer 2012]. The fitting error is however computed from the same data that was used for the estimation of the MCM. As a result, complex models are often favored since it is always possible to find a model that procudes $\hat{E}_g^{\text{fit}} = 0$, which yields an overfitting of the data, especially in the case of nested models like, for instance, MCMs in dMRI. A number of criteria have thus been proposed to assess the goodness of fit of a model while penalizing its complexity [Raftery 1992]. To date, only the Bayesian Information Criterion (BIC) has been introduced for selecting the optimal number of fascicles but has been shown to yield suboptimal results [Schultz 2010].

Recently, [Scherrer 2013] recommended to compare the performances of diffusion models by assessing their generalization error, which measures the ability of the model to predict new unseen data. The generalization error depends on the distribution F of the diffusion data, which is usually unknown. Using the empirical distribution yields the fitting error, which is therefore a biased estimate of the generalization error. Bias can be lowered by leave-one-out cross-validation estimate of F at the cost of a high variance. The variance can in turn be lowered by K-fold cross-validation at the cost of increased bias. There is thus a trade-off to make between bias and variance when estimating the generalization error. [Efron 1983]

introduced the 632 bootstrap estimator, which, in essence, computes the generalization error as the weighted sum of the fitting error and a bootstrap estimate that lower the variance of the cross-validation estimate by computing the distribution F via a bootstrap smoothing method. The fitting error presents a negative bias whereas the bootstrap estimate presents a high positive bias. [Efron 1983] demonstrated, under the assumption of low overfitting, that imputing a weight of 0.632 to the bootstrap estimate balances the bias of the two estimates while keeping an overall low variance thanks to the bootstrap estimate. However, if overfitting is important, the fitting error bias might outweigh the bootstrap estimate bias. A correction to the weight 0.632 has thus been introduced by [Efron 1997] in order to give more weight to the bootstrap estimate of the generalization error when the fitting error presents substantial overfitting.

Model selection based on the assessment of the generalization error produces non-monotonically decreasing generalization errors as the number of fascicles increases. [Scherrer 2013] show on simulated data that a global minimum of the generalization error occurs for the ground truth number of fascicles. Also, they demonstrate on real data that this method is much less sensitive to common artifacts (e.g. head motion) than the F-test method proposed by [Alexander 2002] and implemented in the Camino software[1].

3.2 Averaging approaches

3.2.1 Automatic Relevance Detection: implicit model averaging

Averaging approaches rely on the idea that there is no such thing as a best model that strictly gives a number of fascicles within a voxel. Models with different number of fascicles rather have a relative likelihood given the data. If these likelihoods are known, they can be used to average the different models. In practice however, they are not known. To circumvent this issue, one can use a Bayesian trick known as Automatic Relevance Detection (ARD). To the best of our knowledge, this is the only averaging approach that has been applied for the estimation of MCMs. The method is employed to estimate only the more complex (i.e., the one with the highest desired number of fascicles) ball-and-stick model (see section 2.2.1). Fascicle compartments that are not supported by the data are then automatically pruned by ARD, performing an implicit model averaging [Behrens 2007].

ARD relies on the marginalisation over hyperparameters performed in a fully Bayesian estimation framework. It can be applied to any parameters of a given model. The key idea is to design an appropriate prior distribution for the parameter of interest, such that it will take a value zero with very low variance in the posterior distribution if it is not supported by the data. Usually, such a prior is a zero mean Gaussian distribution with unknown variance. A conjugate prior is subsequently assigned to the unknown variance parameter, which enables the marginalization

[1] http://cmic.cs.ucl.ac.uk/camino/

over the variance hyperparameter.

Let us recall here the ball-and-stick model that needs to be estimated:

$$\Psi_j = \Psi_0 \left(\left(1 - \sum_{i=1}^{M} f_i \right) \exp\left\{ -b_j d \right\} + \sum_{i=1}^{M} f_i \exp\left\{ -b_j d < \boldsymbol{\mu}_i, \boldsymbol{g}_j >^2 \right\} \right), \tag{3.1}$$

where $\Psi_0 > 0$ is the unweighted predicted signal, $d > 0$ is the diffusivity, $\boldsymbol{\mu}_i$ and f_i are respectively the orientation and fraction of occupancy of the i-th fascicle compartment and M is the desired number of fascicles to be estimated. Noise is modeled independently in each voxel by a zero mean Gaussian distribution with variance σ^2.

When $M = 1$, ARD is not used and no averaging is performed. When $M > 1$, ARD is applied to the fractions of occupancy of the fascicle compartments, except for the first fascicle compartment (the one with highest occupancy). Prior distributions on the other parameters are motivated in [Behrens 2003] and only reported here for sake of clarity:

$$\begin{aligned} P(\Psi_0) \sim U(0,\infty)\,,\ P(f_1) \sim U(0,1)\,,\ P(d) \sim \Gamma(a,b)\,,\ P(\sigma) \sim \sigma^{-1}\,, \\ P(\theta_i) \sim \sin\theta\,,\ P(\phi_i) \sim U(0,2\pi)\,, \forall i \in [\![1, M]\!]\,. \end{aligned} \tag{3.2}$$

Briefly, priors on θ_i and ϕ_i are chosen to uniformly cover the 2-sphere $\mathbb{S}^2$, priors on variance parameters are conjugate priors and the other priors are chosen to be noninformative and adapted to the boundaries of each parameter.

In the Bayesian framework proposed by [Behrens 2003], the prior distribution of the fractions of occupancy $f_{2:M}$ were also uniform on $[0,1]$. In order to perform ARD on them, their prior needs to be redefined. The Gaussian distribution is not suited to the fractions of occcupancy since they are bounded to $[0,1]$. [Behrens 2007] thus propose to use a Beta distribution with mode at zero and unknown width $\eta > 0$, with subsequent conjugate prior. Mathematically, this reads:

$$P(f_i|\eta) \sim \beta(1,\eta) \text{ and } P(\eta) \sim \eta^{-1}\,, \forall i \in [\![2, M]\!]\,. \tag{3.3}$$

Marginalization over the width hyperparameter leads to:

$$P(f_i) = -\frac{1}{(1-f_i)\log(1-f_i)}\,, \forall i \in [\![2, M]\!]\,. \tag{3.4}$$

The model is finally estimated using Metropolis Hastings MCMC sampling, which generates samples of all the parameters of the model from the posterior distribution. The i-th fascicle is subsequently discarded if its mean fraction of occupancy is below 0.05.

When translated to clinics, this method however presents some limitations. From a computational point of view, the estimation needs to be fast on a single computer. Unfortunately, [Behrens 2007] resort to MCMC simulations and lowering the number of burnin and/or effective jumps is not often compatible with the level of accuracy needed to get a usable posterior distribution. Moreover, from a practical point of view, clinical exams result in few measured DW images. This is an issue since a

small sample size introduces a strong dependence of the posterior distribution on the priors and thus makes the Bayesian information updating quite ineffective.

Following this idea of model averaging, we therefore propose to directly estimate the relative likelihood of the models given the data. The resulting method does not need a Bayesian framework and is quite effective even with small sample sizes. In the following, (i) we explain how the relative likelihood of a model (with respect to a set of models) given the data can be estimated without resorting to a Bayesian estimation scheme (Section 3.2.2); (ii) we apply the idea to the estimation of the ball-and-stick model recalled in eq. (3.1) for futher comparison with the ARD method proposed by [Behrens 2007] (Section 3.2.3). We present the results in section 3.2.4 and discuss the approach in section 3.2.5.

3.2.2 Proposed approach: explicit model averaging

3.2.2.1 Principle

The approach that we are proposing is only suited to explicitly average nested models. Consequently, it is appropriate for averaging the **same model** with **different levels of complexity**, e.g., with different numbers of fascicle compartments. However, care should be taken if one wants to use this method for averaging two different models (e.g. CHARMED and NODDI) with same or different levels of complexity. A parameter can be averaged following the proposed approach only if it shares the same physical interpretation in both models. Two models can thus be averaged if and only if each parameter of one model shares the same physical interpretation than one parameter of the other model.

For simplicity and since the goal is more to determine the correct averaged fascicle compartments rather than comparing two distinct models, we assume that we are working with the same diffusion model with an increasing number of fascicle compartments from 1 to M. Let $\{\mathcal{M}_m\}_{m=1,\dots,M}$ be a set of M nested MCMs with parameter set Θ. Our fitting procedure with real diffusion data consists in minimizing eq. (2.68) assuming an infinite SNR $\rho \to \infty$, which yields the following least square estimator of Θ provided by model $\mathcal{M}_m$:

$$\widehat{\Theta}_m(\mathbf{S}_N) = \arg\min_{\Theta} \sum_{j=1}^{N} \left(S_j - \Psi_j^{(m)}(\Theta)\right)^2 , \tag{3.5}$$

where $\Psi_j^{(m)}(\Theta)$ is the predicted DW signal from model $\mathcal{M}_m$ with parameter set Θ. We obtain a model-independent estimation of Θ with the following estimator:

$$\widehat{\Theta}_M(\mathbf{S}_N) = \sum_{m=1}^{M} \widehat{\Theta}_m(\mathbf{S}_N)\, p(\mathcal{M}_m|\mathbf{S}_N) \ , \ \text{with} \ \sum_{m=1}^{M} p(\mathcal{M}_m|\mathbf{S}_N) = 1 \ , \tag{3.6}$$

where $p(\mathcal{M}_m|\mathbf{S}_N)$ is the **relative likelihood** of model $\mathcal{M}_m$ given data $\mathbf{S}_N$, where *relative* means that likelihoods depend on the number of nested models included in the set of candidate models. Yet, we can observe that the estimator proposed in eq. (3.6) still depends on the maximum assumed number of possible distinct fascicles

in a given location in the brain. We unravel this issue in the following by showing that $M = 3$ is sufficient for the human brain, which is consistent with the current anatomical knowledge about WM fascicles [Catani 2002, Jellison 2004, Catani 2008, Thiebaut de Schotten 2011].

The estimator defined by eq. (3.6) can only be used if the relative likelihood $p(\mathcal{M}_m|\mathbf{S}_N)$ of each model $\mathcal{M}_m$ given data $\mathbf{S}_N$ is known, which is not the case in practice. We propose to estimate it using the AIC, which is a model selection criterion similar to the BIC, as described in [Posada 2004]. In the following, we first define the AIC and outline its differences with the BIC. Then, we show how the AIC can be used to estimate the relative likelihood of the models given the data.

Model selection criteria. Like the BIC, the AIC assesses the appropriateness of a model from its goodness of fit (evaluated by the fitting error) and a penalization term that increases with the complexity of the model. Both belong to the class of model selection criteria, which are essentially penalized log-likelihoods of the form:

$$-2\log p\left(\mathbf{S}_N|\widehat{\Theta}_m\left(\mathbf{S}_N\right)\right) + \text{penalty}\,. \tag{3.7}$$

Assuming that the underlying errors between the measured DW signals S_j and the predicted DW signals $\Psi_j^{(m)}$ from model $\mathcal{M}_m$ are independent and normally distributed with zero mean and variance σ^2, eq. (3.7) simplifies to:

$$N\log\left(\frac{\text{SSE}\left(\mathbf{S}_N,\mathcal{M}_m\right)}{N}\right) + \text{penalty}\,, \tag{3.8}$$

where $\text{SSE}\left(\mathbf{S}_N,\mathcal{M}_m\right) = \sum_{j=1}^{N}\left(S_j - \Psi_j^{(m)}(\widehat{\Theta}_m\left(\mathbf{S}_N\right))\right)^2$ is the sum of the squared errors.

AIC [Akaike 1974] stems from information theory. It is related to the idea of maximizing the expected likelihood for a new observation. It reads:

$$\text{AIC}\left(\mathbf{S}_N,\mathcal{M}_m\right) = N\log\left(\frac{\text{SSE}\left(\mathbf{S}_N,\mathcal{M}_m\right)}{N}\right) + 2K_m\,, \tag{3.9}$$

where K_m is the number of free parameters of model $\mathcal{M}_m$. The difference in AIC between two models $\mathcal{M}_1$ and $\mathcal{M}_2$ quantifies how much more (or less) information is lost from using $\mathcal{M}_1$ than from using $\mathcal{M}_2$ instead of the true unknown model. This is however only valid asymptotically, i.e., when the sample size is virtually infinite. Indeed, the penalty term in eq. (3.9) is linear in K_ℓ, which often produces overfitting when sample sizes are small. The corrected AIC has been introduced by [Burnham 2002] to circumvent this issue and reads:

$$\text{AIC}_c\left(\mathbf{S}_N,\mathcal{M}_m\right) = N\log\left(\frac{\text{SSE}\left(\mathbf{S}_N,\mathcal{M}_m\right)}{N}\right) + 2K_m + \frac{2K_m(K_m+1)}{N-K_m-1}\,. \tag{3.10}$$

When the sample size N is small, the penalty is amplified with a quadratic term in K_m. [Burnham 2002] strongly recommend AIC_c over AIC when the sample size

is small and, since the two criteria are asymptotically equivalent, AIC_c should be employed regardless. Note that the AIC is asymtotically equivalent to the N-fold cross-validation estimate of the generalization error.

BIC [Schwarz 1978] is related to the idea of maximizing the posterior probability of a model given the data. Generally, this problem involves the computation of complex integrals of which Laplace approximations yield the BIC that reads:

$$\text{BIC}\left(\mathbf{S}_N, \mathcal{M}_m\right) = N \log\left(\frac{\text{SSE}\left(\mathbf{S}_N, \mathcal{M}_m\right)}{N}\right) + \log(N) K_m \,. \tag{3.11}$$

Despite the similarity between the AIC and the BIC, they are of two very different natures. Only the BIC is asymptotically consistent, i.e., the probability of choosing the right model tends to one as $N \to \infty$. However, model selection criteria give no information on the performance of a model in an absolute sense (if all candidate models fit poorly, they still propose an *optimal* one). The AIC can also be derived in the same Bayesian framework as the BIC by using a prior distribution which is a decreasing function of K_ℓ instead of the non-informative prior $1/M$ used for BIC derivation. Additionally, [Burnham 2002] presented a simulation study that favors AIC_c with respect to BIC from a practical point of view.

Another advantage of AIC/AIC_c is that it can be used to obtain an estimate of the relative likelihood $p(\mathcal{M}_m|\mathbf{S}_N)$ of each model $\mathcal{M}_m$ given data $\mathbf{S}_N$ [Posada 2004].

AIC and relative likelihood of the models. The relative likelihood $p(\mathcal{M}_m|\mathbf{S}_N)$ of each model $\mathcal{M}_m$ given data $\mathbf{S}_N$ can then be approximated by [Posada 2004]:

$$p(\mathcal{M}_m|\mathbf{S}_N) \approx \exp\left(-\Delta_m\left(\mathbf{S}_N\right)/2\right) \,,\; m \in [\![1, M]\!] \,, \tag{3.12}$$

where $\Delta_m\left(\mathbf{S}_N\right) = \text{AIC}_c\left(\mathbf{S}_N, \mathcal{M}_m\right) - \text{AIC}_c^\star(\mathbf{S}_N)$ and $\text{AIC}_c^\star(\mathbf{S}_N)$ is the smallest AIC_c. One can then normalize these estimates over the set of candidate models to obtain the so-called **Akaike's weights**:

$$a_m(\mathbf{S}_N) \propto \exp\left(-\text{AIC}_c\left(\mathbf{S}_N, \mathcal{M}_m\right)/2\right) \,,\; m \in [\![1, M]\!] \text{ and } \textstyle\sum_{m=1}^{M} a_m(\mathbf{S}_N) = 1 \,, \tag{3.13}$$

which represents the probability that model $\mathcal{M}_m$ is the expected best model in the sense of the Kullback-Leibler divergence.

We can thus propose our selection-free estimator of Θ by replacing $p(\mathcal{M}_m|\mathbf{S}_N)$ in eq. (3.6) by $a_m(\mathbf{S}_N)$, leading to:

$$\widehat{\Theta}_M\left(\mathbf{S}_N\right) \simeq \sum_{m=1}^{M} a_\ell(\mathbf{S}_N) \widehat{\Theta}_m\left(\mathbf{S}_N\right) \,. \tag{3.14}$$

3.2.2.2 Application to the ball-and-stick model

The principle described in section 3.2.2.1 can be adapted to any MCM discussed in chapter 2. In the following, we apply our model averaging approach to the ball-and-stick model, which enables direct comparison with the ARD method proposed in [Behrens 2007].

The parameters of the ball-and-stick model with m fascicle compartments are:

- the orientation $\boldsymbol{\mu}_i^{(m)}$ of each fascicle ($i = 1, \dots, m$) parametrized by its polar angle $\theta_i^{(m)} \in [0, \pi]$ and its azimuth angle $\phi_i^{(m)} \in [0, 2\pi]$;
- the fraction of occupancy $f_i^{(m)}$ of each fascicle ($i = 1, \dots, m$);
- the free diffusivity $d^{(m)}$;
- the fraction of occupancy $f_0^{(m)} = 1 - \sum_{i=1}^{m} f_i^{(m)}$ of free water.

We first estimate the ball-and-stick models for $m = 0, 1, \dots, M$ by minimizing eq. (2.68) (assuming $\rho \to \infty$). Note that, for $m = 0$, the fraction of occupancy of free water is obviously $f_0 = 1$. Second, we compute the AIC_c for each model according to eq. (3.10). Finally, we generate two sets of Akaike's weights from eq. (3.13):

- a set $\{a_m^{\text{WM}}\}_{m=1,\dots,M}$ of Akaike's weights normalized using the AIC_c from $m = 1$ to $m = M$, in which a_m^{WM} is the probability that the ball-and-stick model with m fascicle compartments is the best model to fit the data in the sense of the Kullback-Leibler divergence, compared to all the ball-and-stick models that include at least one fascicle compartment (which are WM models);
- a set $\{a_m\}_{m=1,\dots,M}$ of Akaike's weights normalized using the AIC_c from $m = 0$ to $m = M$, in which a_m is the probability that the ball-and-stick model with m fascicle compartments is the best model to fit the data in the sense of the Kullback-Leibler divergence, compared to all the ball-and-stick models including the one with no fascicle compartment (which is a simple free water model).

The collection of ball-and-stick models for $m = 0, \dots, M$ is a set of nested models. Parameters $\theta_i^{(m)}$, $\phi_i^{(m)}$ and $f_i^{(m)}$, $i \in [\![1, m]\!]$, only appear in models with $m = 1, \dots, M$. To free these parameters from the dependency upon m, we shall therefore focus on averaging only models for $m = 1, \dots, M$, i.e., models that include at least one fascicle compartment. For sake of clarity, let $M = 3$.

The difficulty with these parameters ($\theta_i^{(m)}$, $\phi_i^{(m)}$ and $f_i^{(m)}$) is that their number differs in the different models: the m-fascicle ball-and-stick model contains m such parameters. Averaging a fascicle occupancy consists in picking one in each model and performing the mean weighted by the Akaike's weights: the 1-fascicle model offers a unique value, the 2-fascicle model offers 2 values and the 3-fascicle model offers 3 values. The same reasoning applies to the angles $\theta_i^{(m)}$ and $\phi_i^{(m)}$.

As a result, the averaging approach provides $3! = 6$ averaged fascicle orientations with corresponding fractions of occupancy. In other words, averaging M ball-and-stick models with an increasing number $m = 1, dots, M$ of fascicle compartments yields a big $M!$-fascicle averaged ball-and-stick MCM. This averaged model comprises $M!$ averaged fascicle orientations $\boldsymbol{\mu}_i$ and corresponding occupancies f_i ($i = 1, \dots, M!$), an averaged free diffusivity d and an averaged free water fraction of occupancy f_0. In details, these parameters are obtained as follows:

- Free diffusivity and free water occupancy: selection-free estimates d and f_0 of the free diffusivity and the fraction of occupancy of free water respectively are easily obtained from the set $\{a_m\}_{m=1,\dots,M}$ of Akaike's weights that includes the free water isotropic model; they read:

$$d = \sum_{m=0}^{M} a_m d^{(m)} \text{ and } f_0 = a_0 + \sum_{m=1}^{M} a_m \left(1 - \sum_{i=1}^{m} f_i^{(m)}\right), \tag{3.15}$$

- Fascicle orientations: we compute the 6 following weighted direction cosine matrices:

$$\begin{aligned}
\mathrm{DCM}_1 &= a_1^{\mathrm{WM}} \boldsymbol{\mu}_1^{(1)} \boldsymbol{\mu}_1^{(1)T} + a_2^{\mathrm{WM}} \boldsymbol{\mu}_1^{(2)} \boldsymbol{\mu}_1^{(2)T} + a_3^{\mathrm{WM}} \boldsymbol{\mu}_1^{(3)} \boldsymbol{\mu}_1^{(3)T}, \\
\mathrm{DCM}_2 &= a_1^{\mathrm{WM}} \boldsymbol{\mu}_1^{(1)} \boldsymbol{\mu}_1^{(1)T} + a_2^{\mathrm{WM}} \boldsymbol{\mu}_1^{(2)} \boldsymbol{\mu}_1^{(2)T} + a_3^{\mathrm{WM}} \boldsymbol{\mu}_2^{(3)} \boldsymbol{\mu}_2^{(3)T}, \\
\mathrm{DCM}_3 &= a_1^{\mathrm{WM}} \boldsymbol{\mu}_1^{(1)} \boldsymbol{\mu}_1^{(1)T} + a_2^{\mathrm{WM}} \boldsymbol{\mu}_1^{(2)} \boldsymbol{\mu}_1^{(2)T} + a_3^{\mathrm{WM}} \boldsymbol{\mu}_3^{(3)} \boldsymbol{\mu}_3^{(3)T}, \\
\mathrm{DCM}_4 &= a_1^{\mathrm{WM}} \boldsymbol{\mu}_1^{(1)} \boldsymbol{\mu}_1^{(1)T} + a_2^{\mathrm{WM}} \boldsymbol{\mu}_2^{(2)} \boldsymbol{\mu}_2^{(2)T} + a_3^{\mathrm{WM}} \boldsymbol{\mu}_1^{(3)} \boldsymbol{\mu}_1^{(3)T}, \\
\mathrm{DCM}_5 &= a_1^{\mathrm{WM}} \boldsymbol{\mu}_1^{(1)} \boldsymbol{\mu}_1^{(1)T} + a_2^{\mathrm{WM}} \boldsymbol{\mu}_2^{(2)} \boldsymbol{\mu}_2^{(2)T} + a_3^{\mathrm{WM}} \boldsymbol{\mu}_2^{(3)} \boldsymbol{\mu}_3^{(3)T}, \\
\mathrm{DCM}_6 &= a_1^{\mathrm{WM}} \boldsymbol{\mu}_1^{(1)} \boldsymbol{\mu}_1^{(1)T} + a_2^{\mathrm{WM}} \boldsymbol{\mu}_2^{(2)} \boldsymbol{\mu}_2^{(2)T} + a_3^{\mathrm{WM}} \boldsymbol{\mu}_3^{(3)} \boldsymbol{\mu}_3^{(3)T}.
\end{aligned} \tag{3.16}$$

  The principal eigenvector $\boldsymbol{e}_1$ (i.e., the one with the largest eigenvalue λ_1) of the DCMs yields the averaged fascicle orientations $\boldsymbol{\mu}_i$ ($i = 1, \dots, 6$). However, two orthogonal orientations cannot be uniquely averaged if the first two eigenvalues of the DCM are equal $\lambda_1 = \lambda_2$. We therefore retain only the DCMs for which $\lambda_2/\lambda_1 < 0.95$.

- Fraction of occupancy: the corresponding fractions of occupancy are then averaged accordingly:

$$\begin{aligned}
f_1' &\propto a_1^{\mathrm{WM}} f_1^{(1)} + a_2^{\mathrm{WM}} f_1^{(2)} + a_3^{\mathrm{WM}} f_1^{(3)}, \\
f_2' &\propto a_1^{\mathrm{WM}} f_1^{(1)} + a_2^{\mathrm{WM}} f_1^{(2)} + a_3^{\mathrm{WM}} f_2^{(3)}, \\
f_3' &\propto a_1^{\mathrm{WM}} f_1^{(1)} + a_2^{\mathrm{WM}} f_1^{(2)} + a_3^{\mathrm{WM}} f_3^{(3)}, \\
f_4' &\propto a_1^{\mathrm{WM}} f_1^{(1)} + a_2^{\mathrm{WM}} f_2^{(2)} + a_3^{\mathrm{WM}} f_1^{(3)}, \\
f_5' &\propto a_1^{\mathrm{WM}} f_1^{(1)} + a_2^{\mathrm{WM}} f_2^{(2)} + a_3^{\mathrm{WM}} f_2^{(3)}, \\
f_6' &\propto a_1^{\mathrm{WM}} f_1^{(1)} + a_2^{\mathrm{WM}} f_2^{(2)} + a_3^{\mathrm{WM}} f_3^{(3)},
\end{aligned} \tag{3.17}$$

  and normalized. These occupancies are further corrected to account for free water proportion f_0, yielding the final fascicle fractions of occupancy:

$$f_i = (1 - f_0) f_i', i \in [\![1, 6]\!], \tag{3.18}$$

  where f_0 is the averaged free water fraction of occupancy computed according to eq. (3.15).

In the following, we compare this new proposed approach to the ARD method on real clinical data.

3.2.3 Material & Methods

3.2.3.1 Description of the data

Raw data. We conducted a pilot study consisting of one normal adult brain, on which a clinical diffusion sequence was performed 10 times by the same operator on the same Siemens 3T Verio MR scanner. It comprised a single measured unweighted image B0 and 30 measured DW images acquired at the same low b-value ($b = 1000$ s/mm^2) using 30 non-collinear DSG directions uniformly distributed over the unit north hemisphere. The following parameters were used: $128 \times 128 \times 60$ image resolution with $2 \times 2 \times 2$ mm^3 isotropic spatial resolution, TR= 11 s and TE= 99 ms.

Preprocessing. We preprocessed the data as follows:

1. For each scan, we performed (i) a rigid registration of the measured DW images on the B0 to correct for subject motion and rotated the gradient tables accordingly and (ii) an affine registration of the measured DW images on the B0, initializing with the previously estimated rigid transformation, to correct for distortions.

2. We performed a rigid registration of the B0 images of the different scans on the B0 image of the first scan and applied the corresponding transformation to the subsequent measured DW images. We rotated the gradient tables accordingly. We used the FLIRT algorithm [Jenkinson 2012] for all the registration steps.

3. We filtered the data of each scan to reduce the noise in the images, using the Rician-adapted non-local means filter [Wiest-Daesslé 2008].

4. We extracted the brain in all the images using the BET algorithm [Jenkinson 2012].

3.2.3.2 Experimental setup

First, we aimed at confirming that $M = 3$ is actually sufficient in real data. In the WM, it has been shown that the fraction of occupancy of free water is very low [Zhang 2012]. In general, the AIC$_c$ favors the simplest model unless more complex models accurately fit the data. This means that our approach greatly favors the isotropic model ($m = 0$) with respect to any WM models ($m > 0$) if the **true** number of fascicles exceeds the **maximum assumed** number of fascicles M. Therefore, regions in the WM where f_0 is close to 1 indicate that M should be increased. On our experiments, we observed that such regions do not appear with $M = 4$. We thus performed the selection-free estimation of the ball-and-stick model with M ranging from 1 to 4. Regions with high f_0 values in the WM are expected to be rarer as M increases. Our goal is to show that the f_0 map obtained with $M = 3$ is very close to the one obtained with $M = 4$. To this purpose, we computed the Dice scores between the f_0 map obtained with $M = 4$ and those obtained with $M = 1, 2, 3$. The maps were binarized with ones when $f_0 = 1$ and zeros otherwise. Dice scores close to 1 indicate high similarity between the maps.

Next, we propose to compare our approach to the BEDPOSTX (BPX) algorithm, which is the implementation of the ARD method proposed in [Behrens 2007]. In

this method, only one ball-and-stick model is estimated assuming M fascicles and the ARD is performed on the fractions of occupancy f_i^{BPX}, $i \in [\![2, M]\!]$. The resulting averaged model is an M-fascicle ball-and-stick MCM, where $f_1^{\text{BPX}} \geq \cdots \geq f_M^{\text{BPX}} \geq 0$ are the fascicle occupancies. BPX only requires to fit the more complex model and, consequently, cannot provide the probability for the m-fascicle MCM to be the best model, except for the M-fascicle MCM, which is actually the only fitted one. We can thus directly compare the fraction of occupancy f_M^{BPX} of the fascicle compartment with lowest occupancy obtained with BPX to the Akaike's weight a_M of the M-fascicle ball-and-stick MCM. We also provide visualization of the other fractions of occupancy obtained by BPX and of the other Akaike's weights to give an insight into how the model averaging is performed by each method.

We therefore generated 100 bootstrap measured DW images out of the original 10 and computed these indices for $M = 3$. We then averaged them over the 100 bootstrap volumes and additionally performed an FDR-corrected t-test of nullity on f_3^{BPX} and a_3. We also computed the estimated orientations for both methods and propose a visualization of them for qualitative assessment.

3.2.4 Results

Figure 3.1 shows the f_0 maps obtained assuming a maximum number of expected fascicles equal to $M = 1$ (fig. 3.1a), $M = 2$ (fig. 3.1b), $M = 3$ (fig. 3.1c) and $M = 4$ (fig. 3.1d). We qualitatively see that $M = 3$ and $M = 4$ produce very similar maps, while those obtained with $M = 1$ or $M = 2$ contain substantial differences (more regions with high f_0 in the WM). The Dice scores between the thresholded f_0 map obtained with $M = 4$ and those obtained with $M = 1, 2, 3$ are respectively equal to 6.0×10^{-3}, 0.30 and 0.88, which quantitatively demonstrates that $M < 3$ yields to an positively biased estimates of f_0 while from $M = 3$ the f_0 does not change much and contains very few regions of high f_0 in the WM. Assuming a maximum number of expected fascicles $M = 3$ is thus sufficient, at least for clinical data.

Figure 3.2 shows the fascicle fractions of occupancy obtained by BPX with ARD (1st row) and the Akaike's weights output by our approach (2nd row). f_3^{BPX} (fig. 3.2c) and a_3 (fig. 3.2g) can be directly compared as they both quantify the presence of 3 fascicles. Qualitatively, there are few regions where f_3^{BPX} is not null but its value is always very small. On the contrary, much more regions are identified by a_3 and are in concordance with anatomical knowledge. We statistically assessed (by means of FDR-corrected paired t-tests of nullity) whether these coefficients, when not null, were significantly non-null. Figure 3.2d (resp. fig. 3.2h) shows statistically significant 3-fascicle regions detected by BPX (resp. by our approach). No such regions are detected by BPX showing that it does not robustly identify the same 3-fascicle regions from one bootstrap image to another. On the contrary, our approach shows significant 3-fascicle regions that were expected by previous anatomical knowledge. We can also observe that our approach detects few voxels where the single fascicle compartment model is predominant because the orientation averaging method is able to preserve even very small angles between them.

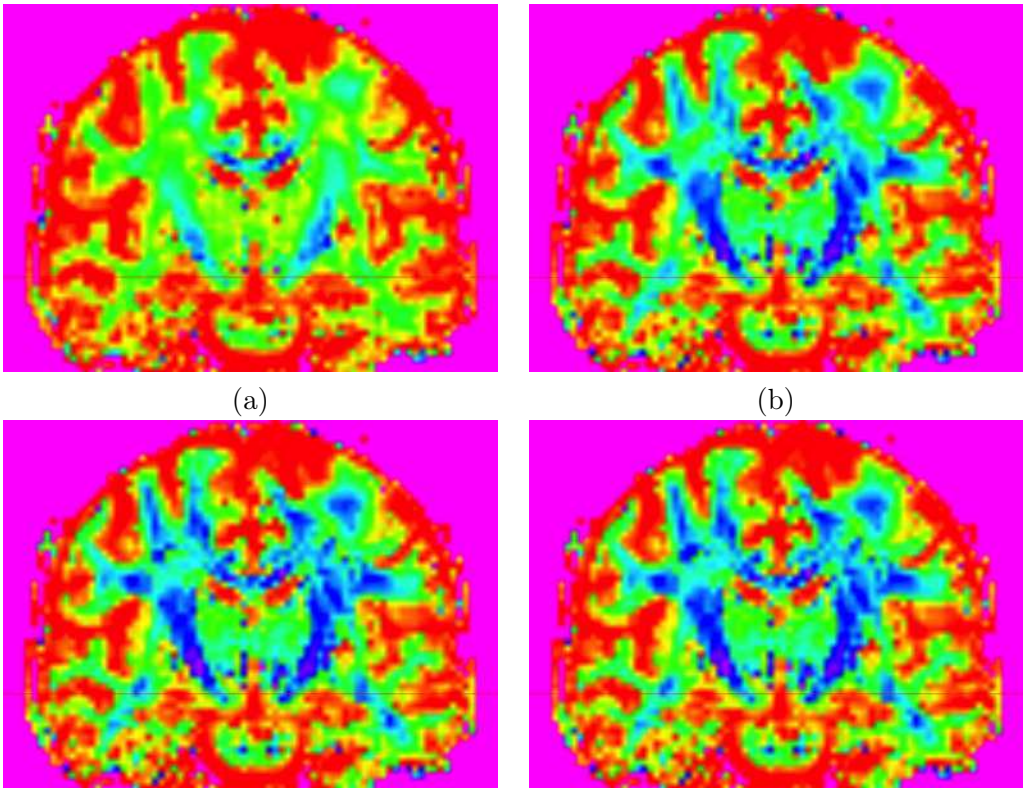


Figure 3.1: **Maps of ball-and-stick free water fraction of occupancy.** Coronal view of the f_0 map obtained applying our method to average ball-and-stick models up to $M = 1$ (a), $M = 2$ (b), $M = 3$ (c) and $M = 4$ (d) fascicle compartments. The value of f_0 is color-coded: from 0 to 1 (purple-blue-cyan-green-yellow-red). Cases $M = 3$ and $M = 4$ exhibit the same blue pattern in the WM (low f_0) while cases $M = 1$ and $M = 2$ exhibit larger f_0 values.

Figure 3.3 shows an example of visualization of the fascicle compartments. Each fascicle is represented by a cone whose axis lies on the fascicle orientation and whose angle is proportional to the corresponding fraction of occupancy. Fascicles obtained by BPX are shown in fig. 3.3a,c. Fascicles obtained by our averaging approach are shown in fig. 3.3b,d. Results are very similar, which demonstrates the appropriateness of our averaging method. Overall, our averaging approach produces more orientationally homogeneous results and better delineates 3-fascicle regions.

3.2.5 Discussion

Reliable selection of the number of fascicles in a voxel is still not a solved issue, which precludes the extensive use of MCMs in dMRI. In this section, we have proposed an explicit averaging method that proceeds in two steps: (i) estimating the relative likelihoods of the models given the data using the AIC corrected to account for

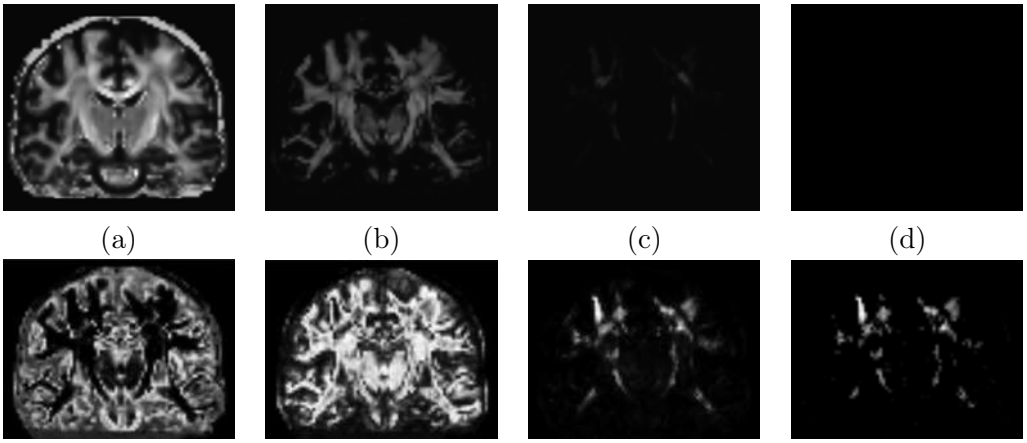


Figure 3.2: **Model averaging weights.** Coronal view of the mean fascicle fractions of occupancy f_1^{BPX} (a), f_2^{BPX} (b), f_3^{BPX} (c) and statistically significant f_3^{BPX} (d) produced by BPX and of the mean Akaike's weights a_1 (e), a_2 (f), a_3 (g) and statistically significant a_3 (h) produced by our approach. Our approach robustly detects 3-fascicle regions (h), contrary to BPX (d).

overfitting with small sample sizes and (ii) using these likelihoods to obtain model-independent estimates of the parameters. We successfully applied the method on clinical data and managed to identify regions with 3 fascicles with strong statistical support. Explicit averaging also able to keep very small angles between orientations when this is relevant. Last but not least, the computation time for $M = 3$ is about five hours for BPX and only two hours for our approach on an 8-core computer, which makes it closer to a clinically acceptable running time.

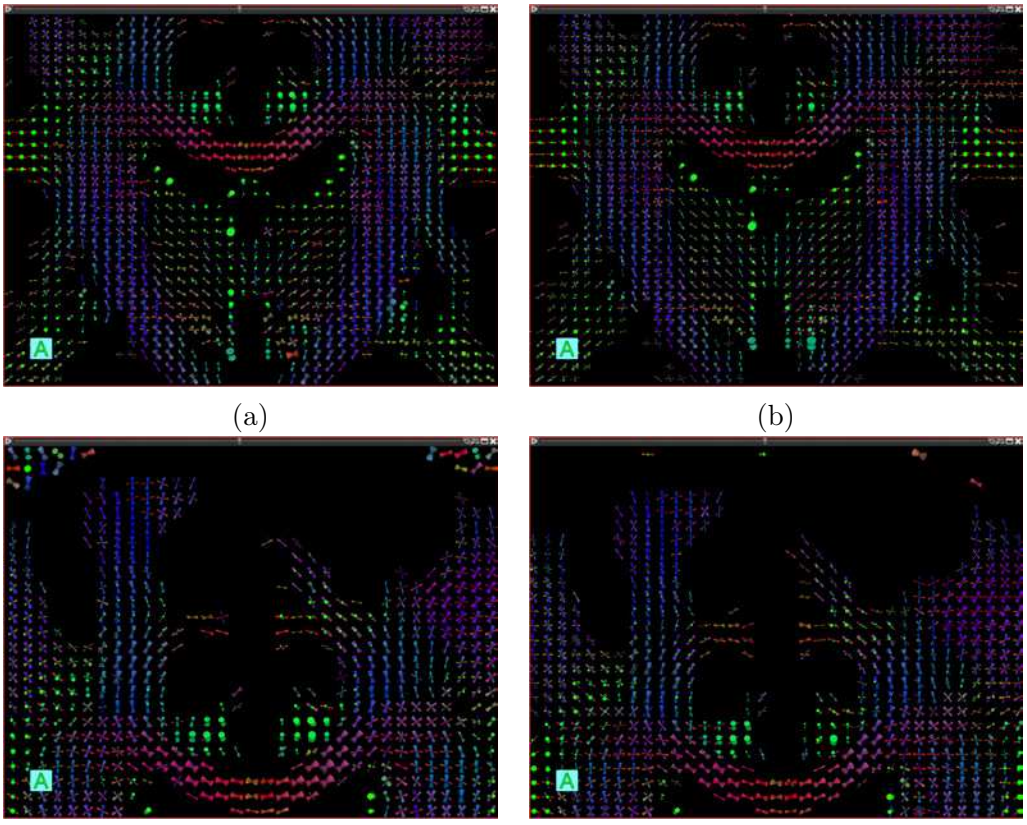
A slight limitation is that the estimated orientations in regions with more than two fascicles are strongly dependent on the initialization. In these cases, we currently perform 4 multiple restarts with a random initialization of the fascicle orientations. Increasing the number of restarts improves the detection of 3-fascicle regions at the cost of computational time.

3.3 Model averaging applied to Diffusion Directions Imaging

3.3.1 Averaging procedure

Our model averaging approach is very generic and can be applied to any MCM. Consequently, we aim at showing how it can be applied to the DDI MCM that we proposed in chapter 2, section 2.3 and how the averaging improves the results, leading to corrected versions of the usual MD and FA maps as well as correct identification of 3-fascicle regions.

The parameters of the m-fascicle DDI model with fixed diffusivities are:



(a) (b)

(c) (d)

Figure 3.3: **Fascicle representation.** Coronal view of the estimated orientations produced by BPX (first column) and our approach (second column): global view (a,b) and zoom on the Corpus Callosum and its extremities (c,d) where 3-fascicle regions are expected. Each fascicle is represented by a cone whose axis lies on the fascicle orientation and whose angle is proportional to the corresponding fraction of occupancy.

- the orientation $\boldsymbol{\mu}_i^{(m)}$ of each fascicle ($i = 1, \ldots, m$);
- the fraction of occupancy $w_i^{(m)}$ of each fascicle ($i = 1, \ldots, m$);
- the orientational dispersion index $\kappa^{(m)}$ in the fascicle compartments;
- the extra-axonal fraction of occupancy $\nu^{(m)}$ within fascicle compartments;
- the free diffusivity along the fascicles $d^{(m)}$ given by eq. (2.50);
- the free diffusivity in the CSF $d_0^{(m)} = D_{\text{iso}} = 3.0 \times 10^{-3}$ mm^2/s;
- the fraction of occupancy $w_0^{(m)} = 1 - \sum_{i=1}^{m} w_i^{(m)}$ of free water.

Before applying the averaging procedure detailed in section 3.2.2.2, we first apply it between the DDI MCM with extra-axonal fraction of occupancy constrained to

$\nu = 0$ (which is basically a ball-and-zeppelin model) and the full DDI MCM. Indeed, when dealing with clinical data at low b-value, it might be the case that we cannot observe the effect of intra- and extra-axonal diffusion on the measured DW images. In such a case, taking $\nu = 0$ is as good as any other value but enables a more accurate estimation of the remaining parameters.

Consequently, for a given number m of fascicle compartments ($m > 0$), we fit the DDI MCM with and without (ball-and-zeppelin model) extra-axonal fraction of occupancy, compute the corresponding AIC_c from eq. (3.10) and the two corresponding Akaike's weights from eq. (3.13). We average the two models accordingly (all the parameters). Since they share the same number m of fascicle compartments, the resulting averaged model also contains m fascicle compartments that are the averaged compartments obtained by matching pairwise the different fascicle compartments of the two input models. This first step is performed for each assumed number of fascicle compartments from $m = 1$ to $m = M$.

The second step consists in averaging these resulting models between $m = 0, 1, \ldots, M$ to obtain a single big DDI MCM with $M!$ fascicle compartments. The averaging is accomplished with a procedure very similar to the one described in section 3.2.2.2. The only differences are that (i) diffusivities are now set to fixed values and thus do not need to be averaged and (ii) the orientational dispersion index κ and the extra-axonal fraction of occupancy ν are both parameters that are involved in WM DDI MCMs, i.e., m-fascicle DDI MCMs with $m > 0$ and are thus averaged using the set of WM Akaike's weights $\{a_m^{\text{WM}}\}_{m=1,\ldots,M}$ (see section 3.2.2.2). In details, selection-free estimates of κ and ν read:

$$\kappa = \sum_{m=1}^{M} a_m^{\text{WM}} \kappa^{(m)} \text{ and } \nu = \sum_{m=1}^{M} a_m^{\text{WM}} \nu^{(m)} . \tag{3.19}$$

In the following section, we present some results provided by the combinated use of the DDI model estimation described in chapter 2 and the model averaging procedure detailed in this chapter on both an healthy subject (diffusion data introduced in section 3.2.3) and a patient with a tumor in presurgical planning scanned a few days before the surgery (diffusion data courtesy of 2013 MICCAI DTI Challenge).

3.3.2 Results

In section 3.2.4, we showed that, under the ball-and-stick model, it is sufficient to assume $M = 3$. In this section, since we are analyzing the diffusion data under the DDI model, we performed a similar analysis to validate that $M = 3$ is also sufficient with our proposed model. Figure 3.4 gives the free water fraction of occupancy estimated by the 1-fascicle averaged DDI model (a), the 2-fascicle averaged DDI model (b) and the 3-fascicle averaged DDI model (c). There is almost no inflated free water occupancies in the WM in the case $M = 3$ with respect to the other cases, which qualitatively demonstrates that $M = 3$ is also sufficient with the DDI model.

(a) (b) (c)

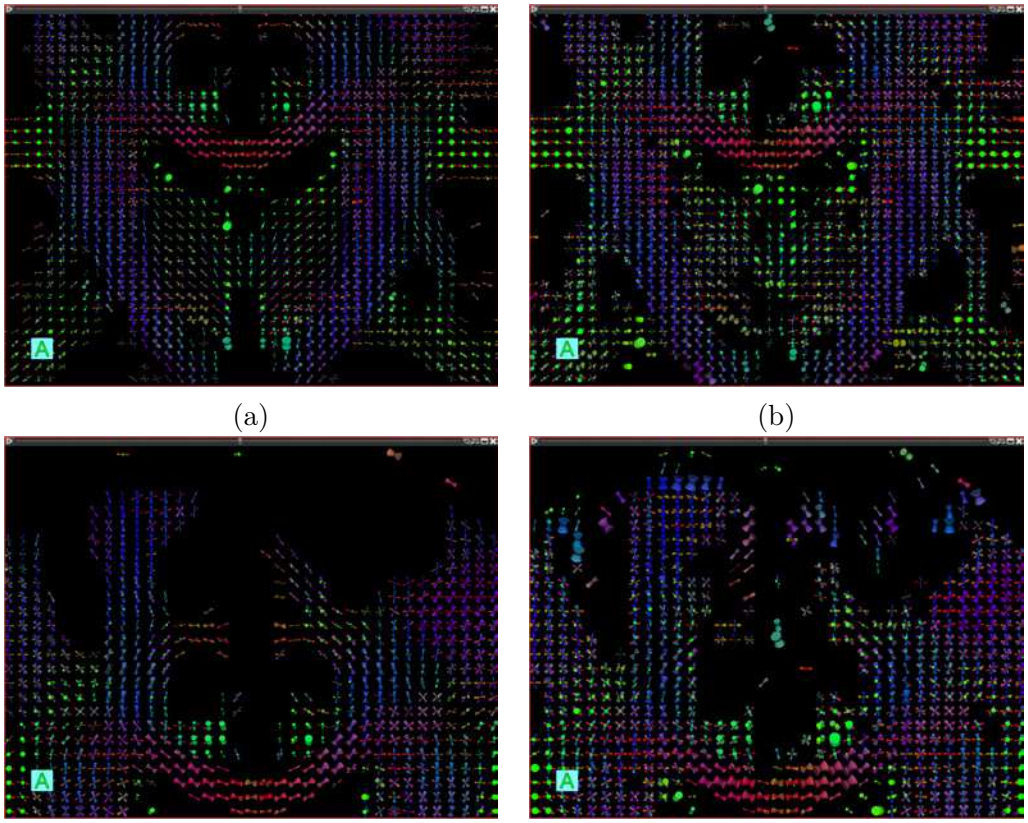
Figure 3.4: **Maps of DDI-derived free water fraction of occupancy.** Coronal view of the free water fraction of occupancy estimated by the DDI model after model averaging with $M = 1$ (a), $M = 2$ (b) and $M = 3$ (c). Large values in the WM progressively disappear in the WM and are almost absent for $M = 3$, which demonstrates that a maximum expected number $M = 3$ of fascicles is sufficient for the human brain.

3.3.2.1 Results on healthy subject

In this section, we aim at showing results obtained with the diffusion data described in section 3.2.3, which was acquired on an heathy volonteer. We performed the estimation of the m-fascicle DDI MCM, with $m = 0, 1, 2, 3$ (we thus choose $M = 3$) and subsequently averaged these models using the averaging approach proposed in section 3.2.2.

Figure 3.5 offers a visual comparison of the reconstructed fascicles provided by the ball-and-stick model and the DDI model. Both models have been estimated up to 3 fascicles and averaged according to section 3.2.2.2. Each fascicle is represented by a cone whose axis lies on the fascicle orientation and whose angle is proportional to the corresponding fraction of occupancy. Focus is on the extremities of the CC where 3-fascicle regions are expected since projection, association and commissural fascicles cross. We qualitatively see that DDI provides much more orientationally homogeneous 3-fascicle regions where expected than the ball-and-stick. Also, cone angles are systematically bigger with DDI, which reflects the fact that DDI accurately estimates a small free water occupancy in the WM unlike the ball-and-stick model that overestimates it to compensate for the over-simplified assumed geometry of the fascicles.

MCMs also provide corrected maps of all the usual scalar markers used in clinics such as MD or FA. Figure 3.6 shows some of these maps derived from the DDI model. The first row corresponds to the MD, the second row to the FA, the third row to the OD computed from eq. (2.19) and the fourth row to the extra-axonal occupancy. The last two are new indices provided by the DDI model. The first column shows the scalar maps obtained estimating the model only up to $M = 1$ while $M = 2$ in the second column and $M = 3$ in the third column. Values are color-coded from blue (smallest values) to red (largest values). We see the benefits of fitting up to $M = 3$ on the MD, FA and OD maps: MD estimates are smooth over the entire WM only for $M = 3$, FA drops to 0 (blue) in orientationally heterogeneous



(a) (b)

(c) (d)

Figure 3.5: **Conic representation of fascicles in the healthy Human brain.** Coronal view of the estimated orientations using ball-and-stick modeling (first column) and DDI modeling (second column) including averaging post-processing with $M = 3$: global view (a,b) and zoom on the Corpus Callosum and its extremities (c,d) where 3-fascicle regions are expected. Each fascicle is represented by a cone whose axis lies on the fascicle orientation and whose angle is proportional to the corresponding fraction of occupancy.

regions when $M < 3$ where a strong Orientation Dispersion (OD) is also observed, the OD is greatly reduced and corresponding FA well corrected when $M = 3$. The estimated extra-axonal occupancy is noisier but one can observe that it shows a large value (leading to negative kurtoses) in deep WM structures, corroborating the findings of [Veraart 2011b] who claimed that these negative kurtoses were due to wrong estimations.

3.3.2.2 Results on patient with a tumor

In this section, we briefly shows how the scalar maps of MD, FA, OD and extra-axonal occupancy are affected in peritumoral regions. The diffusion data has been provided by the organizers of the 2013 MICCAI DTI Tractography challenge and

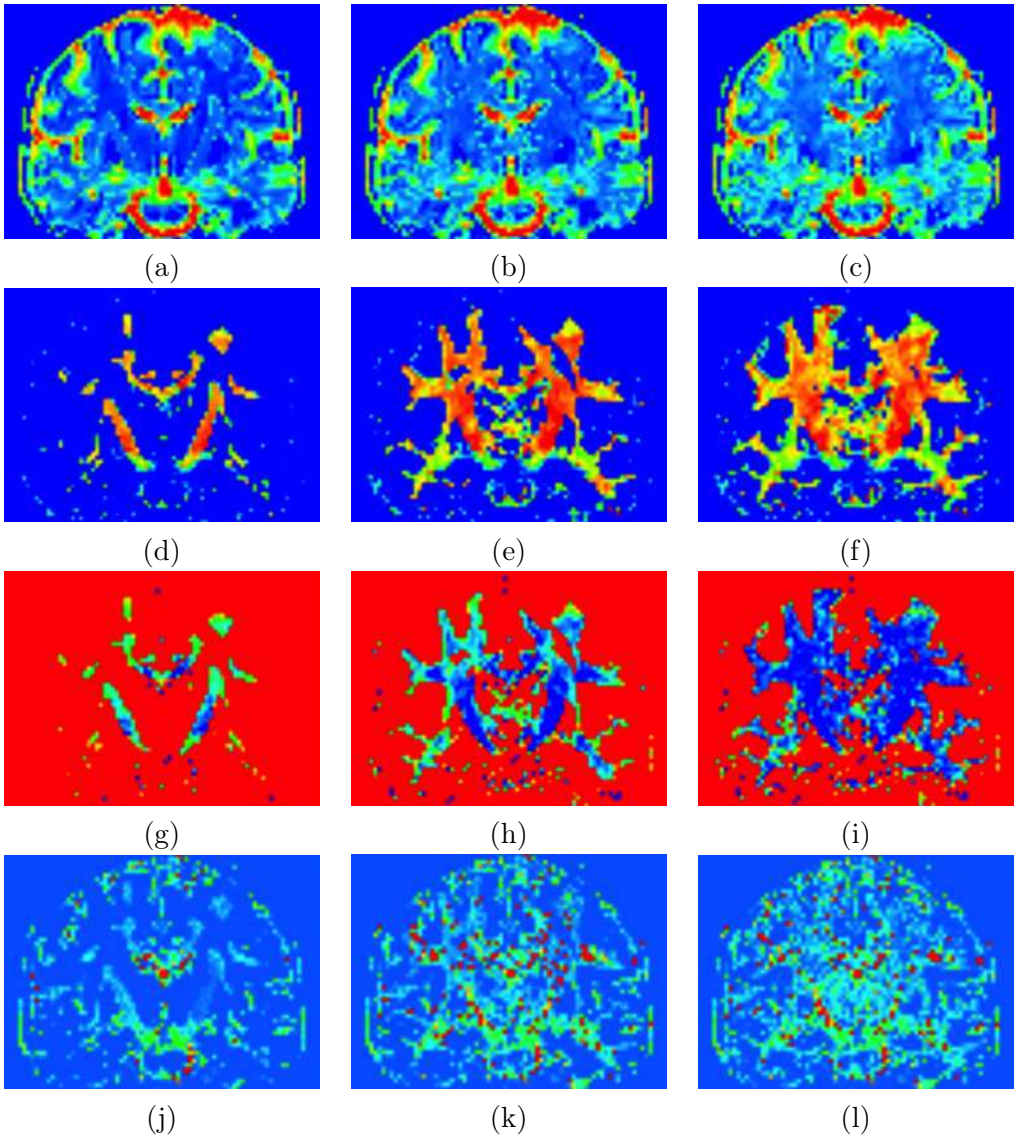
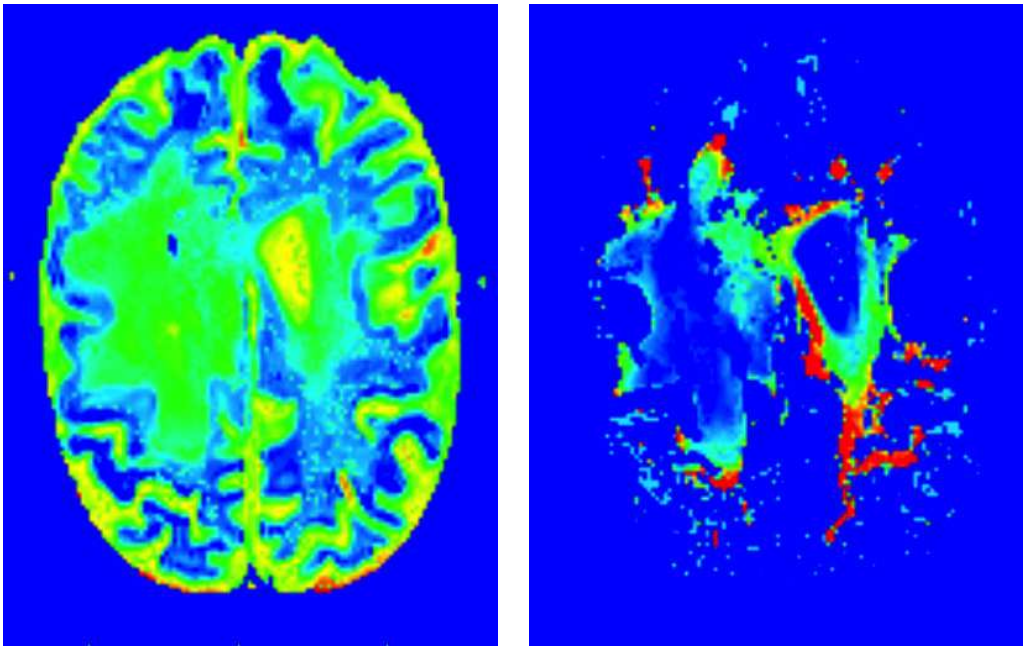


Figure 3.6: **DDI-derived scalar maps in the healthy Human brain.** Coronal view of some DDI-derived scalar maps obtained after model averaging with $M = 1$ (1st column), $M = 2$ (2nd column) and $M = 3$ (3rd column). 1st row: MD, 2nd row: FA, 3rd row: OD and 4th row: extra-axonal occupancy. Values are color-coded from blue (smallest values) to red (largest values).

consists in a presurgical scan with one b-value at 1000 s/mm^2 and only 20 DSG directions with a slice thickness of 5 mm. This is a typical example of low spatial and angular resolution diffusion data acquired in clinical settings. We performed the estimation of the DDI MCM up to $M = 3$ and averaged the models using $M = 3$.

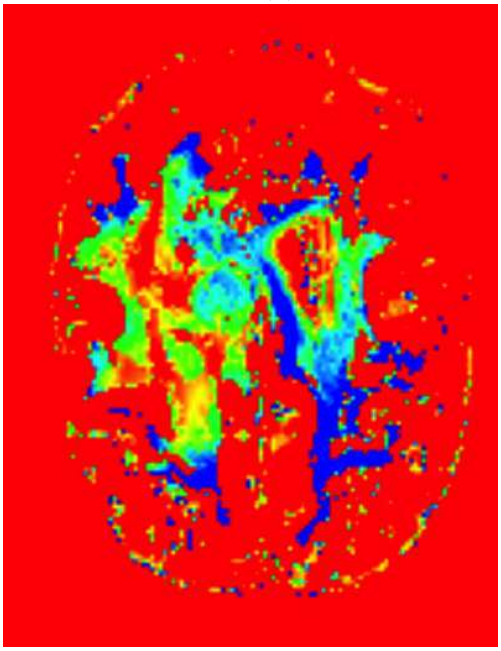
Figure 3.7 shows the MD (a), the FA (b), the OD (c) and the extra-axonal occupancy (d) provided by the averaged DDI MCM with $M = 3$. In the peritumoral region, MD is inflated ($\approx 1.67 \times 10^{-3}$ mm^2/s, green area in the peritumoral region)

with respect to the usual expected MD in the WM ($\approx 0.8 \times 10^{-3}$ mm^2/s, blue-cyan areas). The FA is also uniformly reduced to values close to 0 (blue-cyan area in the peritumoral region) whereas, in normal appearing WM, it exhibits values greater than 0.7 (green-yellow-red areas). The FA is a combined function of OD and extra-axonal occupancy (see eq. (2.44)). It is interesting to see that the uniformly reduced FA in the peritumoral region is the result of two very different combinations of these values: a very high OD (close to 1, red areas) and a low but non null extra-axonal occupancy (≈ 0.03, green areas) produce the same reduced FA as a moderate OD (≈ 0.5, green areas) and a high extra-axonal occupancy ($\approx 0.3 - 0.5$, red areas). The intuition is that maybe OD and extra-axonal occupancy may be better specific biomarkers than the FA.

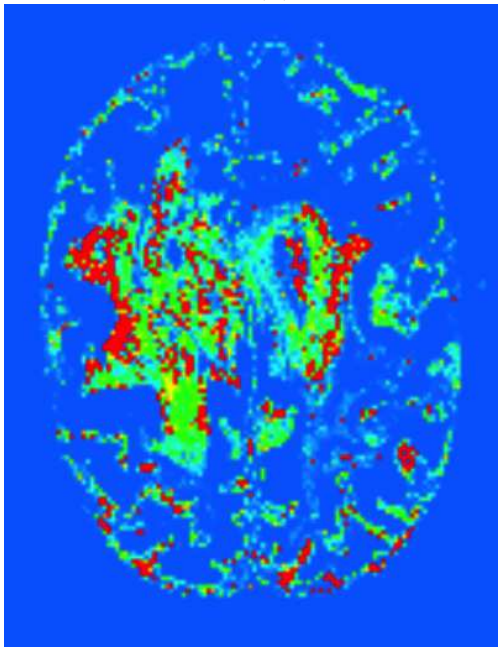


MD (a)

FA (b)



OD (c)



Extra-axonal occupancy (d)

Figure 3.7: **DDI-derived scalar maps in peritumoral region.** Axial view of some DDI-derived scalar maps obtained after model averaging with $M = 3$. Fig.(a): MD (blue: $\approx 0.8 \times 10^{-3}$ mm^2/s, green: $\approx 1.67 \times 10^{-3}$ mm^2/s), fig.(b): FA (blue: close to 0, red: close to 1, fig.(c): OD (blue: close to 0, red: close to 1) and fig.(d): extra-axonal occupancy (blue: close to 0, red: $\approx 0.3 - 0.5$).

Chapter 4

Brain White Matter tractography

Contents

The advent of dMRI and the development of DT-MRI [Basser 1994b] and subsequent more complex diffusion models to make the most out of the measured DW images offer a unique non-invasive way to study the neural processes in the human brain, including

- depicting the microstructure of the brain (geometry, size and proportion of its different component cells),
- assessing the existing interconnections between different anatomical regions in the brain.

The first application has proven to be very prolific since the beginning of the 90′ when DTI was found to be a good description of diffusion within anisotropic media. Many scalar metrics to quantify diffusivity (e.g. MD, parallel and perpendicular diffusivities) and anisotropy (e.g. FA, RA, VR) were derived from DTI analysis. Consistent maps of these scalar metrics were first obtained on the healthy human brain and, subsequently, clinical researchers found that some of these maps were really good biomarkers for a number of neurological disorders (see section 1.4.3).

The second application is often dubbed **tractography** and aims at providing a 3-dimensional reconstruction of the different WM fascicles that drive the information between the different neural centers in the CNS. Despite the ability of DTI to provide an average local orientation of diffusion, DTI tractography gained little interest at first because of its inherent complexity. For instance, if a fascicle leaving from point A reaches point B, then "should we not also find a fascicle (the same

one) leaving from B that reaches A?" In practice, it is not always the case and this simple question highlights the problem of computational cost: indeed, if one wants to ensure this symmetry property for instance, it requires a complex combination of forward and backward trackings.

After a time lag to overcome these problems, the first tractography algorithms, based on DTI at that time, appeared in the literature and generated a lot of enthusiasm and expectations. Many DTI tractography studies were published on both healthy and pathological brains (see section 1.4.4). Indeed, tractography can be very helpful to understand how exactly a specific neurological disorder affects the WM fascicles that carry the nervous information: in addition to understand whether or not the WM integrity is preserved (which is quantified by anisotropy and/or diffusivity measures), tractography could enable the identification of the damaged fascicle and quantify the severity of this alteration, or it could help to understand whether the presence of a tumor makes the measured DW signals uninterpretable in terms of fascicles or simply alters the measured DW signals to reflect a relocalization of the tracts. All these information would be of great help for further medical decisions (e.g. medications or surgical planning).

The basic idea behind tractography is to represent a fascicle as successive pairs of positions and arrival directions $(\boldsymbol{x}_k, \boldsymbol{v}_{k-1})$ starting from an initial position $\boldsymbol{x}_0$ known as the **seed voxel** (see fig. 4.1). The arrival directions, and subsequent positions

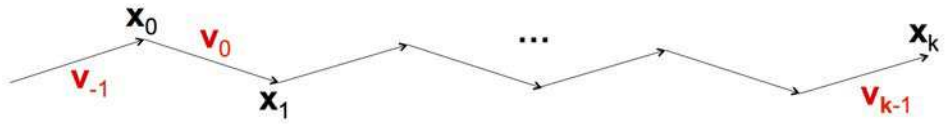


Figure 4.1: **Schematic representation of a fascicle.** A fascicle is a sequence of successive pairs of positions $(\boldsymbol{x}_k)$ and arrival directions $(\boldsymbol{v}_k)$.

visited by the fascicle, are determined by following the local orientations of diffusion provided by a diffusion model that approximates the distribution of the DIMD (or its angular profile, which suffices to extract orientations of diffusion). Two strategies are then possible: one can be interested either in connections between two voxels or ROIs, in which case only fascicles that start from one of them and end in the other are reconstructed or in understanding which areas of the brain are connected to a particular seed region. Tractography relies on three generally admitted assumptions that (i) fascicles present low curvature, (ii) they are smooth and (iii) they evolve within the WM but interconnect regions of the GM and thus should not end in the middle of the WM.

Tractography algorithms can be roughly divided into four categories: **streamline** tractography, **energy-based** tractography, **locally filtered streamline** tractography and **probabilistic streamline** tractography. Deterministic streamline algorithms reconstruct a fascicle in a stepwise fashion from a seed voxel by following the orientations identified by an underlyling diffusion model. Energy-based algorithms reconstruct a globally optimal fascicle from a seed voxel to any other

voxel by minimizing appropriate cost functions. Roughly speaking, locally filtered streamline algorithms regularize streamline fascicles by perturbing the position or orientation at each step and by keeping the more consistent one. Probabilistic streamline algorithms generalize locally filtered streamline algorithms by keeping all perturbed positions or orientations but assigning weights to the corresponding fascicle.

Deterministic streamline tractography algorithms: These algorithms were the first to appear in the literature. They reconstruct a fascicle in a stepwise fashion from a seed voxel by following the orientations identified by an underlyling diffusion model. The reconstruction is usually stopped when one of the following criteria is met:

- the fascicle reaches an area outside the brain;
- the fascicle reaches an area of low anisotropy (e.g. low FA in DTI). In this case, either it reaches an isotropic area meaning that the fascicle does end here or the local orientation mainly reflects the measurement noise and should therefore not be taken into consideration;
- the next step of the fascicle involves too much evoluation of its curvature. In this case, it may reflect the fact that another fascicle is crossing the one that is being reconstructed and the local orientation should not be considered as reliable. Indeed, it is either an orientation associated to the crossing fascicle or a powder average of the local orientations of the two crossing fascicles.

Deterministic streamline tractography algorithms can be regrouped into two general algorithms:

- *The FACT algorithm*: Starting with a voxelwise map of local orientation(s) of diffusion provided by an underlying diffusion model (e.g. the principal eigenvector of the DT) and a seed voxel, the basic concept is to follow the local orientation of diffusion from the seed voxel. The key parameter of this process is the distance that the fascicle will be allowed to travel in the direction $\boldsymbol{v}_k$ from the current position $\boldsymbol{x}_k$ to the next position $\boldsymbol{x}_{k+1}$. This distance is called the **step length** and its value determines both the curvature and smoothness of the reconstructed fascicle: the smaller the step length, the smoother the fascicles but the larger their curvature.

In order to obtain a continuous tracking, [Mori 1999, Xue 1999] proposed a method, that they coined FACT, in which the local orientation of diffusion $\boldsymbol{v}_k$ is followed linearly from the current position $\boldsymbol{x}_k$ until the fascicle quits the voxel in which $\boldsymbol{x}_k$ belongs by one of its faces. Such a process results in (i) successive positions of the reconstructed fascicle that always lie on the faces of the voxels and not inside of the voxels and (ii) step lengths that are varying. This technique produces fascicles that are only piecewise linear, which reflects poorly the assumed smoothness of the fascicle.

Smoother fascicles can be produced with the same streamline strategy by defining the step length differently: it is set either to a fixed constant [Conturo 1999] or inversely proportional to the local curvature of the fascicle [Basser 2000], i.e., when the curvature is low, large steps are authorized because the local shape is mostly linear whereas high curvature indicates more complex shapes that are better captured

using small steps.
Positions lying on the faces of the voxels make it easy to select the next local orientation to follow. The definitions of the step length proposed in [Conturo 1999, Basser 2000] allows for smaller steps that result in smoother trajectories less prone to discretization error. However, small step lengths lead to positions that are inside the voxels and not on the their faces where diffusion information is generally not available. They thus resolve the issue by interpolating the needed information from the diffusion information available in the surrounding voxels. Depending on the underlying diffusion model, the interpolation can be performed either directly on the diffusion model or on the measured DW images, in which case online estimation of the diffusion model is required at each step of the tractography algorithm. There are several methods of interpolation in the literature, the effect of which has been studied in [Lori 2002] who showed that, the noisier the measured DW images, the more dependent the tractography on interpolation.
It is often the case that WM fascicles are made of many branches that DTI misses because it can only identify one local orientation per voxel. Moreover, from a computational point of view, even if branches could be identified as well, we could not afford exploring all of them. A way to circumvent this issue while still using the DTI model to provide the local orientations has been proposed in [Conturo 1999, Stieltjes 2001]. In these works, the streamline tractography is performed from all the voxels in the brain and resulting fascicles are filtered in order to keep only those that visited the voxel of interest.
• The TENsor Deflection (TEND) algorithm: The major drawback of FACT algorithms is that they reconstruct the fascicles by following the local orientation of diffusion estimated by a diffusion model. This estimated orientation is however corrupted by various sources of noise such as inherent measurement noise in the measured DW images or crossing, kissing, merging or splitting fascicles in a single voxel. As a result, it does not always match the orientation of the fascicle. Following exclusively this estimated orientation can thus generate accumulations of errors and send a fascicle to an unexpected ending position. Also, local irregularities in the diffusion data can dramatically change the shape of the reconstructed fascicles. A first improvement has been proposed by the TEND algorithm [Lazar 2003], in which arrival direction $\boldsymbol{v}_k$ is deflected at each step by the entire DT D, which means that a tensor-deflected arrival direction is computed as $D\boldsymbol{v}_k$. Following the tensorline algorithm proposed in [Weinstein 1999], the next direction to follow ($\boldsymbol{v}_{k+1}$) is computed as a weighted linear combination of the local orientation of diffusion (e.g. the principal eigenvector $\boldsymbol{e}_1$ of the DT), the arrival direction ($\boldsymbol{v}_k$) and the tensor-deflected arrival direction ($D\boldsymbol{v}_k$). The weights of this combination are computed based on scalar metrics that differentiate prolate, oblate and spherical DT shapes [Weinstein 1999, Westin 2002, Zhang 2004].
Even if the TEND algorithm indeed reduces the accumulation of errors with respect to the FACT algorithm, the deflection is based on the estimated DT, the uncertainty of which is not taken into consideration. All the following presented algorithms aim at either circumventing the problem of accumulated errors due to noisy local orien-

tations of diffusion (energy-based algorithms) or accounting for it (locally filtered and probabilistic streamline algorithms).

Energy-based tractography algorithms. These algorithms reconstruct a globally optimal fascicle from a seed voxel to any other voxel by minimizing appropriate cost functions, which circuments to some extent the problem of accumulated errors.

• *Tractography as a fast marching problem*: The purpose of fast marching techniques is to generate three-dimensional time of arrival maps starting from a seed voxel by propagating a 3D wave front though a given volume: in essence, the front is a surface of propagation determined by a velocity function. In the context of tractography, direction and speed of such a velocity function are determinged based on the DT [Parker 2002, Prados 2006, Staempfli 2006]. Any voxel can be set as seed voxel to propagate the wave front. An optimal fascicle that connects two voxels can then be reconstructed by minimizing the travel time from one voxel to the other. Fast marching-based tractography algorithms also provide indices of connectivity between any two voxels. The use of fast marching techniques for tractography was pioneered in [Parker 2002], who used the principal eigenvector of the DT to define the direction and speed of the propagation front. The front evolves voxelwise and therefore suffers from discretization errors that [Tournier 2003] proposed to overcome with an adaptive sub-voxel evolution grid. [Prados 2006] further compute geodesic distances between two voxels in a Riemannian framework, which also enables to reconstruct optimal fascicles connecting the two voxels. Interestingly, their method does not require to actually reconstruct the fascicles to access the indices of connectivity, the computation of which is performed during the propagation of the 3D wave front. Finally, [Staempfli 2006] account for the possible presence of multiple fascicles in a voxel by defining the velocity function of the propagation front differently whether the current and next position in a fascicle reconstruction contain prolate or oblate DTs. Prolate/oblate classification of the DTs is based on the shape coefficients defined in [Westin 2002].

• *Tractography as a Hamilton-Jacobi-Bellman (HJB) problem*: The HJB framework has also been emcompassed within a tractography algorithm to robustify the tractography algorithms in presence of noise [Parker 2002, O'Donnell 2002, Jackowski 2005]. The principle of these methods is to compute the shortest fascicle that minimizes a cost function from a seed voxel to any other voxel in the brain. This cost is evaluated based on the integration of the positions visited by the fascicle and the arrival directions followed by the fascicle. The cost further penalizes fascicles that followed directions different from the ones identified by the DTs. Such a formulation leads to partial differential equations that are usually solved using propagation front methods. Consequently, fast marching-based tractography algorithms can be seen as HJB-based algorithms as well. In [Parker 2002], the propagation front is defined by a speed that increases (respectively, decreases) as the normal of the front gets closer (respectively, farther) to the principal eigenvector of the DT. In [O'Donnell 2002], the front is propagated in such a way to obtain geodesics. The DTs is used to further define a Riemannian metric that is used to penalize a connectivity index defined as the geodesic length over the Riemannian distance between the seed voxel and any

other voxel. [Jackowski 2005] further improve the algorithm of [Parker 2002], by integrating the parallel diffusivity (i.e., the largest eigenvalue of the DT) into the definition of the speed function of the front. [Pichon 2005] determine optimal fascicles between two voxels by solving the Hamilton-Jacobi-Bellman equation using dynamic programming, with an orientationally dependent cost function particularly well suited to HARDI. [Fletcher 2007] noticed that all previous HJB-based tractography algorithms dealt with the reconstruction of optimal fascicles from a single seed voxel or region. They propose an alternative algorithm that identifies fascicles interconnecting two regions, which results in a volumetric representation of the fascicles rather than in individual fascicles.

• *Tractography as a spin glass problem*: [Mangin 2002, Kreher 2008, Fillard 2009, Reisert 2011] propose to adapt the spin glass model introduced in statistical physics to the tractography problem. The word *spin* here needs not to be confounded with the spin of protons. In a spin glass problem, a spin is an oriented particle defined by its position and orientation. A spin glass is a set of spins in a closed domain. A fascicle is modeled as a chain of spins, which is equivalent to the representation proposed in fig. 4.1. Spins can thus be interpreted as pieces of fascicles. The work of [Fillard 2009] can be viewed as a improvement over the algorithms proposed by [Mangin 2002, Kreher 2008]. In this improved algorithm, each spin is endowed with three energy functions: (i) a diffusion one that tells the spin to align itself to the local orientation of diffusion, (ii) an interaction one that favors long chains of spins with low curvature and (iii) a generative one that allows duplication of spins to prevent a fascicle from ending within the WM. The previously estimated diffusion model defines the diffusion energy function and, although the authors used QBI [Tuch 2004], any diffusion model can be employed. In the recent work of [Reisert 2011], energy functions are defined differently to adapt the problem to HARDI data and to remove the necessity of boundary conditions to stop the reconstruction of the fascicles (like low FA).

• *Tractography as a genetic global optimization problem*: Another way to identify globally optimal fascicles connecting two regions is to cast the tractography problem into a genetic global optimization problem [Wu 2009], borrowing ideas from evolutionary biology, which proceeds in three stages: selection, recombination and mutation. The fascicles are assumed to be 4D curves, in which each of their three coordinates are 1D functions that are represented as a truncated Fourier series. Their Fourier coefficients are initialized randomly. First, they undergo a selection step, which retains optimal fascicles based on a trade-off between geometric smoothness and consistency with the DT field. In a second recombination step, a subset of best fascicles is kept and the rest is recombined. The resulting fascicles are slightly perturbed in a third mutation step. These three processes are repeated several times to obtain globally optimal fascicles connecting the two voxels.

• *Tractography as a network problem*: Innovately, [Lifshits 2009] use combinatorial tracking, which sees the brain as a graph in which each voxel has 26 weighted connections (the neighbouring voxels) and reconstruct fascicles by a random walk on this graph. For each voxel, the weights of its 26 connections are evaluated as

the probability of each neighbour to be connected with the current voxel. This is achieved by generating samples from the distribution of 3D molecular displacements from the center of the current voxel under DTI that assumes this distribution is a 3D centered Gaussian. Each sample is further classified as belonging to the neighboor towards which the displacement was the most collinear. Finally, the probability of each neighbour to be connected to the current voxel is estimated as the proportion of samples assigned to the corresponding neighbour over the total number of samples. In a second step, globally optimal fascicles are reconstructed either by integrating prior connectivity knowledge within a shortest path algorithm or by using an estimate of the hitting time matrix.
Tractography algorithms based on fast marching, HJB or spin glass frameworks reduce the accumulation of errors induced by following sequentially the estimated local orientations of diffusion by producing globally optimal solutions. However, diffusion information estimated from the measured DW images is still used in a deterministic way. These methods thus circumvent the issue of accumulated errors rather than accounting for orientational uncertainty. Tractography algorithms based on genetic global optimization represent the fascicles as 4D curves parametrized by the coefficients of a truncated Fourier series. The genetic global optimization framework accounts for uncertainty but on the Fourier coefficients rather than on the fascicles themselves, which lacks interpretability of such uncertainty and subsequent sources.

Locally filtered streamline tractography algorithms. These algorithms grossly regularize deterministic streamline fascicles by perturbing the position or orientation at each step and keeping the more consistent one. We term this strategy local filtering. [Gössl 2002] were the first to propose the use of a classic linear Kalman filter to locally filter streamline fascicles. At each step of the reconstruction of a fascicle, a current *observed* position is computed as the last position prolongated by the principal eigenvector of the DT. The current observed position is then seen as a noisy realization of the current expected position, which is further updated using the error between its observed value and the previous position, weighted by the Kalman gain.
Approaches based on locally filtering the arrival directions at each step have also been proposed recently [Malcolm 2010, Savadjiev 2010, Rowe 2013]. At each step of a streamline reconstruction, a number of successive directions are proposed according to a local filtering distribution and the most consistent one(s) is (are) followed. In [Malcolm 2010], the SID is represented by a finite mixure of Watson distributions (which is equivalent to a ball-and-stick model [Behrens 2007]). Parameters of the model are the orientations of diffusion with associated concentration parameter that measures the orientational dispersion. Using an unscented Kalman filter in which these parameters define the state vector, they propose perturbed versions of the state that they use to predict new SIDs. Errors between these new SIDs and the measured one, weighted by the Kalman gain, are finally used to update the estimates of the state and the direction with highest concentration is selected to proceed the fascicle to its next position. Similarly, [Savadjiev 2010, Rowe 2013] employ a particle

filter to approximate the local filtering distribution. In [Savadjiev 2010], local filtering is defined on the distribution of so-called 3D *streamline flows*, which are locally spatially-varying according to a given differential geometry that defines the local orientation and shape of the flow. At each step of a streamline reconstruction, the maximum *a posteriori* streamline flow is computed and a local deterministic streamline tractography is performed from the resulting flow in a small neighborhood. In [Rowe 2013], particle filtering is locally performed on orientations generated from the orientations of diffusion provided by the NODDI model.
The idea of a filtering distribution to obtain more reliable fascicles is very appealing but such local approaches focus on locally filtering either the positions or the arrival directions. The quantities of interest in these algotihms are thus more the individual positions visited by the fascicle or the directions it followed rather than the fascicle itself.

Probabilistic streamline algorithms. A common problem of deterministic streamline tractography algorithm is that they do not account for the orientational uncertainty inherent to the estimation of the local orientations of diffusion [Jones 2003]. This uncertainty can be due to measurement noise [Gudbjartsson 1995] and/or partial volume averaging (e.g., crossing, kissing, diverging, bifurcating fascicles or CSF contamination in a single voxel) [Alexander 2001]. Energy-based tractography algorithms cope with this problem by finding globally optimal fascicles. This circumvents the problem but does not really account for the orientational uncertainty. Locally filtered tractography algorithms introduce a filtering distribution that allows one to locally identify directions or positions that are more consistent than the ones provided by the diffusion model. However, the filtering is not performed on the fascicles themselves, which means that such algorithms do not produce fascicles that represent the distribution of all possible fascicles emanating from a single voxel.
On the contrary, probabilistic streamline algorithms are particularly appealing for integrating this uncertainty. The prefix *probabilistic* stems from the fact that resulting fascicles are weighted, which measures the strength of that specific connection to the seed voxel. These methods proceed in two stages. First, the uncertainty in local orientations of diffusion (and/or of other diffusion parameters) is modeled using an appropriate PDF, which is referred to as the model uncertainty PDF in the following. Then, the model uncertainty PDF is used to sample many orientations of propagation and streamline reconstruction of fascicles is carried out for each sampled orientation, assigning suitable weights to each fascicle.
The first work that formulated the model uncertainty PDF was [Behrens 2003]. The posterior PDF of the DT parameters given the measured DW images is theoretically computed according to Bayes' rule. Facing the intractability of the involved integrals, they propose a complete Bayesian framework for the estimation of the underlying diffusion model. The framework is solved by resorting to an MCMC procedure that produces samples of each diffusion parameter according to its marginal posterior distribution using a Gibbs sampler. The diffusion model is a ball-and-stick model with one fascicle compartment. The framework is first improved in [Behrens 2007] by

estimating a ball-and-stick model with several fascicle compartments and retaining only those supported by the data according to automatic relevance detection. A second improvement over the framework of [Behrens 2007] is achieved in [Jbabdi 2007] by incorporating global connectivity knowledge in addition to local diffusion information to perform tractography. The idea is that, wherever the local diffusion model fitting performs poorly and provides inconsistent diffusion information (e.g. biased local orientations of diffusion), global connectivity knowledge will substitute local diffusion information and robustify the reconstruction of the fascicles. In [Jbabdi 2007], a complete Bayesian framework for the estimation of a ball-and-stick MCM is derived, in which priors on the parameters are different whether the tangent direction of the fascicle lies on a pathway known to connect the two regions of interest or not. A Metropolis-Hasting MCMC algorithm is then used to compute the posterior distribution of the parameters. Finally, they use the orientations obtained in the posterior distribution to sample similar orientations according to a proposal distribution built in a pre-processing step without integrating the global connectivity knowledge. Fascicles are then reconstructed from these orientations. To avoid the use of a computationally heavy MCMC procedure, [Friman 2006] proposed a similar Bayesian framework in which the model uncertainty PDF is evaluated on a set of uniform samples on the 2-sphere.
Efficient sampling according to the model uncertainty PDF is crucial in these methods. MCMC procedures were considered but require unaffordable computational time. Evaluation of the PDF on a small sample on the 2-sphere provides too low angular resolution. Alternative algorithms have thus been proposed to circumvent these problems by formulating the tractography problem in such a way that the Gaussian PDF can be used to generate appropriate samples [Parker 2003b, Parker 2005, Berman 2008]. In the first step, these algorithms start by estimating the local orientations of diffusion, using an adequate diffusion model such as DTI [Parker 2003b], Persistent Angular Structure (PAS)-MRI [Parker 2005] or QBI [Berman 2008]. The second step consists in generating samples (simulated measured DW images) by simulating the SID using a *perturbed* diffusion model. [Parker 2003b] define a zero-mean 1-dimensional Gaussian PDF whose standard deviation is a function of the DT parameters that they use to generate slightly perturbed orientation samples around the local orientation of diffusion. The rotation that align the local orientation of diffusion to the sampled one is then applied to the DT and the rotated DT is used to generate SIDs. The work is further improved in [Parker 2003a] to account for multiple fascicles. [Parker 2005] come up with a mixture of Gaussian PDFs with cylindrically symmetric fixed-shape DTs aligned on the local orientations found by the PAS and use it to generate the SID, which is further corrupted by artifical Rician noise. A single component is used in the mixture if the underlying DT is prolate and two components are used when the DT is oblate. [Berman 2008] use the estimated dODF SH coefficients from QBI and the residuals of their estimation to generate bootstrap samples (simulated measured DW images). In a third step, diffusion models are estimated for each sampled measured DW image and the final step performs a deterministic streamline tractography for each sampled measured

DW image. The strength of connectivity of given voxel is evaluated as the number of tracts that visited it over the number of sampled measured DW images.
However, previous algorithms employed the Gaussian distribution to produce orientation samples because it is very handy to sample from it but it poorly models the uncertainty on orientations of diffusion. Indeed, when sampling from the Gaussian distribution, only perturbed SIDs can theoretically be generated, since orientations are not evolving in the 3-dimensional Euclidean space. Distributions suited to orientational data analysis have thus been proposed: [Cook 2004] captures the orientation uncertainty by a Watson distribution, [Kaden 2007] uses a mixture of Bingham distributions estimated by spherical deconvolution and [Jeurissen 2011] uses a fODF, whose SH coefficients are estimated using constrained spherical deconvolution as proposed in [Tournier 2008].
Nevertheless, all the previously described probabilistic streamline tractography algorithms assume that the ten of thousands of generated fascicles have the same weight and a probability index of connectivity is defined between any voxel and the seed voxel as the number of fascicles that visited the voxel over the number of trials. The generated fascicles are thus still not a representative sample of the distribution of all possible fascicles emanating from the seed voxel. This distribution is usually called the **filtering distribution** and can be well approximated by particle filters. The use of particle filters for tractography was pioneered by [Brun 2002, Björnemo 2002] and enables the reconstruction of fascicles by sequential importance sampling. After [Brun 2002] introduced the idea that particle filtering might help in obtaining a reliable reconstruction of fascicles, [Björnemo 2002] implemented a particle filter in which orientations of diffusion are sampled according to a Gaussian distribution and fascicle weights are assumed constant thoughout the sequential sampling. As a result, this first proposed particle filter for tractography presents the same drawbacks as the other probabilistic streamline algorithms. Recently, more elaborate particle filters have been proposed in the literature [Zhang 2009, Pontabry 2013]. The common foundation of these algorithms is that they resort to the vMF distribution to model uncertainty on the local orientations of diffusion, which is a suitable distribution for directional data that can be efficiently sampled fast [Ulrich 1984, Wood 1994, Jakob 2012]. The two algorithms differ by the choice of the diffusion model that is used to extract the diffusional information out of the measured DW images. In [Zhang 2009], the DT model is used, distinguishing prolate from oblate DTs based on the linear coefficient of the tensor given in eq. (1.21) [Westin 2002]. The weights of the fascicles are updated at each step of their reconstruction, with different update equations whether the DT is prolate or oblate. The particle filter proposed by [Pontabry 2013] relies on QBI [Tuch 2004]. When fitted to adequate diffusion data (high b-value and HARDI), such a diffusion model is able to locally detect multiple orientations of diffusion. Consequently, a single update equation is defined for the fascicle weights and the local detection of multiple orientations help in capturing the multi-modality of the filtering distribution.
However, a common problem of classical particle filters is that, they have hard

time in maintaining the multi-modality of the filtering distribution. In the context of tractography, this means that they will hardly be able to track multiple fascicles over extended volumes, regardless of the capability of the underlying diffusion model to identify correctly multiple orientations locally. This effect is stressed by the resampling of the fascicles that is occasionnally performed in the process. Indeed, in order to avoid degeneracy of the particle weights, a resampling step is sometimes necessary to keep only the most relevant particles [Doucet 2000]. After this step, bifurcating sub-fascicles are often discarded.

In the tractography literature on deterministic streamline algorithms, we noticed that, since these algorithms (namely FACT and TEND) were developed when only DTI was available to extract local orientations of diffusion, there has been no attempt to combine diffusion models that are capable of identifying multiple orientations with FACT or TEND algorithms. This could be very promising since it could avoid having to perform full brain tractography, which is prohibitively time-consuming, even with fast DTI estimation. In section 4.1, we will therefore propose an adapated FACT algorithm to account for multiple local orientations of diffusion. The behavior of the proposed algorithm will then be compared using three different diffusion models, namely DTI, QBI and our proposed DDI MCM.

Among the most recent tractography algorithms, it seems that particle filtering on the distribution of all possible fascicles emanating from a seed voxel is one of the most promising techniques to account for orientational uncertainty while focusing on the fascicles themselves. However, even if [Pontabry 2013] made significant improvements, classic particle filters intrinsically have hard time maintaing the multi-modality and thus are unlikely to depict well all the birfucations of many fascicles in the brain [Catani 2002, Jellison 2004, Catani 2008, Thiebaut de Schotten 2011]. In section 4.2, we will thus describe how the particle filter works and how it can be improved to better maintain the multi-modality of the filtering distribution. The adapative multi-modal particle filter for tractography is implemented with both DTI and QBI, which allows for a comparison with their mono-modal counterparts proposed respectively by [Zhang 2009] and [Pontabry 2013]. The benefits of our proposed DDI MCM will also be shown through this probabilistic tractography algorithm.

4.1 Multi-modal FACT algorithm

4.1.1 Introduction

dMRI [Le Bihan 2003] allows in vivo and non-invasive imaging of tissue structure. It is based on the facts that i) the diffusion of water molecules is constrained by the micro-structure of the tissues (such as, typically, the WM fascicles in the brain), and ii) MRI can be made sensitive to this diffusion, using specific *diffusion* sequences. Diffusion models can then be devised, and their parameters can be estimated for further study and analysis of tissue architecture. The simplest model is that of a Gaussian EAP, which amounts to characterizing the diffusion with a tensor (*i.e.* a

3×3 symmetric definite positive matrix), giving its name to DTI [Basser 2002].

Tractography has been developed to "reconstruct" or "dissect" the fascicles in vivo, and then infer brain anatomy [Mori 2002]. Many of the association (*e.g.* uncinate fasciculus, cingulum) and commissural (*e.g.* transverse fascicles of the corpus callosum) fascicles have been successfully reconstructed using i) clinical dMRI sequences (with few gradient directions and low b-values), ii) the simple Gaussian diffusion model and iii) simple deterministic streamline tractography algortihms [Mori 2005]. On the contrary, it has proved much more difficult to reconstruct projection fascicles, and especially the motor tracts of the CST using such standard protocols. On one side, the anatomy of these tracts between the spinal cord and the internal capsule has been well-studied using DTI [Holodny 2005, Kim 2008, Hong 2010]. On the other side, the study of these tracts between the internal capsule and the cortex, and in particular of those dedicated to a specific motor function [Han 2010, Seo 2012], is much more challenging using DTI, mostly due to the numerous crossings, kissing, merging or diverging fascicles in the corona radiata. A particularly difficult tract to reconstruct within the CST is the portion corresponding to the motor hand area, because it is located laterally on the motor cortex [Yousry 1997] (as shown by the homunculus of Penfield & Rasmussen), compared to the leg or the trunk areas for instance. In the following, we call this tract the Hand Motor Tract (HMT).

The HMTs are a crucial part of the CST to investigate, in the context of normal anatomy, within the more general study of handedness, cerebral dominance, and brain asymmetry [Toga 2003]. More generally, the development of diffusion models and tractography methods for the CST in general, and the HMTs in particular, which could be used in clinical routine, is key for a better understanding of pathologies of the CST such as, typically, ALS [Iwata 2008], Wallerian degeneration of the CST after ischemic stroke [Yu 2009], motor dysfunctions in infants [Ludeman 2008] or in patients with relapsing-remitting MS [Lin 2007].

New types of image acquisition schemes (*e.g.* HARDI sequences), diffusion models (*e.g.* multiple tensors, QBI, *etc.*) [Lenglet 2009, Assemlal 2011] and tractography methods [Mori 2002, Mukherjee 2008a, Mukherjee 2008b] have been introduced to account for intricate tract configurations, but these techniques i) have been reported to often miss entirely the lateral portions of the CST, and thus the HMTs [Behrens 2007] and ii) are not applicable at hand in a clinical setting, mostly due to long acquisition times.

In this section, we propose to investigate the usefulness of our proposed diffusion model, termed DDI and introduced in section 2.3, to reconstruct the HMTs on clinical data, using a deterministic streamline tractography algorithm. This model allows the capture of several diffusion directions within a voxel, with a low number of parameters.

The two goals of this section are: 1) to evaluate the ability of this new diffusion model and of two other standard models (DTI/QBI) to reconstruct the left and right HMTs, using a common tractography algorithm, in a clinical setting, *i.e.* with **few diffusion gradients** (typically, less than 15) and **low b-values** (typically, less

that 1000 s/mm^2), and 2) to study whether the number of reconstructed HMTs in one hemisphere is different from that in the contralateral hemisphere, for right-handed and left-handed subjects. Note that we do not test multi-tensor models, as these have been shown to be unable to provide a unique solution in the context of single-shell (one unique b-value) acquisitions [Scherrer 2010], as is the case here and in most standard clinical protocols.

In section 4.1.2.1, we present the three tested models, and we outline our implementation thereof. In section 4.1.2.2, we describe the common multi-modal deterministic streamline tractography algorithm we use for these three models. The same algorithm is used to make sure that the subsequently reported results can be interpreted as **differences between models**, rather than differences in dMRI sequences or tractography algorithms. The data are described in section 4.1.2.3, and we perform statistical tests and numerical evaluation in section 4.1.3, before discussing these results, concluding and giving some perspectives in section 4.1.4.

4.1.2 Material & Methods

4.1.2.1 Diffusion models

DT model : It assumes that the diffusion process can be captured by a tensor (6 parameters), which is proportional to the covariance matrix of the unknown Gaussian PDF. The DT D is parametrized as $D = \exp(M)$, where M is an unknown 3×3 symmetric matrix, and its estimation is done using a least square fitting on the measured DW signals [Fillard 2007]. The least square criterion is optimised numerically using the NEW Unconstrained Optimization Algorithm (NEWUOA) algorithm [Powell 2006]. Within a given voxel, the single *putative* fascicle orientation is considered to be aligned with the orientation of the eigenvectors associated to the largest eigenvalue of the DT. The tractography algorithm uses a log-Euclidean interpolation scheme [Arsigny 2006].

Q-ball model : The dODF describes the orientational structure of the EAP [Tuch 2004]. The measured DW signals are modeled with a modified basis of spherical harmonics, whose c coefficients are estimated using a least square fitting including a Laplace-Beltrami regularisation term. The number of unknown coefficients depends on the order l of the basis: $c = (l+1)(l+2)/2$. This least square problem has a closed-form solution, from which the **optimal** dODF (or to be precise, an approximation thereof) can be computed analytically using the Funk-Hecke theorem [Descoteaux 2007]. Then, dODF sharpening is performed using spherical deconvolution to compute the fODF [Descoteaux 2009]. Within a given voxel, the putative fascicle orientations are selected as the local maxima of the normalized fODF for which the fODF value is above a user-specified threshold set here to 0.1. These local maxima are computed using NEWUOA (with starting points homogeneously distributed on the unit sphere), and they are sorted according to their fODF value. The tractography algorithm uses a trilinear interpolation scheme [Lenglet 2009].

DDI model : The EAP is modeled by the full DDI MCM introduced in sec-

tion 2.3.4.1, to which we applied some constraints that are however different from those applied to the constrained DDI MCM proposed in section 2.3.4.3. The number of mixture components is equal to the number m of different fascicle orientations within the voxel. The extra-axonal occupancy in each fascicle compartment is set to $\nu_i = 0.5$. The free diffusivity along each fascicle d_i is assumed to be related to the concentrations κ_i via $d_i = (\kappa_i + 1)\lambda$, where λ is a diffusivity common to all fascicle compartments. The diffusivity in the CSF is assumed to be equal to λ and finally the fractions of occupancy of the fascicle compartments are given by $w_i = \mathrm{FA}_i^{\mathrm{approx}}/m$, with $\mathrm{FA}_i^{\mathrm{approx}}$ being:

$$\mathrm{FA}_i^{\mathrm{approx}} = \frac{\kappa_i}{\sqrt{(\kappa_i + 1)^2 + 2}}, \tag{4.1}$$

based on the exponential part of the DDI PDF, which resembles a Gaussian distribution with DT eigenvalues $\lambda_1 = d$ and $d/(\kappa+1)$ yielding the FA expression given by eq. (4.1). Therefore, to allow for m fascicles, this constrained DDI model requires $3m+1$ parameters. The approximated FA given by eq. (4.1) is akin to the FA (resp. the generalized FA [Tuch 2004]) in the DT (resp. Q-ball) model. The $3m+1$ unknown parameters are estimated using a least square fitting on the measured DW signals, and this optimization is performed using NEWUOA. Within a given voxel, the putative fascicle orientations are a natural output of this model, sorted according to their EAP values. Note that, as of now, the number of fascicles m in this model is set to 2, since the model averaging tool described in section 3.2.2 was not available at this moment. The tractography algorithm uses a trilinear interpolation scheme.

4.1.2.2 The common multi-modal deterministic streamline tractography algorithm

Our goal is to reconstruct the fascicles connecting two ROIs. Our deterministic streamline algorithm can be viewed as an extension of the original FACT algorithm [Mori 1999], adapted to QBI and DDI, using a *breadth-first*-type search. It must be made clear that for DTI, the tractography is led **without** considering multiple orientations; we omit this important detail below for the sake of clarity.

Starting from one of the two ROIs, we define n starting points within each voxel of the ROI. The DT, Q-ball and DDI models at these starting points are estimated using the previously described interpolation schemes. For each of these starting points, we compute the two *principal* putative orientations (defined using the previously described sorting out procedures), we follow the *first* orientation with a step size of l millimeters and we record the *second* one for future use, as it can be indicative of crossing, kissing, merging or diverging tracts. We then reestimate the DT, Q-ball and DDI models at this new spatial position (using the previously described interpolation schemes), and we compute *all* the putative orientations for each model. Among these, we follow the one closest (*i.e.* with minimal angular difference) to the previously estimated first orientation. A second orientation, having the highest fODF/EAP value among the remaining putative fascicle orientations, is

recorded for future use. The reconstruction of the *main* tract is achieved when i) the angle between two successively estimated first orientations is higher than α, or when ii) FA/generalized FA/ξ is lower than β, or when iii) the tract reaches the border of a precomputed brain mask [Smith 2002]. The algorithm for the reconstruction of the main tract can be summarized as follows:

Starting from one of the ROIs, we define n starting point(s) at each voxel of the ROI. Given one point along its path, we build the *main* tract iteratively as follows:

1. If the number of putative fascicle orientations $m = 0$, we stop the reconstruction.
2. If $m = 1$, we compute FA_1 and the angle α_1 between the input direction and orientation $\pm\boldsymbol{\mu}_1$. We ensure the inner product between the input direction and $\boldsymbol{\mu}_1$ is positive: if not, we apply a 180° rotation to $\boldsymbol{\mu}_1$. If $\alpha_1 < \alpha_t$ and $\text{FA}_1 > \text{FA}_t$, then we follow the single putative fascicle direction $\boldsymbol{\mu}_1$ with a step size of l millimeters. Else, we stop the reconstruction.
3. If $m = 2$, we compute FA_1, FA_2, the angle α_1 (resp. α_2) between the input direction and $\boldsymbol{\mu}_1$ (resp. $\boldsymbol{\mu}_2$). We ensure the inner product between the input direction and $\boldsymbol{\mu}_1$ (resp. $\boldsymbol{\mu}_2$) is positive: if not, we apply a 180° rotation to $\boldsymbol{\mu}_1$ (resp. $\boldsymbol{\mu}_1$). If:
   - $\alpha_t < \alpha_1, \alpha_2$: we stop the reconstruction.
   - $\alpha_1 < \alpha_t < \alpha_2$: cf. the case $m = 1$.
   - $\alpha_2 < \alpha_t < \alpha_1$: if $\text{FA}_2 > \text{FA}_t$, then we follow the direction $\vec{\mu}_2$ with a step size of l millimeters, else we stop the reconstruction.
   - $\alpha_1, \alpha_2 < \alpha_t$: if $\text{FA}_2 < \text{FA}_t$, then cf. the case $m = 1$; else if $\kappa_2 > r \times \kappa_1$ then we sort the two tracts in ascending order according to the angles α_i. We follow the new direction $\boldsymbol{\mu}_1$ with a step size of l millimeters and we record the second putative fascicle direction $\boldsymbol{\mu}_2$ (*branch*) for future use, as it can be indicative of crossing, kissing, merging or diverging tracts.

This is the global description of the algorithm, in which the switch on the number m of putative fascicle orientations is used to generalize the algorithm to any diffusion model. For instance, with the DT model, only the case $m = 1$ is relevant. With the Q-ball model, cases $m = 1, 2$ can occur. With the DDI model, only the case $m = 2$ is relevant since no model selection is performed in this analysis.

Figure 4.2 illustrates the process described above for reconstructing a tract with bifurcations. Starting from a single starting point in a given voxel, the main tract is followed as long as the diffusion model proposes a single ($m = 1$) fascicle orientation (Figure 4.2a). At some point, the tract reaches a location at which the model proposes two fascicles orientations (Figure 4.2b). The main tract is first prolongated

by following the green one, which is more collinear to the incoming direction than the orange one (Figure 4.2c). Later, the main tract reaches another location at which the model proposes two fascicle orientations (Figure 4.2d). Since no stopping criterion has yet been met, the main tract is still prolongated by following the green one, which is more collinear to the incoming direction than the orange one (Figure 4.2e). At some point, one or more of the stopping criteria is/are met and the reconstruction of the main tract stops. Bifurcating directions previously recorded **along the main tract** are then explored in order of appearance (Figure 4.2f). If all locations explored during the reconstruction of a bifurcation contain only one fascicle orientation then the bifurcation is stopped when a stopping criterion is met (Figure 4.2g). Else, if, during the reconstruction, a bifurcation reaches a location with two fascicle orientations, the green one is followed (more collinear to incoming direction) and the orange one is not recorded anymore since it is not a bifurcating direction along the main tract (Figure 4.2h). Figure 4.2i shows the complete reconstruction using our adapted multi-modal FACT algorithm and fig. 4.2j shows the complete reconstruction using the original FACT algorithm.

Once this main tract has been reconstructed, we perform the same reconstruction from all the possible crossing, kissing, merging or diverging points that have been recorded along its path. Importantly, for these reconstructions, the stepping rule and stopping criteria are identical as those for the main tract, but we do not record any possible mixed tract configuration along these secondary paths, for which we only follow the first orientation at each step. The same reconstruction pattern is then led on the second ROI, and only the tracts linking the two ROIs are kept for further analysis. In practice, we choose the parameters $n = 1$, $l = 1$, $\alpha = 60°$, $\beta = 0.15$ and $r = 0.8$.

4.1.2.3 Data

The data consist of dMRI, Anatomical MRI (aMRI) and Functional MRI (fMRI) on 14 right-handed (8 males, 6 females) and 9 left-handed (6 males, 3 females) healthy volunteers. The mean age was 30.3 (21 to 45). Handedness was determined using the Oldfield questionnaire [Oldfield 1971]. The aMRI, dMRI and fMRI data were acquired using standard sequences on a Philips Achieva 3T system:

- aMRI: T1-w 3D Turbo Field Echo (TFE), 184 sagittal slices of size 256×256 ($1 \times 1\text{mm}^2$) and 1mm thickness.
- fMRI: gradient echo Echo Planar Imaging (EPI) using Blood Oxygen Level Dependent (BOLD) contrast, 24 contiguous axial slices of size 128×128 ($1.8 \times 1.8\text{mm}^2$) and 4mm thickness. The hand motor task consisted in opening and closing the hand, and was implemented in a standard block design. Motion correction, slice-timing and detection of the activation areas for both right and left hands were performed within SPM5.
- dMRI: single shot EPI, 60 contiguous axial slices of size 128×128 (2×2 mm^2) and 2mm thickness. DSGs were applied in 15 non-collinear directions with $b =$

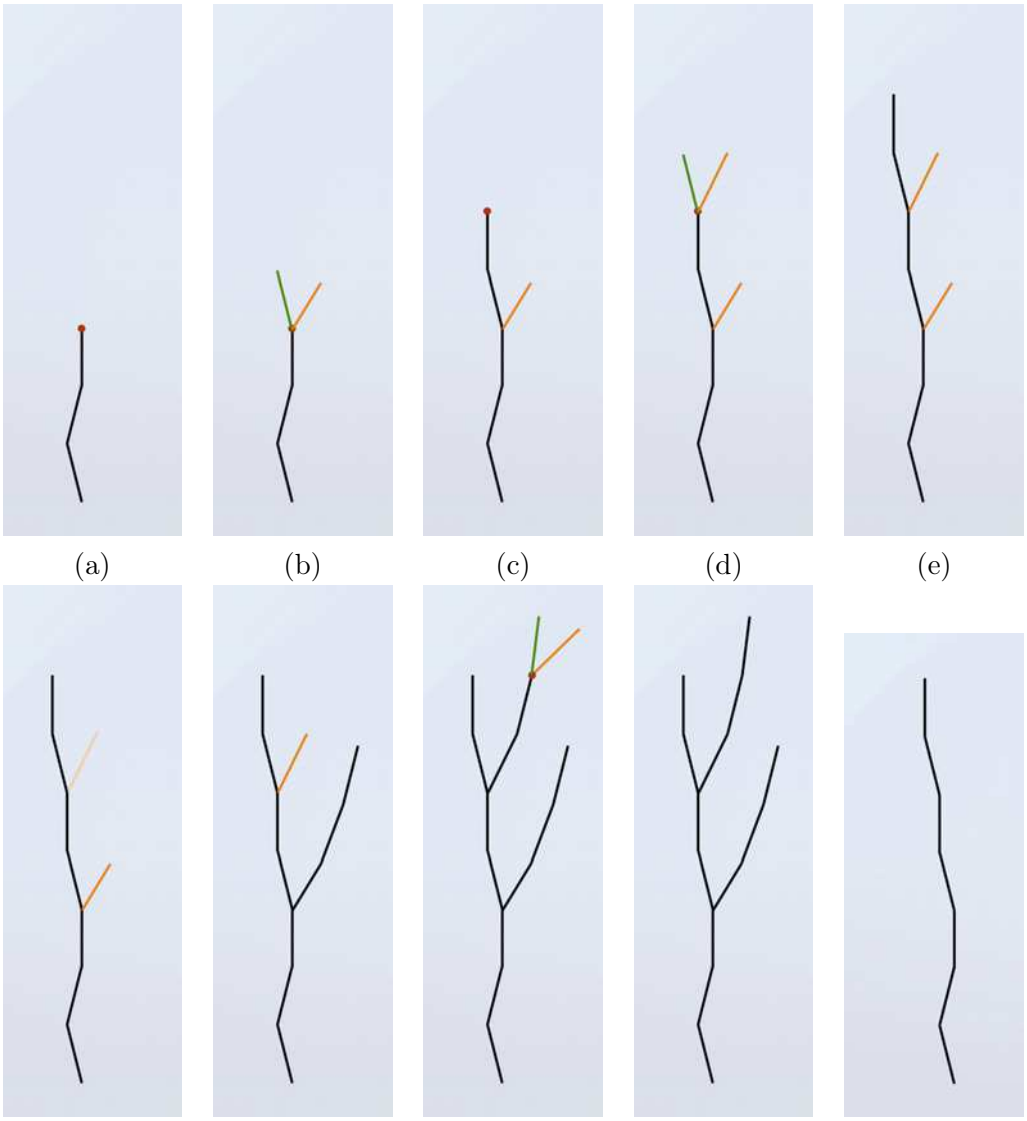


Figure 4.2: **Illustrative example of the multi-modal FACT reconstruction.** Reconstructed tract with bifurcations from our multi-modal FACT algorithm (a-i) and from the original FACT algorithm (j).

800 s/mm^2. Each measured DW image was corrected for eddy current-induced geometric distortions [Netsch 2004] and denoised using the Rician non-local means algorithm [Wiest-Daesslé 2008]. Given this low number of directions, the (modified) spherical harmonics basis of order 4 (15 parameters) was used for dODF estimation, while 6 parameters (resp. 7) were to be estimated for DTI (resp. DDI).

For each subject, the images from aMRI and fMRI were rigidly registered to the B0 image of the dMRI sequence [Maes 1997]. A first ROI was manually delineated

by an expert neuroanatomist in an axial slice on the aMRI data through the superior part of the mesencephalum, both on left and right sides. Tractography was then performed between these two anatomical ROIs and the two (left and right) cortical functional ROIs to reconstruct the HMTs using the three above-mentioned diffusion models and the previously described tractography algorithm.

4.1.3 Results

4.1.3.1 Connections between the ROIs

Our objective here was to evaluate whether the three diffusion models, coupled with the tractography algorithm, were able to connect fully, partially, or not at all, the anatomical and functional ROIs. First of all, we split each functional ROI into a *medial* and a *lateral* area, the latter corresponding to the extremity of the hand representation on the motor homunculus, *i.e.* the thumb. Then for each model (DTI/QBI/DDI), each hemisphere, and each of the 23 subjects, we evaluated the strength (subjectively based on the number of tracts) of the connection between the anatomical ROI and the medial part of the functional ROI by a discrete score of 0 (almost no tracts), 1 or 2 (lots of tracts); in a word, we estimated $3 \times 2 \times 23 = 138$ scores. Similarly, we computed another set of 138 scores for the connection with the lateral part of the functional ROI. At last, the overall number of tracts composing the reconstructed HMTs, *i.e.* connecting the two (anatomical and functional) ROIs, was also computed.

The Pearson χ^2 test is particularly adequate to handle such qualitative, discrete scores. We performed pairwise Pearson χ^2 tests with a significance level of 0.05, corrected for multiple comparisons (Bonferroni) to compare DTI vs QBI, QBI vs DDI, and DDI vs DTI for the medial and lateral areas. To compare the overall number of tracts, we first showed that the data were not Gaussian-distributed using the Jarque-Bera test, and then we performed pairwise sign rank tests (which allows the test of differences in medians) with a significance level of 0.05, corrected for multiple comparisons (Bonferroni).

The p-values are reported in table 4.1, left, and mainly show that i) the Q-ball model was able to reconstruct more medial tracts than the DT model, but as many lateral tracts, and that ii) the DDI model did not reconstruct more medial tracts than the Q-ball model, but did reconstruct more lateral tracts, at the 0.05 significance level. These two results are confirmed by the sign rank test on the overall number of tracts.

4.1.3.2 Asymmetry of the hand motor tracts

Our objective was to evaluate whether the three diffusion models, coupled with the tractography algorithm, were able to show significant differences (in terms of number of tracts) between the left and right reconstructed HMTs, in right-handed (14) and left-handed (9) subjects. We pooled males and females for increased statistical

	Medial	Lateral	# Tracts
DTI vs QBI	< 0.001	0.340	< 0.001
QBI vs DDI	0.037	< 0.001	< 0.001
DDI vs DTI	< 0.001	< 0.001	0.001

	RH	LH
DTI	0.092	1.000
QBI	0.581	0.180
DDI	0.019	0.508

Table 4.1: ***p*-values of the statistical tests.** Left table: "Is there a significant difference (level=0.05) between the 3 models in recovering tracts reaching the medial and lateral areas of the functional ROI, in each hemisphere, for the 23 subjects? And in the overall number of tracts?". Right table: "Are the 3 models able to show significant differences (level=0.05), in terms of number of tracts, between the left and right HMTs, for the 14 right-handed (RH) and the 9 left-handed (LH) subjects?".

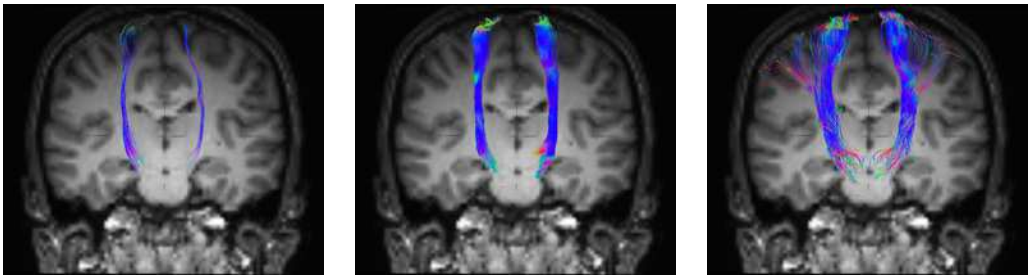


Figure 4.3: **Tractography of the HMTs in the left and right hemispheres.** From left to right: DTI, QBI and DDI. Note that we use the neurological convention, *i.e.* the left (resp. right) hemisphere is displayed on the left (resp. right). The DDI model is the only one able to consistently reach the lateral area within the functional ROI.

power. After showing that the data were not Gaussian-distributed using the Jarque-Bera test, we performed pairwise sign rank tests with a significance level of 0.05, corrected for multiple comparisons (Bonferroni).

The *p*-values are reported in table 4.1, right, and mainly show that i) none of the models is able to show differences in left-handed subjects, but that ii) the DDI model is the only model showing that the bundle of HMTs is *larger* (in terms of number of tracts) in the left hemisphere than in the right for right-handed subjects.

4.1.4 Conclusive remarks

In this section, we showed that i) the DDI model outperforms both the DT and the Q-ball models to reconstruct the HMTs (see fig. 4.3), and, maybe more importantly, that ii) it is the only model able to show that the number of HMTs in the left hemisphere is larger than in the contralateral hemisphere for right-handed subjects. The latter finding meets the intuition that the hand motor tracts in the hemisphere contralateral to the dominant hand *should* be somewhat *more developed* than those in the other hemisphere. Interestingly, the DDI model, as the other two models, failed to find such a difference for left-handed subjects, which may

suggest that left-handedness is not simply a *mirrored* right-handedness. This is already known from *e.g.* the notoriously differing patterns of hemispheric dominance for language between left and right-handed subjects, but to our knowledge, this is the first time such results are reported on white matter tracts, at least on clinical data. These results must now be further investigated in light of the huge literature on brain asymmetry and cerebral dominance [Toga 2003]. In particular, recruiting more males/females and right/left-handed subjects for increased statistical power and population-specific analysis would be necessary to confirm these first results, of potentially important anatomical significance.

As expected, the DT model performs very poorly when reconstructing the HMTs. Importantly, we stress here that we do *not* state that the DDI model outperforms the Q-ball model in general, but only in this particular experimental setting. Low angular resolution of the order-4 dODF model can explain why it is outperformed here. It would be of high interest to try to replicate our experiments on HARDI data using higher-order dODF models and improved (*e.g.* probabilistic) tractography methods, to further support our first findings.

We adapted to the FACT algorithm to account for the multiple fascicle orientations provided by MCMs and shows that it leads to interesting findings. However, it still suffers from two main limitations: (i) it strictly follows the one or several orientations provided by the MCM so that the reconstructed tract may be subject to accumulated errors due to the uncertainty on these estimated orientations and (ii) exploring all the orientations provided by the MCM at each voxel is prohibitively time- and space-consuming. We therefore moved towards probabilistic algorithms and, in particular, those based on the use of particle filters that try to approximate the distribution of all possible tracts from a single voxel which, at least theoretically, precisely answer the question raised by the tractography problem.

4.2 Multi-modal particle filtering

Particle filters are a common way to approximate a **filtering distribution**. One can define WM tractography as a filtering problem, where the filtering distribution is defined for some fixed tract length, as the distribution of all possible tracts of this length emanating from a single seed voxel. The use of particle filters to approximate such a filtering distribution has been pioneered by [Zhang 2009], where the diffusion information is estimated through DTI. The particle filter is improved in [Pontabry 2013] by using the Q-ball model to locally account for multiple fascicle orientations. However, particle filters often fail at consistently capturing the multi-modality of the filtering distribution [Vermaak 2003]. Multiple fascicles are thus unlikely to be tracked over extended volumes. We propose an adaptive multi-modal particle filter for WM tractography that improves the multi-modality capture. We apply it to both DTI and QBI (Section 4.2.1). We then design an experimental framework for validation of the proposed methods (Section 4.2.2) and show results on both synthetic data with an extensive comparison to their traditional particle

filter counterparts and real clinical data (Section 4.2.3). Another contribution is that the number of modes of the filtering distribution is dynamically estimated. We show that our proposed algorithm greatly improves WM tractography and, when the associated diffusion model accurately captures multiple fascicle orientations, it is able to distinguish crossings from bifurcations.

4.2.1 Proposed algorithm

4.2.1.1 Tractography as a mixture filtering problem

At step k, a tract is a sequence $X^k = [(\mathbf{x}_0, \mathbf{v}_{-1}), \dots, (\mathbf{x}_k, \mathbf{v}_{k-1})] \in \left(\Omega, \mathbb{S}^2\right)^{k+1}$ of successive pairs of positions and arrival directions, related by $\mathbf{x}_{k+1} = \mathbf{x}_k + \rho \mathbf{v}_k$, where $\rho > 0$ is the step size, which is assumed to be constant. The state space at step k is the set of all possible fiber pathways X^k originated from a specified position $\mathbf{x}_0$ and a specified arrival direction $\mathbf{v}_{-1}$.

The filtering distribution at step k is the distribution of X^k, given the diffusion data $\mathcal{Y}_k = \{y_\mathbf{x} : \|\mathbf{x} - \mathbf{x}_0\| \leq k\rho\}$ where $y_\mathbf{x} = \{S_{\mathbf{x},i}, \Psi_{\mathbf{x},i}\}_{i=1}^n$. S_i designates a measured DW image/signal and Ψ_i the corresponding predicted DW image/signal from some given diffusion model. At step $k+1$, this distribution is sequentially determined by successively computing the prediction distribution $p(X^{k+1}|\mathcal{Y}_k)$ out of the previous filtering distribution (at step k) and then using Bayes' rule to obtain $p(X^{k+1}|\mathcal{Y}_k)$. These **prediction** and **update** stages require to specify respectively the evolution model $p(X^{k+1}|X^k)$ and the likelihood $p(\mathcal{Y}_{k+1}|X^{k+1})$.

We assume that the tracts are first-order Markov chains. Consequently, the evolution model simplifies to $p(\mathbf{v}_k|\mathbf{v}_{k-1})$. Following [Zhang 2009, Pontabry 2013], we use the evolution model defined as:

$$p(\mathbf{v}_k|\mathbf{v}_{k-1}) = \mathrm{vMF}(\mathbf{v}_k; \mathbf{v}_{k-1}, \kappa) = \frac{\kappa}{4\pi \sinh \kappa} \exp\{\kappa < \mathbf{v}_{k-1}, \mathbf{v}_k >\}, \tag{4.2}$$

where $\mathrm{vMF}(\cdot; \mathbf{v}_{k-1}, \kappa)$ is the vMF distribution [Jupp 1989] on the 2-dimensional sphere with mean direction $\mathbf{v}_{k-1} \in \mathbb{S}^2$ and concentration parameter $\kappa \geq 0$. The concentration parameter κ of the evolution model controls the smoothness of the reconstructed tracts.

Denoting $\mathcal{B}_k = \{\mathbf{x} \in \Omega : \|\mathbf{x} - \mathbf{x}_0\| \leq k\rho\}$ and assuming that the diffusion data are spatially independent given a tract, the likelihood $p(\mathcal{Y}_{k+1}|X^{k+1})$ can be decomposed as follows:

$$p(\mathcal{Y}_{k+1}|X^{k+1}) = \prod_{\mathbf{x} \in \mathcal{B}_{k+1} \setminus \mathbf{x}_{0:k+1}} p_0(y_\mathbf{x}) \prod_{j=0}^{k} p_1(y_{\mathbf{x}_{j+1}}|\mathbf{v}_j) \propto \prod_{j=0}^{k} \frac{p_1(y_{\mathbf{x}_{j+1}}|\mathbf{v}_j)}{p_0(y_{\mathbf{x}_{j+1}})}, \tag{4.3}$$

where $\mathbf{x}_{0:k+1}$ is the polygonal chain defined by X^{k+1} and $p_1(\cdot|\mathbf{v}_j)$ (distribution of the measurement noise if $\mathbf{v}_j$ is the true diffusion orientation) and p_0 (distribution of the measurement noise) depend on the diffusion model and will be defined in section 4.2.1.4.

In order to better capture the possible multi-modality of the filtering distribution, we follow the idea of [Vermaak 2003] and formulate it as a mixture of M_k

components:

$$p(X^k|\mathcal{Y}_k) = \sum_{m=1}^{M_k} \pi_{m,k} p_m(X^k|\mathcal{Y}_k)\,, \tag{4.4}$$

where $\sum_{m=1}^{M_k} \pi_{m,k} = 1$. Such a formulation allows us to perform the filtering recursion for each component p_m individually, provided that each mixture weight is updated as the normalized weighted likelihood for the associated component.

4.2.1.2 Mixture particle filter

In general, there is no closed-form expressions for the filtering recursion equations. A popular strategy is to resort to particle filters. They approximate the filtering distribution by a set of sample (the particles) that are properly weighted to represent the filtering distribution at each step.

Traditional particle filters approximate distributions of the type of eq. (4.4) with $M_k = 1$ at each step k. The only difference that occurs when accounting for $M_k > 1$ is that particles are grouped into clusters and resampling is performed within the clusters, which gives better chances to secondary leads. Weighting is thus provided at two levels, the particles and the clusters, with adequate updating rules for compliance with sequential importance sampling. Consequently, one can see each cluster as a super-particle. The set of particles within a given cluster m aims at approximating the component p_m of the filtering distribution given in eq. (4.4).

Using the notations in [Vermaak 2003], let $\mathcal{P}_k = \{M_k, \Pi_k, \mathcal{X}_k, \mathcal{W}_k, \mathcal{C}_k\}$ be the particle representation of the filtering distribution where M_k is the number of components, $\Pi_k = \{\pi_{m,k}\}_{m=1}^{M_k}$ the set of mixture weights, $\mathcal{X}_k = \{\mathbf{x}_k^{(\ell)}\}_{\ell=1}^{N}$ the set of N particles, $\mathcal{W}_k = \{w_k^{(\ell)}\}_{\ell=1}^{N}$ the set of particle weights and $\mathcal{C}_k = \{c_k^{(\ell)}\}_{\ell=1}^{N}$ the set of component indicators (i.e., $c_k^{(\ell)} = m$ if particle ℓ belongs to component m). Given $\mathcal{P}_k$, the particle approximation with mixture filtering distribution proceeds to step $k+1$ in five stages:

Proposition of new samples: New samples are generated according to a proposal density $q(\cdot|\mathbf{v}_{k-1}^{(\ell)}, \mathcal{Y}_k)$ which depends on the previous direction and the diffusion information at step k:

$$\mathbf{v}_k^{(\ell)} \sim q(\mathbf{v}_k|\mathbf{v}_{k-1}^{(\ell)}, \mathcal{Y}_k) \text{ and } \mathbf{x}_{k+1}^{(\ell)} = \mathbf{x}_k^{(\ell)} + \rho \mathbf{v}_k^{(\ell)}\,. \tag{4.5}$$

Update of particle weights: The weights of the new particles are updated in order to be representative of the filtering distribution according to [Doucet 2000]:

$$\tilde{w}_{k+1}^{(\ell)} = \frac{w_k^{(\ell)} p_1(y_{\mathbf{x}_{k+1}^{(\ell)}}|\mathbf{v}_k^{(\ell)}) p(\mathbf{v}_k^{(\ell)}|\mathbf{v}_{k-1}^{(\ell)})}{p_0(y_{\mathbf{x}_{k+1}^{(\ell)}}) q(\mathbf{v}_k^{(\ell)}|\mathbf{v}_{k-1}^{(\ell)}, \mathcal{Y}_k^{(\ell)})}\,. \tag{4.6}$$

The normalization of these weights is performed within each component:

$$w_{k+1}^{(\ell)} = \frac{\tilde{w}_{k+1}^{(\ell)}}{\sum_{j \in \mathcal{I}_{m,k}} \tilde{w}_{k+1}^{(j)}}\,, \tag{4.7}$$

where $\mathcal{I}_{m,k} = \{\ell \in [\![1, N]\!] : c_k^{(\ell)} = m\}$ is the set of indices of the particles that belong to the m-th mixture component at step k.

Update of mixture weights: The mixture weights need to be updated properly to ensure that the filter still acts on each component individually:

$$\pi_{m,k+1} = \frac{\pi_{m,k}\tilde{w}_{m,k+1}}{\sum_{i=1}^{M} \pi_{i,k}\tilde{w}_{i,k+1}} \text{ with } \tilde{w}_{m,k+1} = \sum_{\ell \in \mathcal{I}_{m,k}} \tilde{w}_{k+1}^{(\ell)} . \tag{4.8}$$

Resampling within each component: To avoid the degeneracy of the particle weights, occasional resampling is necessary [Doucet 2000]. The resampling stage can be performed within each subset of particles associated to a mixture component independently, according to the component particle weights [Vermaak 2003]: we compute the effective number of particles in a mixture component as:

$$\mathrm{ESS}_m = \left(\sum_{\ell \in \mathcal{I}_{m,k}} \left(w_{k+1}^{(\ell)} \right)^2 \right)^{-1} , \tag{4.9}$$

and perform within-cluster resampling from a categorical distribution with $|\mathcal{I}_{m,k}|$ categories and event probabilities given by eq. (4.7), if ESS_m is below a threshold $\alpha|\mathcal{I}_{m,k}|$, where $|\cdot|$ denotes the set size operator.

Reclustering of the particles within new components: The number of components M_k in the mixture is not known. At the end of each step, it is dynamically estimated by merging and/or splitting some of the components: M_k, $\mathcal{C}_k$ and $\mathcal{I}_{m,k}$ are updated to M_{k+1}, $\mathcal{C}_{k+1}$ and $\mathcal{I}_{m,k+1}$ accordingly.

In section 4.2.1.3, we describe how the reclustering of the mixture filtering distribution is performed. In section 4.2.1.4, we define the proposal density and the likelihood for the DT and Q-ball models following respectively [Zhang 2009] and [Pontabry 2013].

4.2.1.3 Reclustering of the particle set

After the resampling stage, we characterize each cluster by

1. the weighted mean of the cluster particles, referred to as the *global cluster mean particle*;

2. a vMF distribution, referred to as the *local cluster directional distribution*, with mean direction $\boldsymbol{\mu}_{m,k}$ and concentration $\kappa_{m,k}$. The parameters of the distribution are estimated using the following equations:

$$\mathbf{r}_{m,k} := \frac{\sum_{\ell \in \mathcal{I}_{m,k}} \mathbf{v}_k^{(\ell)}}{|\mathcal{I}_{m,k}|} , \qquad \boldsymbol{\mu}_{m,k} = \frac{\mathbf{r}_{m,k}}{\|\mathbf{r}_{m,k}\|} , \qquad \kappa_{m,k} = \frac{\|\mathbf{r}_{m,k}\|(3 - \|\mathbf{r}_{m,k}\|^2)}{1 - \|\mathbf{r}_{m,k}\|^2} . \tag{4.10}$$

These estimators have been proposed in [Banerjee 2006] and have been introduced for dMRI in [Bhalerao 2007] for their unbiasedness and robustness.

We first test the clusters pairwise for merging. The decision to merge two clusters shall be based on the difference in shape between the global cluster mean particles.

Consequently, we merge two clusters if the Hausdorff distance [Rockafellar 1998] between the global cluster mean particles drops below a threshold h_t.

We then test each cluster for splitting. The decision to split a cluster shall be based on some local measure of directional dispersion. Consequently, we split a cluster if the concentration parameter $\kappa_{m,k}$ of the local cluster directional distribution drops below a threshold κ_t.

Finally, the number of mixture components is updated to M_{k+1}, the component indicators to $\mathcal{C}_{k+1}$ and the set of indices to $\mathcal{I}_{m,k+1}$. In order to maintain a properly weighted sample from the filtering distribution and thus to preserve the convergence properties of the particle filter, we perform the following update of mixture and particle weights [Vermaak 2003]:

$$\pi^{\star}_{m,k+1} = \sum_{\ell \in \mathcal{I}_{m,k+1}} \pi_{c^{(\ell)}_k,k+1} w^{(\ell)}_{k+1}\,, \qquad w^{(\ell)\star}_{k+1} = \frac{\pi_{c^{(\ell)}_k,k+1} w^{(\ell)}_{k+1}}{\pi^{\star}_{c^{(\ell)}_{k+1},k+1}}\,. \tag{4.11}$$

4.2.1.4 Diffusion models: associated proposal densities and likelihoods

The DT model. It provides a 2nd order DT represented by its eigensystem $\{d_1, d_2, d_3, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, of which we extract the FA (eq. (1.20)), the linear coefficient c_l (eq. (1.21)), the MD $\overline{D}$ (eq. (1.19)), the perpendicular diffusivity $D_\perp$ (eq. (1.18)), the principal eigenvector $\mathbf{e}_1$ and the minor eigenvector $\mathbf{e}_3$.

Noise in dMRI is known to be Rician [Gudbjartsson 1995]. The logarithm of the measured DW signal is thus approximately normally distributed around the predicted DW signal in absence of noise with standard deviation equal to the inverse SNR [Salvador 2005]. Following [Zhang 2009], the likelihood $p_1(y_{\mathbf{x}_{k+1}}|\mathbf{v}_k)$ at position $\mathbf{x}_{k+1}$ given the local diffusion direction $\mathbf{v}_k$ thus reads:

$$p_1(y_{\mathbf{x}_{k+1}}|\mathbf{v}_k) = \begin{cases} \displaystyle\prod_{i=1}^{n} \frac{\Psi^{\star}_{\mathbf{x}_{k+1},i}}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(\Psi^{\star}_{\mathbf{x}_{k+1},i})^2(\log S_{\mathbf{x}_{k+1},i} - \log \Psi^{\star}_{\mathbf{x}_{k+1},i})^2}{2\sigma_i^2}\right\}, & c_l > \tau\,, \\ \displaystyle\frac{1}{\sigma\sqrt{(2\pi)^3}} \exp\left\{-\frac{(\arccos < \mathbf{v}_k, \mathbf{e}_3 > -\pi/2)^2}{2\sigma^2}\right\}, & c_l \le \tau\,, \end{cases} \tag{4.12}$$

where the diffusion tensor is estimated at position $\mathbf{x}_{k+1}$, σ_i and $\mathbf{g}_i$ are the standard deviation and the DSG direction of the i-th gradient image respectively, estimated by **LS!** (**LS!**) estimation and pseudo-residuals [Gasser 1986], σ is a user-defined standard deviation and $\Psi^{\star}_{\mathbf{x}_{k+1},i} = \Psi_{\mathbf{x}_{k+1},0} \exp\{-b(D_\perp + 3 < \mathbf{v}_k, \mathbf{g}_i >^2 (\overline{D} - D_\perp))\}$ is the predicted DW signal from the DT model cylindrically constrained along the sampled direction $\mathbf{v}_k$.

Similarly, the likelihood $p_0(y_{\mathbf{x}_{k+1}})$ at position $\mathbf{x}_{k+1}$ can thus be expressed as follows:

$$p_0(y_{\mathbf{x}_{k+1}}) = \begin{cases} \prod_{i=1}^{n} \frac{\Psi_{\mathbf{x}_{k+1},i}}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(\Psi_{\mathbf{x}_{k+1},i})^2(\log S_{\mathbf{x}_{k+1},i} - \log \Psi_{\mathbf{x}_{k+1},i})^2}{2\sigma_i^2}\right\}, & c_l > \tau\,, \\ \frac{1}{\sigma\sqrt{(2\pi)^3}}\,, & c_l \leq \tau\,, \end{cases} \tag{4.13}$$

where $\Psi^{\star}_{\mathbf{x}_{k+1},i} = \Psi_{\mathbf{x}_{k+1},0} \exp\{-b(D_\perp + 3 < \mathbf{e}_1, \mathbf{g}_i >^2 (\overline{D} - D_\perp))\}$ is the predicted DW signal from the DT model.
We define the proposal density as:

$$q(\mathbf{v}_k|\mathbf{v}_{k-1}, y_{\mathbf{x}_k}) = \begin{cases} \text{vMF}(\mathbf{v}_k; \mathbf{e}_{1,k}, \nu_k)\,, & c_l > \tau\,, \\ p(\mathbf{v}_k|\mathbf{v}_{k-1})\,, & c_l \leq \tau\,, \end{cases} \tag{4.14}$$

where the DT is estimated at $\mathbf{x}_k$ and ν_k is a function of FA [Zhang 2009].

The Q-Ball model. It provides an estimate of the fODF of which we extract the set Λ of maxima $\boldsymbol{\mu}$, the value of the fODF at its maxima $\mathcal{M}(\boldsymbol{\mu})$ and the mean curvature of the fODF at its maxima $\mathcal{H}(\boldsymbol{\mu})$. Borrowing ideas from [Pontabry 2013], we define the likelihood $p_1(y_{\mathbf{x}_{k+1}}|\mathbf{v}_k)$ at position $\mathbf{x}_{k+1}$ given the local diffusion direction $\mathbf{v}_k$ as follows:

$$p_1(y_{\mathbf{x}_{k+1}}|\mathbf{v}_k) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(S_{\mathbf{x}_{k+1},i} - \Psi^{\star}_{\mathbf{x}_{k+1},i})^2}{2\sigma_i^2}\right\}, \tag{4.15}$$

where the fODF is estimated at position $\mathbf{x}_{k+1}$ and $\Psi^{\star}_{\mathbf{x}_{k+1},i}$ is the diffusion signal simulated according to [Aganj 2010] from the fODF that has been rotated to align the sampling direction to the sampled one.
Similarly, the likelihood $p_0(y_{\mathbf{x}_{k+1}})$ at position $\mathbf{x}_{k+1}$ can thus be expressed as follows:

$$p_0(y_{\mathbf{x}_{k+1}}) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(S_{\mathbf{x}_{k+1},i} - \Psi_{\mathbf{x}_{k+1},i})^2}{2\sigma_i^2}\right\}, \tag{4.16}$$

where $\Psi_{\mathbf{x}_{k+1},i}$ is the diffusion signal simulated from the fODF according to [Aganj 2010].
We define the proposal density as:

$$q(\mathbf{v}_k|\mathbf{v}_{k-1}, y_{\mathbf{x}_k}) = \begin{cases} \sum_{\boldsymbol{\mu}\in\Lambda} \omega_{\boldsymbol{\mu}} \text{vMF}(\mathbf{v}_k; \boldsymbol{\mu}, \kappa_{\boldsymbol{\mu}})\,, & \Lambda \neq \varnothing\,, \\ p(\mathbf{v}_k|\mathbf{v}_{k-1})\,, & \Lambda = \varnothing\,, \end{cases} \tag{4.17}$$

where the fODF is estimated at position $\mathbf{x}_k$, $\omega_{\boldsymbol{\mu}} \propto \mathcal{M}(\boldsymbol{\mu})$ (normalized) and $\kappa_{\boldsymbol{\mu}} \propto \mathcal{H}(\boldsymbol{\mu})$.

4.2.1.5 Summarized algorithm

The proposed algorithm has the following user-defined parameters:

- FA_t is the threshold on FA: tracts reaching a voxel with too low FA are stopped,
- ρ is the step length: at each step, we update the tracts one step of such length ahead,
- N is the size of the particle set,
- $L_{\max}$ is the common maximum length of the tracts: tracts that reach this maximum length are stopped,
- $N_{\min}$ is the minimum allowed cluster size: particle clusters that reach this size are not allowed to be splitted anymore,
- α is the relative resampling threshold: if the effective sample size within cluster drops below this relative threshold, particles are resampled according to their weights in the cluster,
- κ is the concentration of the prior vMF distribution: the higher κ, the smoother the tracts,
- h_t is the threshold on the Hausdorff distance between the global cluster mean particles: clusters are merged if this distance drops below the treshold,
- κ_t is the threshold on the concentration of the local cluster directional distribution: a cluster is splitted if this concentration drops below the threshold.

A summary of the proposed tractography algorithm is given in algorithm 1. We used the following stopping criteria for generated tracts: (i) the maximum number $L_{\max}$ of steps is reached, (ii) the anisotropy becomes too low or (iii) it reaches a position outside the brain. Other model-specific stopping criteria may also be included in the algorithm. For instance, if the diffusion model is an MCM, the CSF fraction of occupancy w_0^{CSF} is often available. It is then reasonable to stop the generation of a tract that reaches a voxel in which $w_0^{\text{CSF}} \geq w_t^{\text{CSF}}$, where a typical value for the threshold is $w_t^{\text{CSF}} = 0.8$.

An important algorithmic issue arises with the use of the vMF distribution. Its PDF and the sampling procedure described in [Ulrich 1984] indeed involve exponentials of the concentration parameter κ that make the method numerically unstable for large κ. We circumvent this issue using some tricks proposed in [Jakob 2012], who provides numerically stable variants.

4.2.2 Experimental setup & evaluation metrics

4.2.2.1 Phantom data

Two synthetic DW phantoms were created for validation and are illustrated in fig. 4.4: a case of two crossing fascicles at a 90° angle and a case of one bifurcating fascicle at a 60° angle. For both phantoms, one measured unweighted image S_0 and 81 measured DW images with a single b-value of 3000 s.mm^{-2} were simulated using an equally weighted multi-tensor model at each voxel. Rician noise was then added on the noise-free images with a relative standard deviation of 5%, to generate 50 samples of each phantom.

Algorithm 1: Multi-modal particle filter for WM tractography

Require: A set of n measured DW images $\{S_i\}_{i=1}^{n}$ and corresponding predicted DW images $\{\Psi_i\}_{i=1}^{n}$ from a diffusion model, one true unweighted image Ψ_0, one diffusion model image $\mathcal{M}$ and a seed voxel $\mathbf{x}_0$.

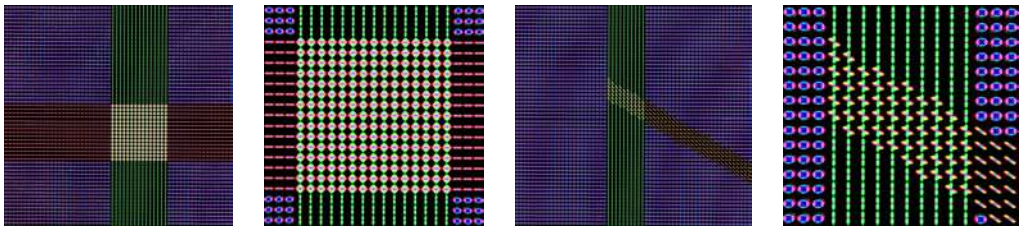
Output : Clustered tract samples originated from $\mathbf{x}_0$.

initialize stepCount = 0, stoppingCriteria = 0, $M_0 \leftarrow$ n° of principal diffusion directions (PDD) in $\mathcal{M}_{\mathbf{x}_0}$

for $\ell = 1$ **to** N **do** $\mathbf{v}_{-1}^{(\ell)} \leftarrow$ PDD given by $\mathcal{M}_{\mathbf{x}_0}$; $w_0^{(\ell)} \leftarrow 1/N$; $\mathbf{x}_0^{(\ell)} \leftarrow \mathbf{x}_0$

while *not* stoppingCriteria **do**

- update $\mathcal{X}_k$ with eq. (4.5)
- update $\mathcal{W}_k$ with eqs. (4.2), (4.6) and (4.7) and section 4.2.1.4
- update Π_k with eqs. (4.6) and (4.8)
- **for** $m = 1$ **to** M_k **do**
  - compute effective sample size ESS_m with eq. (4.9)
  - **if** $ESS_m \leq \alpha|\mathcal{I}_{m,k}|$ **then** within-cluster resampling from eq. (4.7)
  - compute global cluster mean particle
  - compute local cluster directional distribution with eq. (4.10)
- compute Hausdorff distance d_H between global cluster mean particles
- **if** $d_H < h_t$ **then** merge compared clusters and update M_k
- **for** $m = 1$ **to** M_k **do** **if** $\kappa_{m,k} < \kappa_t$ **then** split cluster
- update M_k, $\mathcal{C}_k$ and $\mathcal{I}_{m,k}$
- update Π_k and $\mathcal{W}_k$ with eq. (4.11)
- update stoppingCriteria and increment stepCount



(a) (b) (c) (d)

Figure 4.4: **Noise-free DW Phantoms.** fODF visualization of the crossing fascicles (a) with close-up view (b), and the bifurcating fascicle (c) with close-up view (d).

4.2.2.2 Evaluation measures for phantom data

Four methods were utilized: DTI-based “mono-modal" (i.e., without particle clustering mechanism) tractography (DTI mono), DTI-based multi-modal tractography (DTI multi), QBI-based mono-modal tractography (QBI mono) and QBI-based multi-modal tractography (QBI multi). For each phantom, a single seed voxel was placed in the upper branch. The initial direction of propagation was set towards the bottom. The following common parameters were used in all algorithms: resampling threshold $\alpha = 0.4$, number of particles $N = 1000$, step length $\rho = 1$ mm, prior concentration $\kappa = 30$, and merge and split thresholds $h_t = 0.5$ mm and $\kappa_t = 40$. Tensors were considered as oblate for $\tau = 0.25$. The output tracts are the final component mean particles (a single one in mono-modal versions).

The tractography results were evaluated visually and with 3 different quantitative measures: (i) the proportion of tracts branching from the main direction, which is an indicator of branching capacity (ii) the root-mean-square error between the end point (after $L_{\max}$ iterations of the particle filter) of each tract following the main path and the expected arrival position (known in the phantoms), which gives an idea of how spread the tracts are around the true one and (iii) the local curvature along each branching tract for the bifurcation phantom, which translates how each branch was created from the main direction (either by an uncertain turn or by a sharper local turn).

The expected arrival position in the bifurcation phantom is not obvious. In this phantom, tracts going straight follow the inaccurate diffusion orientations given by the diffusion model in the heterogeneous region and are thus expected to deviate exclusively towards the right border of the vertical band. Therefore, we have chosen the end position for the bifurcation phantom at the center of the segment joining the center of the vertical band and its right border (red point in figs. 4.5e-h).

4.2.2.3 Clinical diffusion acquisition

The 4 algorithms were also applied on real clinical scans, acquired on a Siemens 3T scanner with a matrix size of 128×128, 60 slices (voxel size $2 \times 2 \times 2$ mm^3). The diffusion acquisition consisted of one measured unweighted image and 30 measured DW images with a b-value of 1000 s.mm^{-2}. Seed ROIs were placed by a radiologist at the basis of the left and right CSTs in the mesencephalon, with filtering ROIs in the posterior limb of the internal capsule to retain only the tracts belonging to the CST. The same parameters as for synthetic data were utilized for real data. A particle filter was initiated at each voxel of the seed regions and the displayed fascicles are the final component mean particles.

4.2.3 Results

4.2.3.1 Experiments on Synthetic Data

We present a representative example of the results achieved by each method in fig. 4.5. We clearly notice that the 2 mono-modal methods fail to capture the multi-modality of the bifurcation phantom and therefore follow only one of the two directions. On the contrary, the 2 multi-modal methods are visually well able to capture the two branching fascicles, thanks to the adaptive clustering based on the proposed directions. It may be noted that DTI multi tends to obtain more fanning tracts, because the observation model is wide for oblate tensors. These visual results are valid for both crossing and bifurcation phantoms. However, in the crossing one, tracts are only expected to go straight since crossing tracts are not part of the same fascicle. Therefore, QBI mono and multi are performing well while DTI multi tends to capture too many branches.

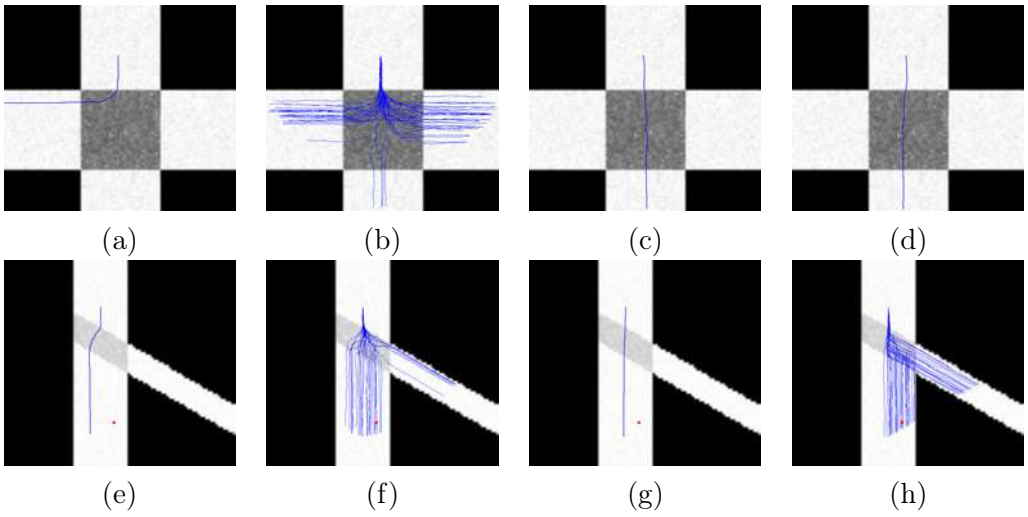


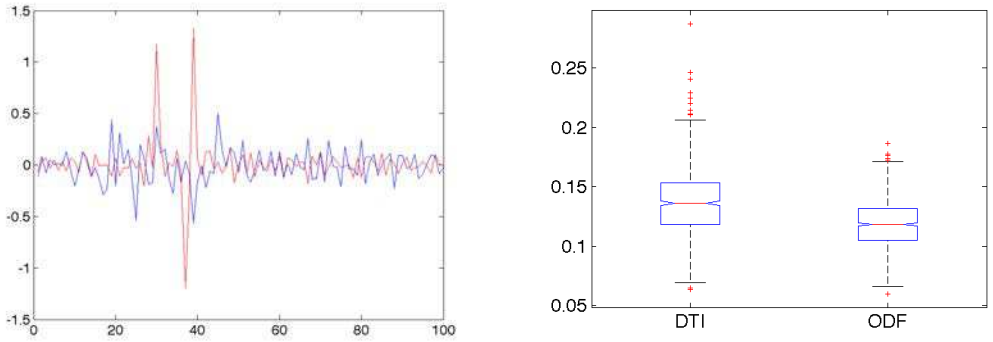
Figure 4.5: **Tractography results on phantom data.** First row: crossing phantom, Second row: bifurcation phantom. Methods used were DTI mono (a,e), DTI multi (b,f), QBI mono (c,g), QBI multi (d,h).

We report in table 4.2 the proportion of branching tracts for each phantom and each method. These quantitative results confirm the visual ones. When utilizing the mono-modal methods, only one of the two branches of each phantom is explored. On the contrary, multi-modal methods capture much better the 2 modes in the bifurcation phantom, with QBI multi being the closest to the half/half ground truth in each branch.

In addition, table 4.2 displays the RMSE towards the expected arrival point of the straight tracts only for those tracts which go in the straight branch of each phantom. For both phantoms, DTI mono and multi perform worse as DTI does not handle multiple directions. QBI multi outperforms the other methods, being able to better recover the final positions of the tracts.

	DTI mono	DTI multi	ODF mono	ODF multi
Crossing Phantom				
Fiber proportion straight (%)	0 ± 0	8.75 ± 3.06	100 ± 0	98.76 ± 6.14
Fiber proportion branch (%)	100 ± 0	91.25 ± 3.06	0 ± 0	1.24 ± 6.14
RMSE (mm)	N/A	9.71 ± 3.71	2.09 ± 1.35	2.05 ± 1.15
Bifurcation Phantom				
Fiber proportion straight (%)	93.88 ± 24.22	66.53 ± 4.06	100 ± 0	56.63 ± 4.59
Fiber proportion branch (%)	6.12 ± 24.22	33.47 ± 4.06	0 ± 0	43.37 ± 4.59
RMSE (mm)	20.16 ± 6.07	15.36 ± 1.28	12.94 ± 0.29	9.08 ± 0.41

Table 4.2: **Evaluation of tracts on phantom data.** Proportion of branching and straight tracts and RMSE of tracts going straight with respect to true expected position, for each phantom.



(a) (b)

Figure 4.6: **Local curvature of branching tracts.** Evaluation of local curvature for the branching tracts of the bifurcation phantom. (a): Local curvature on one single fiber (blue: DTI-based, multimodal, red: QBI-based, multimodal), (b): box-plot representation of inter-quantile range for DTI and QBI over all tracts and repetitions.

The last metric, only for the bifurcation phantom, is the local curvature of each mean tract that deviates from the main vertical path. We report one representative example of the obtained curves as well as a box-plot of inter-quantile ranges of the curvatures along each mean tract for all repetitions (Figure 4.6). The curves of local curvatures clearly show a more peaked behavior for QBI multi (red curve), indicating that it branches more sharply. This was expected as the Q-ball model captures the 2 diffusion orientations in the splitting region, whereas the DT model inaccurately estimates the diffusion orientations and thus the particle filter is mainly driven by the previous direction with a wide observation model. Also, a one-way ANOVA quantitatively shows a significant difference ($p \ll 10^{-3}$) between the standard deviations of the curvature of the 2 methods: away from the peaks, the curvature varies less with QBI multi than with DTI multi.

4.2.3.2 Experiments on real data

To illustrate the capacities of the proposed algorithms on real clinical datasets, we report in fig. 4.7 the left and right CSTs obtained on a normal control subject. The results obtained here are consistent with those obtained on synthetic data. While the mono-modal methods do not capture branchings to lateral parts of the CST, both DTI multi and QBI multi are able to capture branches to the hand area or even sometimes the face area. Interestingly, DTI multi seems a bit more able to capture branches than QBI multi especially on the right side of the brain (left in the images). Since the data was acquired with only 30 DSG directions and a single low b-value of 1000 s.mm^{-2}, it might indeed not be enough for the Q-ball model to identify accurately multiple orientations. However, QBI mono and multi seem overall able to capture more accurate directions of the main tract.

4.2.4 Discussion

We have presented a new adaptive multi-modal particle filter algorithm for WM probabilistic tractography. It relies on the adaptive clustering of the filtering distribution through a new scheme for splitting and merging clusters based on clusters shape and directional dispersion. This strategy is applied after each step of the filtering recursion. We have implemented this algorithm with two different diffusion models: the DT model and the Q-ball model.

We have demonstrated through experiments on synthetic and real data that our proposed algorithms outperform more traditional particle filtering approaches available in the literature, being more able to capture branching and crossing fascicles. In addition, the Q-ball multi-modal algorithm produces more accurate branchings and differentiates crossing fasiclces from bifurcating fascicles, thanks to the diffusion model that already locally accurately captures multiple orientations.

The proposed algorithm is very generic and can thus easily be extended to new diffusion models: only the proposal and the observation densities require to be modified to reflect the underlying diffusion model. As noted in section 4.2.3.2, it

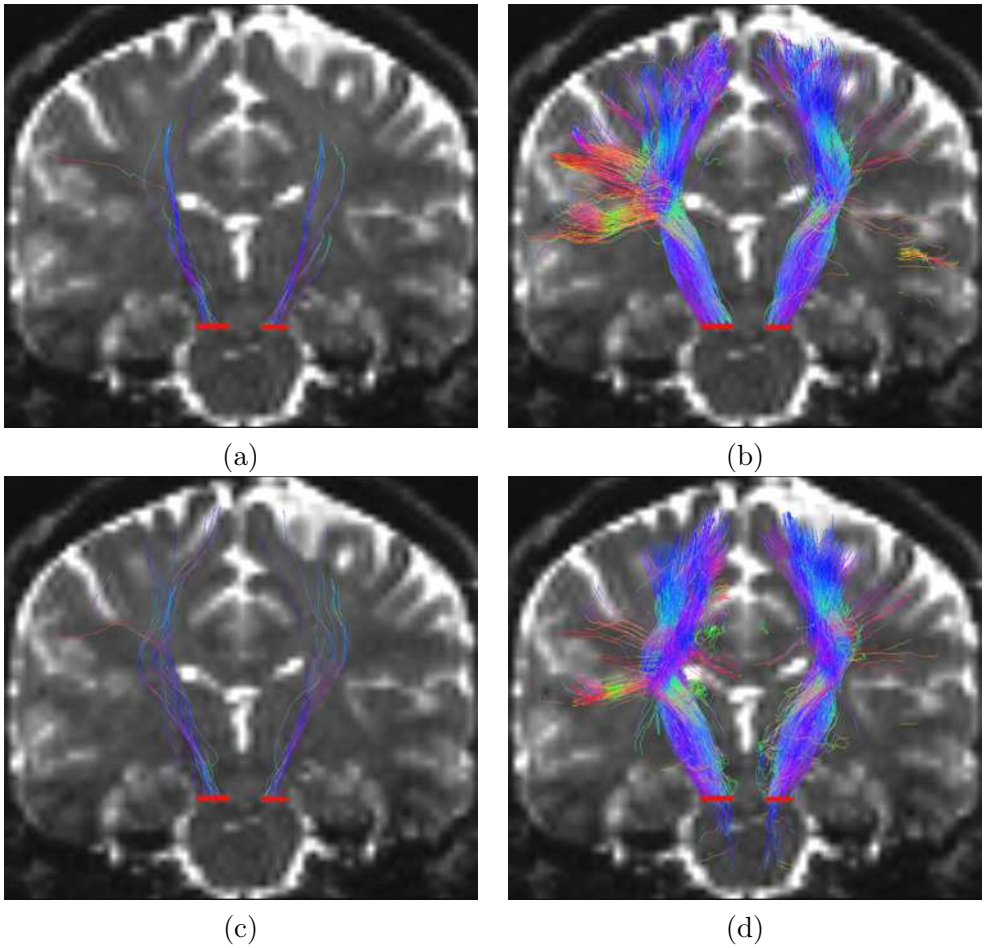


Figure 4.7: **Tractography of the left and right CSTs.** Coronal view of both CSTs using the four proposed methods: (a): DTI mono, (b): DTI multi, (c): QBI mono, (d): QBI multi. Tracts are overlaid on the T2-w volume from the diffusion acquisition, red bars indicate the seeding regions.

will be interesting to apply this algorithm using other diffusion models such as the geometry-based MCMs reviewed in chapter 2 or our proposed DDI MCM. These models are indeed more adapted to data with a small number of DSG directions and a single low b-value and particularly well suited to define proposal densities. In the next section, we briefly demonstrate how the we can easily adapt it to our proposed DDI MCM and how the use of such a model qualitatively improves the tractography results on the reconstruction of the CST of a healthy subject. Future work will include evaluation on more real data cases, also with HARDI, as well as an in-depth study of the user-defined parameters of the algorithm and their influence on the results.

4.2.5 DDI-based multi-modal particle filter

As mentioned in section 4.2.4, the multi-modal particle fiter can be easily adapted to any diffusion model, provided that one can define the observation and proposal densities.

Following the ideas of [Pontabry 2013], the proposal density is modeled by a mixture of vMF PDFs, where the mean direction of each mixture component is one of the diffusion orientations identified by the diffusion model (with occasional 180° rotation to ensure that the inner product with the arrival direction is positive). Weights and concentration parameters required to completely define the proposal density are however not a natural output of the Q-ball model, which forced [Pontabry 2013] to define them from the fODF value and mean curvature respectively.

MCMs are particularly well suited to define the proposal densities. The weights of each compartment can indeed directly be used as the weights for the vMF mixture proposal density. In addition, the DDI MCM introduced in section 2.3 outputs a concentration parameter for each compartment that can be directly (or, to be precise, up to a scaling factor) used as concentration parameter for the corresponding vMF distribution in the mixture defining the proposal density.

The setting of this scaling factor is however not obvious. We propose to compute the FA of each fascicle using eq. (2.44) and then transform it into a suitable concentration parameter for the proposal density according to the fitting rule defined in [Zhang 2009]. As a result, we define the proposal density as:

$$q(\mathbf{v}_k|\mathbf{v}_{k-1}, y_{\mathbf{x}_k}) = \begin{cases} \sum_{i=1}^{M} w_i \mathrm{vMF}(\mathbf{v}_k; \boldsymbol{\mu}_i, \kappa_i)\,, & M > 0\,, \\ p(\mathbf{v}_k|\mathbf{v}_{k-1})\,, & M = 0\,, \end{cases} \tag{4.18}$$

where the DDI MCM is estimated at position $\mathbf{x}_k$, w_i and $\boldsymbol{\mu}_i$ are respectively the weight and the diffusion direction (obtained from the appropriately rotated orientation) of the i-th compartment, κ_i is the concentration parameter obtained from the FA of the i-th compartment according to [Zhang 2009] and M is the number of fascicle compartments of the DDI MCM whose κ_i is greater than the concentration parameter κ of the prior density.

We define the likelihood $p_1(y_{\mathbf{x}_{k+1}}|\mathbf{v}_k)$ at position $\mathbf{x}_{k+1}$ given the local diffusion direction $\mathbf{v}_k$ as follows:

$$p_1(y_{\mathbf{x}_{k+1}}|\mathbf{v}_k) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(S_{\mathbf{x}_{k+1},i} - \Psi^{\star}_{\mathbf{x}_{k+1},i})^2}{2\sigma_i^2}\right\}, \tag{4.19}$$

where the DDI MCM is estimated at position $\mathbf{x}_{k+1}$ and $\Psi^{\star}_{\mathbf{x}_{k+1},i}$ is the predicted DW signal computed from eq. (2.65), after applying to all the directions of the model the rotation that aligns the sampling direction to the sampled one.

The likelihood $p_0(y_{\mathbf{x}_{k+1}})$ at position $\mathbf{x}_{k+1}$ is the same as for the QBI-based algorithm:

$$p_0(y_{\mathbf{x}_{k+1}}) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(S_{\mathbf{x}_{k+1},i} - \Psi_{\mathbf{x}_{k+1},i})^2}{2\sigma_i^2}\right\}, \tag{4.20}$$

where $\Psi_{\mathbf{x}_{k+1},i}$ is the predicted DW signal computed from eq. (2.65).

We used the adaptive multi-modal particle filter for tractoagraphy proposed in this section in conjunction to the constrained DDI MCM detailed in section 2.3.4.3 to reconstruct the CST in patients suffering from tumors of various grades. This was in the context of the three editions of the DTI Tractography challenge organized at the MICCAI conference these last three years and will be the object of the following chapter.

Chapter 5

MICCAI DTI Tractography Challenges

Contents

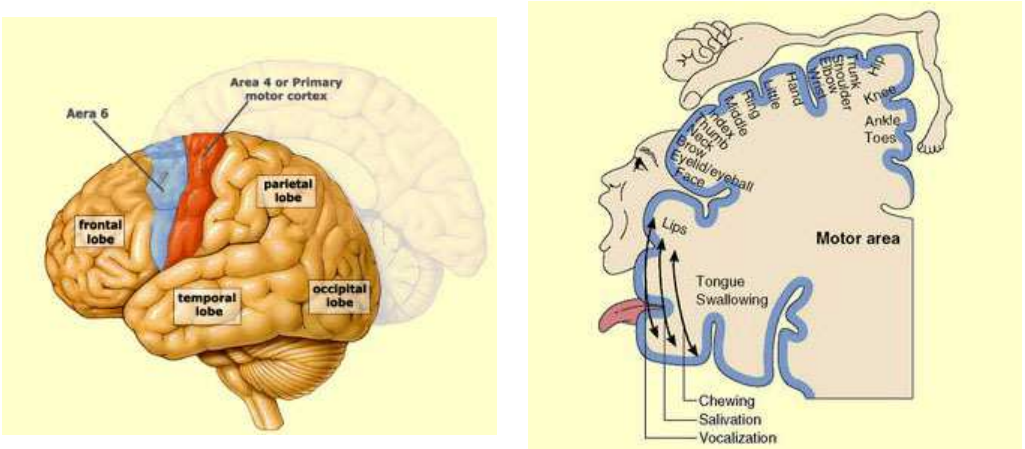
The advent of dMRI and the subsequent development of analysis tools provide unique new information about the brain architecture: so-called tractography enables to gain insights into the organization of white matter fascicles as well as their microstructure. A targeted application is neurosurgical planning, which could be improved with such valuable information. In particular, having an accurate reconstruction of the CST might be of great help for the surgeon for better preparing, for example, the removal of a tumor without damaging the remaining surrounding tissues.

Since 2011, a *DTI Tractography Challenge* is organized each year at the MICCAI conference, aiming at encouraging researchers to deal with the poor quality of the

measured DW images in the neurosurgical context. We participated at each of these three events which we believe to be a good evaluation of all the tools developed in this thesis. In this chapter, we first give a detailed description of the CST and its importance in neurosurgical planning (section 5.1). Then, we present the data that the organizers of the challenge provided each year (section 5.2). Methods and reconstructed CSTs for the three editions of the challenge are given in sections 5.3 to 5.5. Finally, we outline in section 5.6 a discussion about our participation and results.

5.1 The Cortico-Spinal Tract

The Cortico-Spinal Tract (CST) is a projection fascicle that connects the spinal cord to the primary motor cortex [Catani 2002, Jellison 2004, Catani 2008, Thiebaut de Schotten 2011]. The primary motor cortex is located at the posterior portion of the frontal lobe (see fig. 5.1a). It extends from the top of the brain to the lateral part and is divided into areas of non-equal sizes that controls the volontary movement of different parts of the body. The primary motor cortex can thus be seen as a *projection of the body within the brain*. Identification of the function of the different areas of the primary motor cortex in the human brain has been established in [Penfield 1937] by using electrical stimulations on patients that underwent surgery for epilepsy. This is known as Penfield's motor homonculus, which is illustrated in fig. 5.1b. Observe that areas are actually not proportional to the size of the body part they are controlling but to the complexity of the movements that the body part can achieve. This explains why the hand and face areas are so large compared to the other areas, reflecting our great abilities to manipulate things and speak.



(a) (b)

Figure 5.1: **The primary motor cortex.** Localization in the brain (a) and illustration of the function of its different areas (b). Image courtesy of the online encyclopedia *The brain from top to bottom*.

In patients undergoing surgery, the neurosurgeon sometimes has no choice but

to get through the CST. For example, to access a tumor, it might be his/her best shot. In this case, he/she is facing a quite hard dilemma: finding a new access (which might not be optimal and result in a more dangerous surgery), cancelling the surgery (which prevents the patient from being treated) or going through the CST at the risk of disrupting connections that may lead to partial paralysis. An accurate reconstruction of the CST, especially in the vicinity of the tumor, could therefore help the surgeon in taking the right decision. However, as we will see in the next section, the available diffusion data for neurosurgical planning has very low spatial and angular resolution, making it difficult to provide such an accurate reconstruction of the CST.

5.2 Presentation of the data provided by the organizers

In neurosurgical planning, there is no time for long or even moderately long diffusion acquisition schemes. Better yet, the patient often cannot stay long in the MRI scanner and choices need to be made whether or not performing dMRI. Arguments in favor of using it could be the possibility to provide an accurate reconstruction of the CST from very short diffusion acquisition schemes.

In each edition, two patients, each time suffering from tumors of various forms and grades, underwent dMRI. Patients were not the same for each edition. In addition to dMRI, T1, T2 weighted or FLAIR images were available along with the delineation of the tumors (including edematous region too) so that joint views of anatomical and tractography information could be created to help the surgeon, for example, to plan a tumor removal surgery. In the following, Year 1 refers to 2011, which is the year of the 1st challenge. Similarly, Year 2 refers to 2012 and Year 3 to 2013. Subsequently, Patient 1.1 refers to Year 1 Patient 1, Patient 1.2 to Year 2 Patient 1 and Patient 1.3 to Year 3 Patient 1 (same thing for the Patients 2).

In the first and second edition of the DTI tractography challenge, the characteristics of the dMRI data were the following:

- the in-plane spatial resolution of the measured DW images was high (256×256) and the slice thickness relatively small leading to voxels of size $1 \times 1 \times 2.6$ mm^3,
- the acquisition time was shortened using a single-shell acquisition scheme at b = 1000 s/mm^2 on which only 31 DSGs uniformly distributed over the north hemisphere were sampled.

In the first edition of the challenge, the organizers also provided us with two healthy subjects scanned with a multi-shell acquisition scheme at $b_{\max}$ = 1000 s/mm^2 with 31 DSGs. For each, dMRI data were acquired repeatedly (ten repetitions) so that the reproducibility of the tractography method may be evaluated.

In the third edition, two patients underwent dMRI as well and the same shell at b = 1000 s/mm^2 was also adopted. However, the number of sampled DSG directions was reduced to 20 and the slice thickness increased to 5 mm in order to save acquisition time. This yielded very low spatial and angular resolution diffusion data that are of very poor quality.

We applied the following pre-processing pipeline to all challenge datasets. An expert neuroanatomist delineated two ROIs on each side (left and right) of the original T1-weighted images. One is located in the posterior limb of the internal capsule and the other in the superior part of the mesencephalon. In addition, since crossing fascicles such as the association or commissural fascicles may be considered as bifurcations of the CST by the tractography algorithm, the same expert also delineated regions through which the tracts were not allowed. All datasets were then pre-processed according to the following two steps:

1. Diffusion-weighted MRI denoising: dMRI is subject to random noise yielding measures that are different from their real values, and thus biasing the subsequently estimated diffusion models. We filtered the measured DW images with the Rician-adapted non-local means filter [Wiest-Daesslé 2008], which has been shown to efficiently denoise such images while preserving fine anatomical structures.

2. ROI alignment on B0 images: we registered the ROIs on the B0 images according to the following steps:

- global affine registration of the T1-weighted images to the B0 images [Ourselin 2000];
- cropping of the affine-registered T1-weighted images using the mask of the B0 images;
- constrained non-rigid registration [Garcia 2010] of the masked T1-weighted images to the B0 images;
- application of the obtained transformations to the ROIs.

Note that when the T1w image was not available, the T2w or the FLAIR image was used instead.

The subsequent tractography pipeline involves three diffusion tools: diffusion modeling, identification of the most appropriate model and tractography algorithm. Throughout my thesis, these tools underwent some improvements. Before presenting our results for each edition of the challenge, we shall thus briefly make a status report of our diffusion tools at that time.

For each edition of the challenge, the DDI MCM introduced in section 2.3.4.1 was used. However, facing different problems in its elaboration, we did not apply the same constraints on the parameters each year. In its most general form, the M-fascicle model has the following $6M + 2$ parameters:

- the polar angles θ_i of the orientation of each fascicle,
- the azimuth angles ϕ_i of the orientation of each fascicle,
- the orientational dispersion indices κ_i in each fascicle,
- the free diffusivity d_i in each fascicle,
- the extra-axonal fraction of occupancy ν_i in each fascicle,
- the fraction of occupancy w_i of each fascicle,
- the diffusivity d_0 in the CSF,
- the fraction of occupancy w_0 of the CSF.

For each edition, we will then explain which constraints were applied to these parameters and which optimization algorithm was used to solve the least square problem given by eq. (2.68). Our choices will be discussed in the last section.

5.3 2011 Challenge Results

5.3.1 Status report of our diffusion tools

By the time of the first edition of the DTI Tractography Challenge, the status of different tools was the following:

• *Diffusion Modeling.* At that time, we constrained the parameters as follows:
– $\nu_i = 1/2$: we account for intra- and extra-axonal spaces in equal proportions in modeling the diffusion within a fascicle,
– $d_i = (\kappa_i + 1)\lambda$: the perpendicular diffusivity λ was estimated but identical in all fascicles and the free diffusivity d_i in each fascicle was set dependent on the orientational dispersion index κ_i and the common perpendicular diffusivity λ,
– $w_0 = 1 - \sum_{i=1}^{M} w_i$: the CSF fraction of occupancy was set from the fractions of occupancy of the fascicles so that the different mixture weights to sum to one,
– $w_i = \mathrm{FA}_i^{\mathrm{approx}}/M$: the fraction of occupancy of each fascicle was set proportional to its approximated FA computed according to eq. (4.1).
These constraints yielded a $3M + 1$ DDI MCM. The parameters were transformed into unbounded parameters leading to a nonlinear unconstrained minimization problem given by eq. (2.68) that we solved using NEWUOA [Powell 2006]. This optimization algorithm is indeed designed for unconstrained optimization when derivatives are not analytically available.

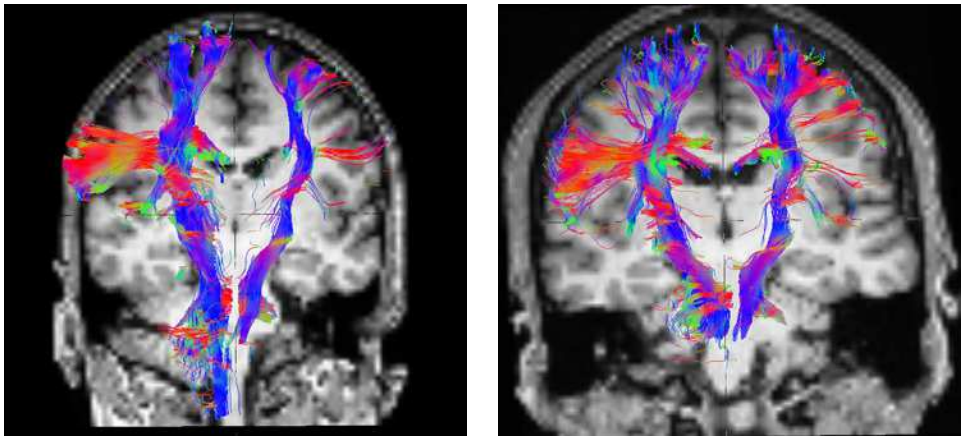
• *Identification of the most appropriate model.* No model selection and/or averaging was performed: a 2-fascicle non-averaged DDI MCM was estimated in each voxel of the brain.

• *Tractography Algorithm.* At that time, we only managed to re-implemented the FACT deterministic streamline algorithm [Mori 1999], adapted to account for multiple non-collinear orientations of diffusion provided by MCMs. The following parameters were used:
– $N = 2$: number of starting points per voxel of the seed region,
– $\mathrm{FA}_t = 0.5$: tracts reaching a voxel with too low FA are stopped,
– $l = 1$ mm: at each step, we update the tract one step of such length ahead,
– $r = 0.8$: relative threshold on orientational dispersion indices above which the less orientationally dispersed orientation of diffusion is favored upon the most collinear one,
– $\alpha_t = 60°$: threshold to stop the tract if the new direction is not sufficiently collinear with the previous one.

5.3.2 Tractography on Healthy Subjects

For each of repetition of each subject, the left and right CSTs were computed using the aligned ROIs delineated according to the process described in section 5.2. We display the tractography of one volume of each of the two subjects in fig. 5.2.

This figure illustrates that we are able to cover the full extent of the CST for these healthy subjects, from the face and hand areas to the medial part. This demonstrates that our diffusion model enables the tractography algorithm to follow



(a) (b)

Figure 5.2: **Illustration of reconstructed CSTs on Year 1 Healthy Subjects.** Representative examples of obtained CSTs for healthy subject 1 (a) and 2 (b) (T1 images are in radiological conventions, i.e. the left hemisphere is on the right side of the image). To see the full extent of the CST spreading, all 3D tracts are displayed. This explains why they may not seem to match exactly the background T1 image.

bifurcations in the WM. For each subject, estimating the DDI models from the measured DW images took approximately 40 minutes (single-threaded), while the tractography of the left and right CSTs took less than 10 minutes.

5.3.3 Tractography on Patients

We also reported results for the CST of each patient on the side of the tumor. The images processed here had a larger spatial resolution than for the healthy subjects and the estimation of the DDI models from the measured DW images took approximately 3 hours (single-threaded). However, it is possible to compute this stage of the processing pipeline offline, leaving only the tractography to perform online (about 5 minutes for each CST) when performing the tractography for surgical planning.

We focus on the qualitative evaluation of the obtained fascicles and their closeness to the tumor (it should be noted that the tumor regions were not used in any way to constrain the tractography algorithm), and on providing the neurosurgeon with helpful views for neurosurgery planning. To this end, we present in fig. 5.3 views combining the reconstructed CSTs, the tumor ROIs provided by the organizers, all on top of the patient's T1 image.

We can observe on this figure that, although the tumor delineation was not used in the algorithm, no tracts are going through the tumor area on patient 1 but some are going through the most central part of the edema (fig. 5.3b,c). This indicates that the tracts were pushed by the tumor mass effect, which is a valuable indication when planning the surgery. Overall, this figure demonstrates the close proximity of the tumors and of the CST for both patients. This is an important insight as the

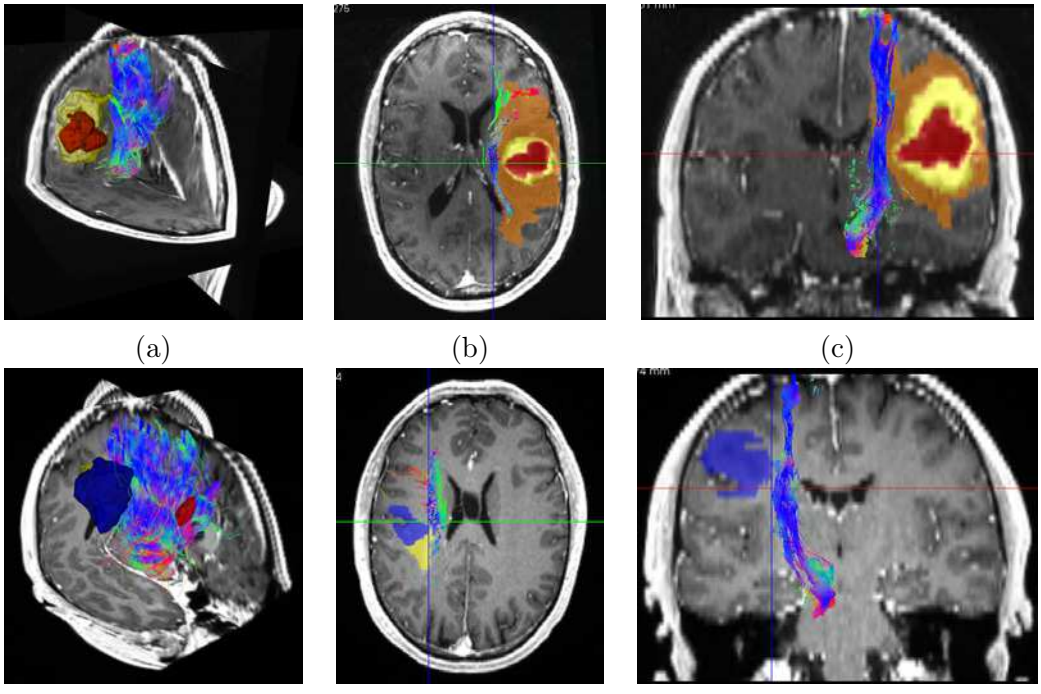


Figure 5.3: **Illustration of reconstructed CSTs on Year 1 Patients.** Combined views of the obtained CSTs for Patient 1.1 (a,b,c) and Patient 1.2 (d,e,f), illustrating the proximity of the tumor to crucial motor pathways. Surfaces for Patient 1.1 correspond to the necrotic part of the tumor (red), the active part of the tumor (yellow) and the edema (orange). For Patient 1.2, each ROI corresponds to a specific tumor. Images (a,d) show overall 3D views and (b,c,e,f) show the tracts and ROIs going through a specific 2D slice to better illustrate their proximity.

neurosurgeon will be able to plan the tumor removal in the optimal way, so as to minimize the possible handicap for the patient after surgery.

5.4 2012 Challenge Results

5.4.1 Status report of our diffusion tools

By the time of the second edition of the DTI Tractography Challenge, the status of different tools was the following:

- *Diffusion Modeling.* At that time, we constrained the parameters as follows:

– $\nu_i = 1/2$: we account for intra- and extra-axonal spaces in equal proportions in modeling the diffusion within a fascicle,
– $w_0 = 1 - \sum_{i=1}^{M} w_i$: the CSF fraction of occupancy was set from the fractions of occupancy of the fascicles so that the different mixture weights sum to one,
– $d_0 = d_0(M = 0)$: the diffusivity in the CSF was estimated with the 0-fascicle DDI MCM and then fixed to its estimated value for estimating any M-fascicle DDI MCM

with $M > 0$.
These constraints yielded a $5M$ DDI MCM. The parameters were directly estimated by defining a nonlinear constrained minimization problem from eq. (2.68) that we solved using Bound Optimization BY Quadratic Approximation (BOBYQA) [Powell 2009]. This optimization algorithm is indeed designed for constrained optimization when derivatives are not analytically available. Approximated FA was still computed from eq. (4.1) for further use in the tractography algorithm.

• *Identification of the most appropriate model.* A first model selection approach was used. The estimation of free-water diffusivity was performed using a 0-fascicle DDI model. Subsequent 1-fascicle and 2-fascicle DDI MCMs were estimated fixing the free water diffusivity to its previously estimated value. In each voxel, the 1-fascicle or 2-fascicle MCM was selected based on the value

$$\kappa_{\mathrm{DTI}} := 2\lambda_1/(\lambda_2 + \lambda_3) - 1\,, \tag{5.1}$$

where $(\lambda_1, \lambda_2, \lambda_3)$ are the eigenvalues of the DT: if $0.3 < \kappa_{\mathrm{DTI}} < 1.2$, the 2-fascicle model was selected whereas $\kappa_{\mathrm{DTI}} \leq 0.3$ was found to relatively well indicate isotropic diffusion and $\kappa_{\mathrm{DTI}} \geq 1.2$ was found to relatively well indicate sufficiently strong anisotropy to assume the presence of only 1 fascicle.

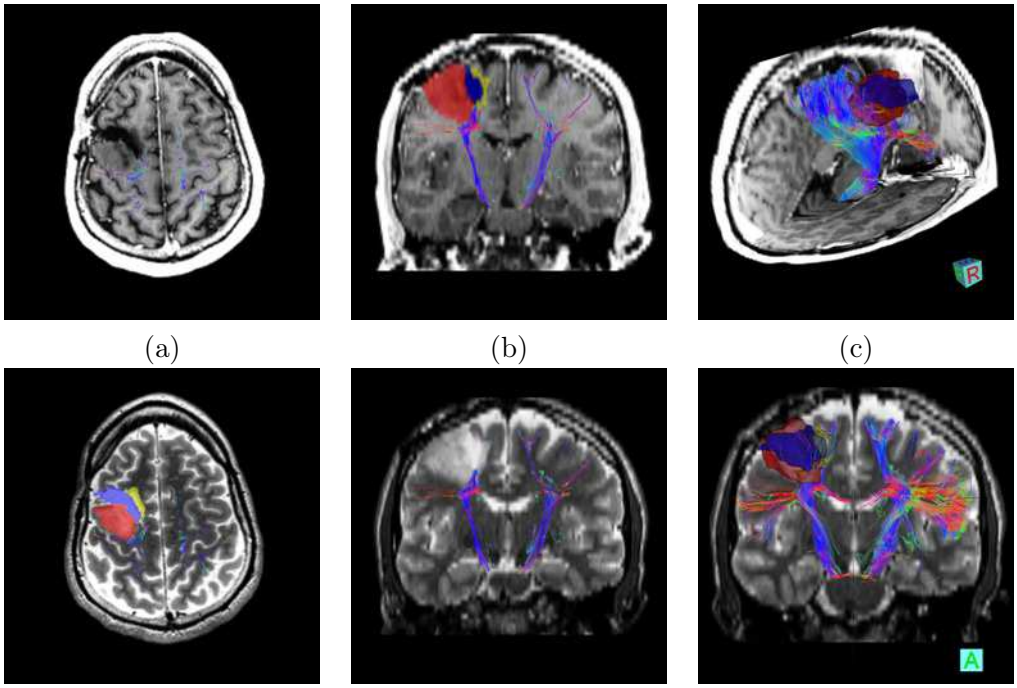
• *Tractography Algorithm.* At that time, we kept using our re-implemention of the FACT deterministic streamline algorithm [Mori 1999], adapted to account for multiple non-collinear orientations of diffusion provided by MCMs. Slightly different parameters were used though:
– $N = 1$: number of starting points per voxel of the seed region,
– $\mathrm{FA}_t = 0.2$: tracts reaching a voxel with too low FA are stopped,
– $l = 1$ mm: at each step, we update the tract one step of such length ahead,
– $r = 0.8$: relative threshold on orientational dispersion indices above which the less orientationally dispersed orientation of diffusion is favored upon the most collinear one,
– $\alpha_t = 60°$: threshold to stop the tract if the new direction is not sufficiently collinear with the previous one.

On both patients, the estimation of the DDI volume took about 10 minutes on an 8-core Xeon 3 GHz computer while the tractography of each CST took about 1 to 2 minutes on the same machine. The model estimation step is longer but still reasonably short and can be performed offline in a clinical context.

5.4.2 Tractography on Patient 2.1

Figure 5.6 presents axial and coronal views of the left and right CST extracted on Patient 2.1, superimposed on the T1w image or the T2w image. In addition, we superimposed the manual delineations and provide 3D combined visualizations to evaluate the relative positions of the tumor and motor pathways.

Several things may be noticed for this patient. First, the left reconstructed CST, which is not affected by the tumor, is able to reach the different motor cortical areas (fig. 5.6b,e,f). On the contrary, the right CST is going neither through the cavity



(a) (b) (c)

(d) (e) (f)

Figure 5.4: **Illustration of the reconstructed CSTs on Patient** 2.1. Combined views of the obtained CSTs overlayed on T1 (a,b,c) and T2 (d,e,f), illustrating the proximity of the tumor to crucial motor pathways. ROIs correspond to the tumor (red), the resection cavity (blue) and gliosis (yellow). (a,d,b,e) are respectively axial and coronal views, while (c,f) are 3D views.

(which was expected) nor through the tumor. Therefore, the number of tracts found on the right side is lower than on the left side of the brain as shown in fig. 5.6f. Some tracts in the right CST are going through the gliosis. Finally, the 3D views clearly highlight a very close proximity between the tumor and the motor pathways, and therefore provide crucial information to the surgeon in case another surgical resection is planned.

5.4.3 Tractography on Patient 2.2

The tumor for Patient 2.2 is different as it is an infiltrating tumor. As for Patient 2.1, we present in fig. 5.5 axial and coronal views of the left and right reconstructed CSTs, superimposed on the T1w image or the T2w image, as well as 3D combined visualizations to evaluate the relative positions of the tumor and motor pathways.

As for Patient 2.1, the left CST is well reconstructed, showing some bifurcations towards several regions of the motor cortex. The number of tracts in the right CST is lower than in the left CST (fig. 5.5b,e,f), which is due to the presence of the edema. However, contrary to patient 1, the right CST is not completely blocked by the tumor as the latter is of different nature. The most central part of the right CST is indeed still visible in our tractography (see axial and coronal views on fig. 5.5b,d). However the outer parts (e.g. corresponding to the hand or face areas) are much sparser or even not found by our tractography algorithm. Again, both 3D views provide an important insight on the relative positions of the tumor, the edema and the CST, which can help the neurosurgeon in his/her decisions.

5.5 2013 Challenge Results

5.5.1 Status report of our diffusion tools

By the time of the third edition of the DTI Tractography Challenge, the status of different tools was the following:

- *Diffusion Modeling.* At that time, we constrained the parameters as follows:

– $d_0 = 3.0 \times 10^{-3}$ mm^2/s: the diffusivity in the CSF was fixed to its value at 37°C except for the isotropic model ($m = 0$) in which it is estimated,

– the diffusivity parallel to the orientation of the fascicle was set to 1.71×10^{-3} mm^2/s, which is the only quantity that has been consistently estimated by several diffusion models (e.g. DTI, DKI); it yields a relationship between the free diffusivity d_i, the orientational dispersion index κ_i and the extra-axonal fraction of occupancy ν_i in each fascicle given by eq. (2.50),

– the concentration parameters κ_i and extra-axonal occupancies ν_i were assumed to be identical in all the fascicle compartments.

This yielded the constrained DDI MCM with $3M + 2$ independent parameters introduced in section 2.3.4.3. The parameters were directly estimated by defining a nonlinear constrained minimization problem from eq. (2.68) that we solved using

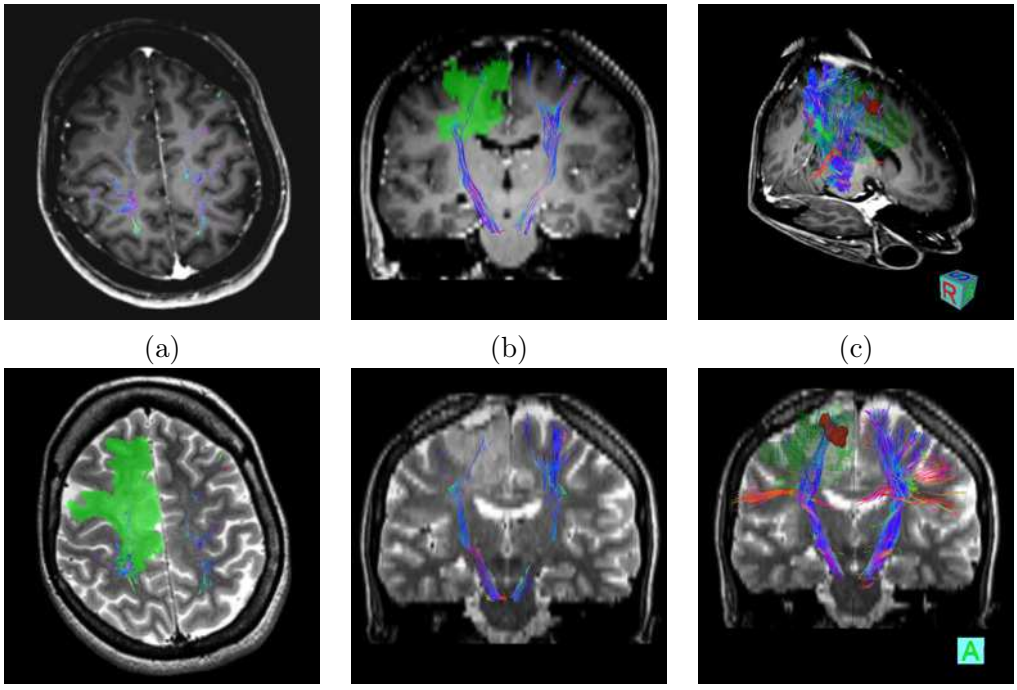


Figure 5.5: **Illustration of the reconstructed CSTs on Patient** 2.2. Combined views of the obtained CSTs overlayed on T1 (a,b,c) and T2 (d,e,f), illustrating the relative positions of the tumor and the motor pathways. ROIs correspond to the solid tumor (red) and the infiltrating tumor (green). (a,d,b,e) are respectively axial and coronal views, while (c,f) are 3D views.

COBYLA [Powell 1994]. Like BOBYQA, this optimization algorithm is indeed designed for constrained optimization when derivatives are not analytically available. The difference is that BOBYQA progresses towards minima by quadratic approximations of the cost function while COBYLA uses linear approximations. As a result, BOBYQA is faster but may lead to fictitious minimas. We thus adopt COBYLA. For each fascicle, the real FA of the distribution was computed according to eq. (2.44) for further use in the tractography algorithm.

• *Identification of the most appropriate model.* The DDI MCMs with 0, 1, 2 and 3 fascicles were estimated and our model averaging approach was used to obtain a 3-fascicle averaged DDI MCM.

• *Tractography Algorithm.* We used our adaptive multi-modal particle filter that approximates the distribution of the fascicles arising from a specific region given the diffusion data while maintaining the multi-modal property of this distribution, which allowed us to track multiple fascicles over extended volumes. A summary of the algorithm and its parameters is given in section 4.2.1.5. In this specific case, the following parameters were used:

– $\text{FA}_t = 0.5$: tracts reaching a voxel with too low FA are stopped,
– $\rho = 1$ mm: at each step, we update the tracts one step of such length ahead,
– $N = 5000$: this is the number of particles used to approximate the true distribution of the tracts in the CST (the more particles, the more clusters and thus the more tracts on the lateral parts),
– $L_{\max} = 20$ cm: tracts that reach this maximum length are stopped,
– $N_{\min} = 10$: particle clusters of this size are not allowed to be splitted anymore,
– $\alpha = 0.8$: if the effective sample size in each cluster drops below this relative threshold, particles are resampled according to their weights,
– $w_t^{\text{CSF}} = 0.8$: tracts reaching a voxel too strongly contaminated by CSF, characterized by a CSF fraction of occupancy greater than this threshold, are stopped,
– $\kappa = 30$: tracts are supposed to be smooth enough, which can be controlled by this parameter (concentration of the prior vMF distribution),
– $h_t = 0.5$ mm: clusters are merged if the Hausdorff distance between the global cluster mean particles drops below this treshold,
– $\kappa_t = \kappa = 30$: clusters can be splitted if their particles lose orientational coherence, characterized by the concentration of the local cluster directional distribution.

On both patients, the estimation and averaging of the DDI volumes took about 150 minutes on a 8-core Xeon 3 GHz computer while the tractography of each CST took about 25 minutes on the same machine for $N = 5000$ initial particles (the computational time is proportional to N). The model estimation step is longer but still reasonably short and can be performed offline in a clinical context.

5.5.2 Tractography Results

Figure 5.6 presents axial and coronal views of the left and right CSTs extracted on both patients, superimposed on the T1w image for Patient 3.1 and on the FLAIR image for Patient 3.2. In addition, we superimposed the manual delineations and

provide 3D visualizations to evaluate the relative positions of the tumor and motor pathways.

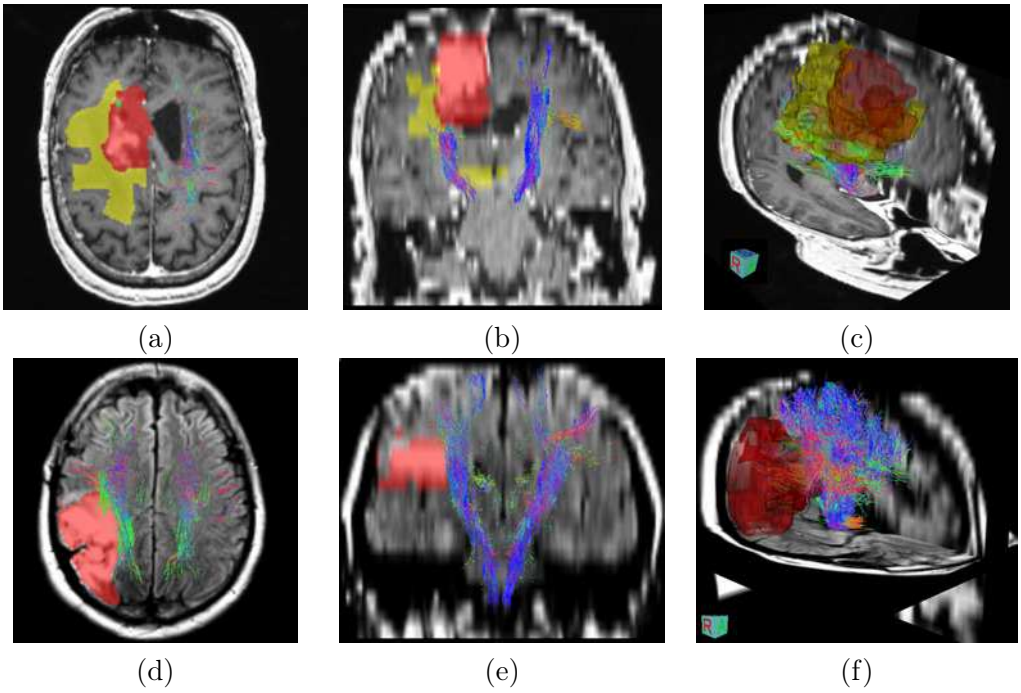


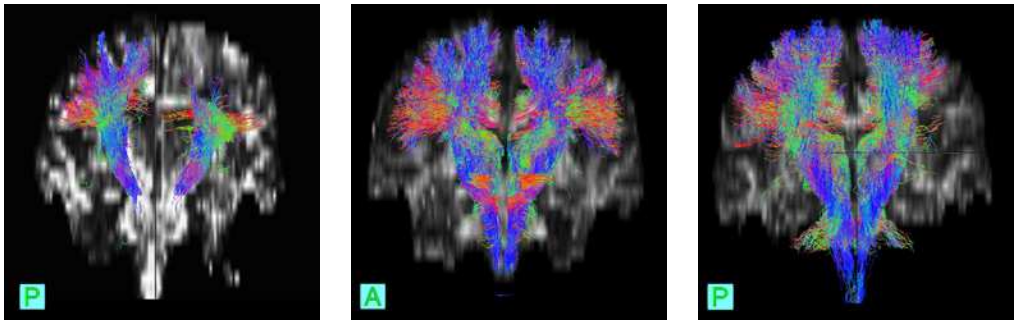
Figure 5.6: **Illustration of the reconstructed CSTs on the Year 3 Patients.** Integrated views of the obtained CSTs of Patient 3.1 overlayed on T1 (a,b,c) and of Patient 3.2 overlaid on FLAIR (d,e,f), illustrating the proximity of the tumor to crucial motor pathways. ROIs on Patient 3.1 correspond to the solid tumor (green), the tumor (red) and edema (yellow). The ROI on Patient 3.2 corresponds to the tumor (red). (a,d,b,e) are respectively axial and coronal views, while (c,f) are 3D views.

Several facts may be noticed for Patient 3.1. First, the left CST, which is not affected by the tumor, is able to spread into the different cortical motor areas (fig. 5.6a,b). On the contrary, the right CST is not going into the tumor part (red and green, fig. 5.6a,b) while it still is able to capture some bifurcations and goes a little in the edema area (fig. 5.6a-c). Therefore, the number of tracts found on the right side is lower than on the left side of the brain. Finally, the 3D view highlights a close proximity between the tumor and the motor pathways, and therefore provide crucial information to the surgeon in case another surgical resection is planned.

The second row of fig. 5.6 presents results for Patient 3.2. Both CSTs are well extracted, showing some bifurcations towards several areas of the motor cortex. The number of tracts in the right CST is lower than in the left CST (fig. 5.6d,e), which is due to the presence of the tumor. The tracts are indeed not going through this specific area and therefore spread less than on the left side. Again, the 3D view provides an important insight on the relative positions of the tumor and the CST, showing a very close proximity of the interior part or the tumor with the CST which

will help the neurosurgeon in his/her decisions.

Figure 5.7 shows the reconstructed CSTs on both patients overlayed on the FA map. For Patient 3.1 (fig. 5.7a), we can see that the tumor completely prevents the CST from reaching the cortical motor areas except for the most lateral areas. For Patient 3.2 (fig. 5.7b,c), we observe that the CST covers all the areas of the motor cortex away from the tumor (fig. 5.7b) but the most lateral parts are not reached in the vicinity of the tumor (fig. 5.7c).



(a) (b) (c)

Figure 5.7: **Reconstructed CSTs overlayed on the FA map.** Coronal view of the reconstructed CSTs overlayed on FA map for Patient 3.1 (a) and Patient 3.2 (b,c). In particular, on Patient 3.2, we observe that the CST covers all the areas of the motor cortex away from the tumor (b) but the most lateral parts are not reached in the vicinity of the tumor (c).

5.6 Synthesis of our participation

Table 5.1 summarizes the diffusion tools that were used for the different editions of the DTI challenge. In this table, we recall the constraints applied for DDI estimation, the number of fascicle compartment used in the DDI MCM, the optimization algorithm used to minimize eq. (2.68), the tractography algorithm used to perform the CSTs reconstruction and the running times for both model estimation and tractography steps. Each year, we improved part of the process. Between Year 1 and Year 2, constraints on the model were relaxed and an hybrid model selection was performed between 1-fascicle and 2-fascicle DDI MCMs. Also, the optimization problem was greatly simplified by the use of BOBYQA, which is adpated to bound constraints. Between Year 2 and 3, we conducted a thorough study of the constraints applied to the model that led us to estimate for each fascicle only its orientation and fraction of occupancy. From our experiments, in order for the estimation problem to be well-posed, it was then required to estimate a single orientation dispersion index *kappa* and extra-axonal occupancy ν for all the fascicle compartments. We also used our explicit model averaging and estimated the DDI MCM up to 3 fascicles. We understand that progressing towards minima of the cost function given in eq. (2.68) by quadratic approximations sometimes led to values that are in fact not minima of

the cost functions, which might be explained by the cost function not being twice differentiable or by a non-smooth second derivative. We thus perform optimization using COBYLA. Last but not least, we developed a probabilistic tractography algorithm that accounts for uncertainty on the estimated fascicle orientations and explore possible tracts much more efficiently than our deterministic multi-modal one used for the first two years.

All the improvements were motivated by the will to provide clinically applicable diffusion tools that can bring more information than DTI without requiring additional training for the medical staff. In clinical applications, time is crucial.

Since the advent of dMRI, time has been saved in the data acquisition step by using large slice thickness and by using few DSGs whose directions are sampled on a single shell and uniformly distributed (often on the north hemisphere). Clinicians want at least to keep it that way, if not to further reduce the scan time in situations like neurosurgical planning. In these cases, we might very well obtain very low spatial (very large slice thickness) and angular resolution (very few DSGs encoding directions on a single shell) diffusion data. Saving time on the second step rely on optimally constraining the diffusion models. By optimal we mean finding appropriate constraints that are physically plausible but not constraining too much in order to still be able to describe the different populations of cells in the WM voxels, which are believed to be very good biomarkers for numerous neurological disorders (see section 1.4.3). In summary, following the MMWMD [Alexander 2010], we shall understand which biomarkers are crucial and which parameters can be constrained to obtain a fast (the less independent parameters to be estimated, the faster the estimation) and accurate (the less independent parameters to be estimated, the less sensitive the estimation procedure to local optima) estimation of the diffusion model. Finally, the last tractography step is usually fast because it uses the diffusion models previously estimated as input and interpolate them when necessary without performing the model estimation at each step of the reconstruction of the fascicles. Still, it can be speeded up by either decreasing the number of starting points per voxel of the seed regions in deterministic streamline algorithms or decreasing the number of particles used to approximate the true distribution of tracts in the reconstructed fascicle in particle filter-based probabilistic algorithms.

In the following sections, we discuss our proposed diffusion tools and their clinical applicability compared to the ones in the literature. When time is of the essence, we need to save it as much as possible. Three diffusion tools have been developed in this work and their efficiency tested concretely on the ground in the very constraining case of neurosurgical planning. In the following, we analyze the choices we made to save time from the data acquisition to the actual reconstructed CSTs for each edition of the challenge. The reconstruction of WM fascicles involves the following steps that are (i) data acquisition (discussed in section 5.6.1, diffusion modeling and subsequent estimation (discussed in section 5.6.2), (iii) model selection/averaging (discussed in section 5.6.3) and (iv) tractography (discussed in section 5.6.4.

5.6.1 Data acquisition

Diffusion data were provided by the organizers of the challenge and consisted over the years of poorer and poorer image quality. We had to deal with data from 31 DSGs repeated 10 times and with 2.6 mm slice thickness (first edition) to 20 DSGs repeated 4 times and with 5 mm slice thickness (third edition). Our decision to confront our diffusion tools to this kind of data demonstrates our intention to make tractography feasible to basic clinical constraints.

5.6.2 Diffusion modeling and subsequent estimation

The DDI model is an MCM. The advantage of using MCMs with respect to other diffusion models is that not only they try to identify the correct orientations of the fascicles but they also characterize the populations of cells which compose the fascicles. However, [Scherrer 2010] mathematically demonstrated that any MCMs belonging to the exponential family - hereafter referred to as exponential MCMs - require diffusion data obtained with multi-shell acquisition schemes, otherwise the model estimation problem is ill-posed: in essence, several sets of parameters exist for exponential MCMs that all yield the same SID on a single shell and thus result indistinguishable. For this reason, the DDI MCM is based on a mixture of PDFs that does not belong to the exponential family, which enables its estimation from single-shell acquisition schemes. On another hand, the estimation of MCMs relies on the minimization of a cost function, which is quite handy in the case of exponential MCMs but becomes much more complex with the DDI MCM. In the first edition of the challenge, we faced this problem and the constraints that we applied back then were mainly designed to simplify the cost function. For instance, setting the extra-axonal fraction of occupancy to $\nu = 1/2$ has no particular physical meaning but has the merit to account for intra- and extra-axonal spaces without favoring one with respect to the other. Also, setting the fascicle fractions of occupancy proportional to their FA does not exactly reflect the volume occupied by the fascicle in the voxel but tends to retain only fascicles with low orientational dispersion, which is not out of sense. Finally, estimating a common perpendicular diffusivity for all fascicles could be seen as a strange constraint but is in fact based on [Scherrer 2010], which showed that the perpendicular diffusivity is responsible for the exponential MCM identificability issue. For the second edition of the challenge, we partially solved our optimization problems using a better implementation of the BOBYQA algorithm and could free most of the constraints. In particular, the constraint on the perpendicular diffusivity appeared useless since the DDI MCM is not an exponential MCM and fascicle fractions of occupancy were not constrained anymore. Our experiments however revealed the necessity to fix the diffusivity in the CSF, which plays a leverage role in the estimation in the sense that if it is not correctly estimated, then all the other parameters will be biased too. We thus naively fixed it to its estimated value using a previously estimated isotropic diffusion model. For the third edition, we noticed that the cost function involved in the model estimation cannot

always be approximated by a quadratic function, in which cases the BOBYQA algorithm performs poorly. We thus replaced it with the COBYLA algorithm that relies on linear approximations. This solved all the remaining optimization issues at the cost of computational time. To speed up the estimation process, we thus initialized the orientations of the fascicles in the model by a previously estimated ball-and-stick model [Behrens 2003], which is an asymptotic DDI MCM (for $\kappa_i \to \infty$ and $\nu_i \to 0$). All previous constraints were also freed and replaced with physically meaningful constraints: we fixed the diffusivity in the CSF to its known value at 37°C, we also fixed the parallel diffusivity in the fascicles to a fixed value, since it is the only metric that has been consistently estimated by several different diffusion models and we were able to analytically express the volume occupied by a fascicle in the voxel according to the parameters that characterize it, which allowed us to set the fractions of occupancy of the fascicles proportional to this volume. This constrained DDI MCM is, to the best of our knowledge, the best compromise to obtain a fine description of tissues microstructure from single-shell low spatial and angular resolution measured DW images.

5.6.3 Model selection/averaging

This additional step before tractography is a specificity of MCMs, in which one has to decide a priori the number M of fascicles to estimate in a given voxel. It usually relies on estimating several M-fascicle MCMs and choosing the one that best fits the data. Differently, our approach consists in arguing that the question "how do we choose *a priori* the right number of fascicles in a given voxel?" does not have an answer. We rather quantify probabilities that a voxel contains 1, 2,..., M fascicles based on the diffusion data and average the different corresponding MCMs accordingly. This step takes less than half a minute when averaging the models up to 3 fascicles.

5.6.4 Tractography

This is the actual reconstruction of the fascicles based on previously estimated diffusion models. It is generally the fastest step. In the first two editions of the challenge, we used a deterministic streamline tractography algorithm. The reason of such a choice was that using an MCM that has been accurately estimated provides orientations of diffusion that present very low bias: as a consequence, following these orientations should not result in an important accumulation of errors. However, MCMs are known to often overfit the data. Even if the model averaging step is meant to reduce this overfitting, the estimation of the probability that a given model is well adapted to fit the data is based on measures that are sensitive to overfitting, which means that overfitting could still persist.

Also, we adapted the streamline algorithm to account for multiple fascicle orientations provided by MCMs as follows: (i) starting from a single voxel, we first reconstruct the main tract by locally choosing as outcoming orientation the ori-

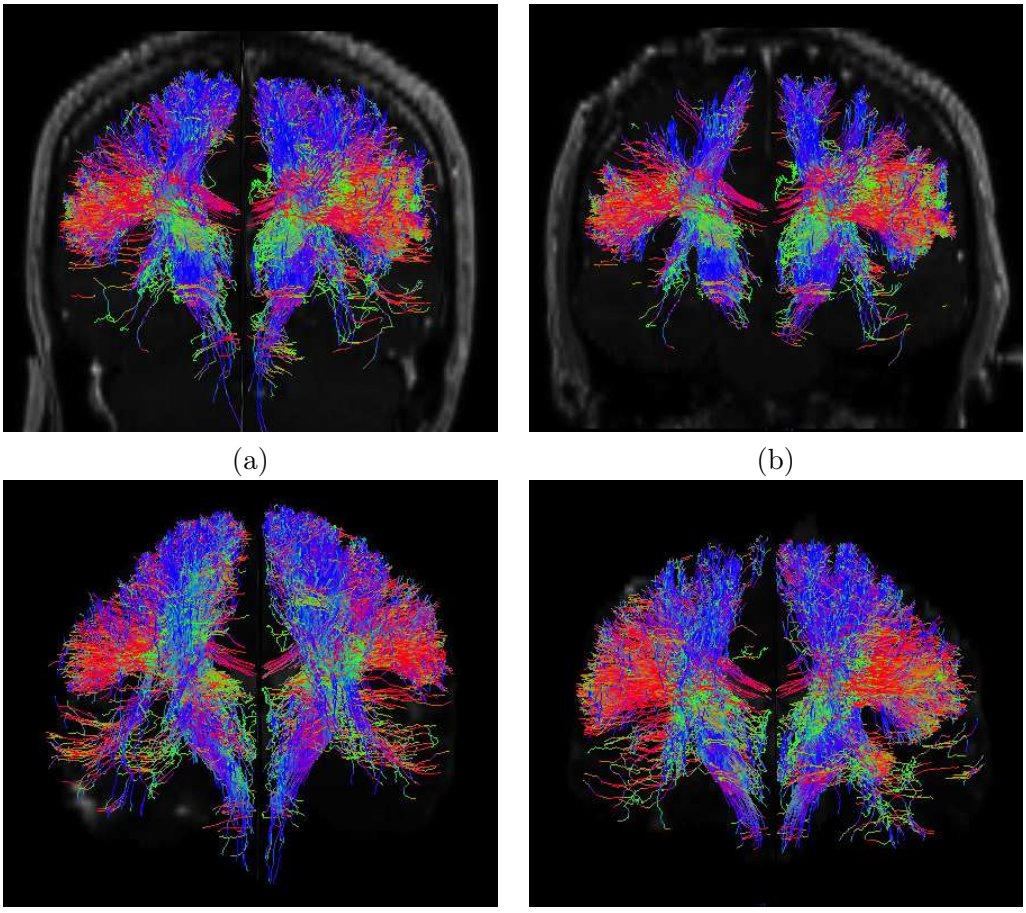
entation the more collinear to the incoming orientation among the set of all the orientations provided the MCM and record the other identified orientations for future use, (ii) when the main tract is reconstructed, we start from scratch to explore the closest bifurcation and (iii) when a bifurcation is reconstructed, we repeat stage (ii). If one wanted, for instance, to cover the full extent of the CST, this would be of an impressive complexity and would lead to such huge tract files that we are not sure to be able to stock. This is visible from fig. 5.2 where the tracts going straight to the top of the head are well represented (main path) and the tracts to the hand and face area also (first explored bifurcations). Tracts in the middle of the motor cortex are absent because the algorithm does not even explore the recorded bifurcations that would lead to these areas. To overcome this issue, our multi-modal probabilistic algorithm is smarter as it does not explore the bifurcations on a first-come, first served basis but rather explore them all sequentially and assign them a probability of existing given the diffusion data. When such a probability is too low they are suppressed, so that the algorithm finally focused on the most probable tracts given the data. This strategy is very promising since, together with the DDI MCM, it manages to cover the full extent of the CST, as shown in fig. 5.7. One might have the impression that the results of the second edition (figs. 5.5 and 5.6) are worse than those of the first edition (fig. 5.3). This is in fact an optical effect due to the fact that, for the first edition, we initialized 2 tracts per voxel of the seed region while we initialized only 1 tract per voxel of the seed region the second year. This has in fact a really low impact on the extent of the reconstructed CST but yields much manageable file sizes. On the contrary, with our probabilistic tractography algorithm, using $N = 5000$ of particles to approximate the true distribution of tracts within the CST allows us to better cover the full extent of the CST than if we had used a smaller number. The explanation is quite simple. The algorithm sequentially classifies the tracts in clusters (that ideally would represent sub-fascicles reaching a specific cortical area), which need to have a minimum size (i.e., a minimum number of tracts inside) for their average tract to actually mean something. We set the minimal size of the clusters to 10 particles. This implies that a maximum of 100 bifurcating tracts will be retained if $N = 1000$. This maximum increases to 500 if $N = 5000$, which allows the algorithm to find more bifurcating tracts that still have high probability of existence given the diffusion data.

5.6.5 Conclusive remarks

In summary, we would like to acknowledge the efforts that the organizers made to open our minds to real-life application and not only to our well-designed and long experiments. Neurosurgeons in particular are constantly in a race against time, in which the data quality will not improve if not impoverishes more. Mathematicians and computer scientists should therefore work together to make the most out of what they can get, rather than making the right data to acommodate their tools.

In this perspective, we designed a diffusion model, the DDI MCM and a multi-modal probabilistic tractography algorithm that are coping well with clinical data,

at least in the reconstruction of the CST. The model estimation step is still quite long (150 minutes) but could be performed offline, leaving the surgeon with only the tractography to perform online. Using $N = 1000$ particles leads to a very reasonable computational time (5 minutes). If there is more time, we recommend to increase N to at least 5000, keeping in mind that the computational time varies linearly with N. Figures 5.8 and 5.9 show the reconstructed CSTs for the two patients of Year 2 using the most recent elaborated diffusion tools that are the constrained DDI MCM proposed in section 2.3.4.3 and the adaptive multi-modal particle filter for tractography proposed in section 4.2. They can be compared to figs. 5.5 and 5.6 and one can conclude, at least qualtatively, that we did make a significant improvement over our diffusion tools available one year ago. The current diffusion tools seem to be able to reconstruct the full extent of the CST, contrary to the tools at Year 2 and provide interesting information in the peritumoral regions that need further evaluation.



(a) (b)

(c) (d)

Figure 5.8: **Year-3 diffusion tools applied on Year-2 data: 3D view of the CST.** Illustration of the reconstructed CSTs in 3D, displayed in coronal view. 1-st row: patient 2.1 (tracts are overlayed on T1 image), 2-nd row: patient 2.2 (tracts are overlayed on B0 image). Front views on (a,c) and back views on (b,d). These results are to be compared to those obtained at Year 2 and given in figs. 5.5 and 5.6.

Year	Data characteristics	Constraints on DDI	Number of fascicle compartments (m)	Optimization algorithm	Tractography algorithm	Running time for estimation (min)	Running time for tractography (min)
1	$b = 1000$ s/mm^2, 31 DSG directions, $1 \times 1 \times 2.6$ mm^3 voxel size.	• $\nu_i = 1/2$; • $d_i = (\kappa_i + 1)\lambda$; • $d_{\text{CSF}} = \lambda$; • $w_i = \text{FA}_i^{\text{approx}}/m$, with $\text{FA}_i^{\text{approx}}$ from eq. (4.1). $3m + 1$ DoF: θ_i, ϕ_i, κ_i and λ.	Always $m = 2$	NEWUOA [Powell 2006] (with transformations of diffusion parameters into unbounded parameters)	Multi-modal FACT (Section 4.1)	5	2 (exploring only bifurcations adjacent to the main tract)
2	$b = 1000$ s/mm^2, 31 DSG directions, $1 \times 1 \times 2.6$ mm^3 voxel size.	• $\nu_i = 1/2$; • d_{CSF} set to pre-estimated value from 0-fascicle model. $5m$ DoF: θ_i, ϕ_i, κ_i, d_i and w_i.	$m = 0$ if $\kappa_{\text{DTI}} < 0.3$, $m = 1$ if $\kappa_{\text{DTI}} > 1.2$ and $m = 2$ otherwise, where κ_{DTI} is given by eq. (5.1).	BOBYQA [Powell 2009] (quadratic approx. of the cost function)	Multi-modal FACT (Section 4.1)	10	2 (exploring only bifurcations adjacent to the main tract)
3	$b = 1000$ s/mm^2, 20 DSG directions, $1 \times 1 \times 5$ mm^3 voxel size.	• $d_{\text{CSF}} = 3.0 \times 10^{-3}$ mm^2/s except for $m = 0$ in which case it is estimated; • $\kappa_i = \kappa$; • $\nu_i = \nu$; • d_i given by eq. (2.50). $3m + 2$ DoF: θ_i, ϕ_i, w_i, κ and ν.	$m = 0$, 1, 2 and 3 and explicit model averaging (Section 3.2.2)	COBYLA [Powell 1994] (linear approx. of the cost function)	Adaptive multi-modal particle filter (Section 4.2)	150	5 (for $N = 1000$ particles)

Table 5.1: **Synthetic View of the Diffusion Tools used for each Edition of the MICCAI DTI Tractography Challenge.** Running times are given for an 8-core Xeon 3 GHz computer.

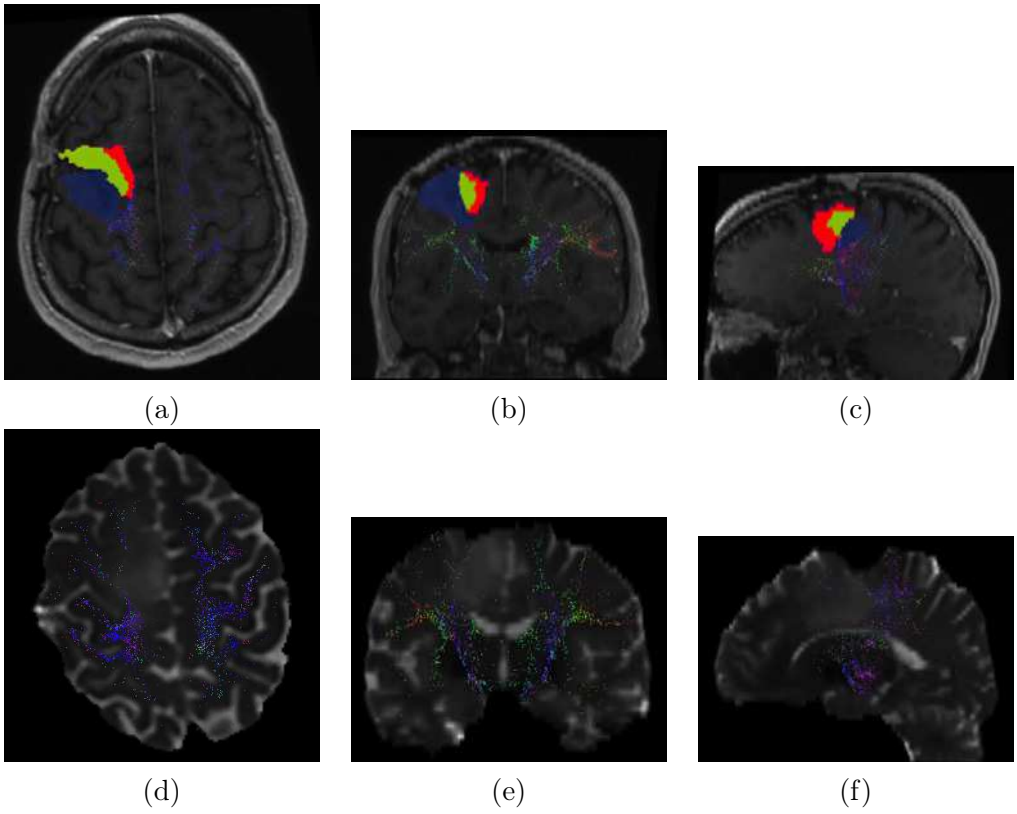


Figure 5.9: **Year-3 diffusion tools applied on Year-2 data: CST in the peritumoral regions.** Illustration of the reconstructed CSTs in the vicinity of the tumor using our Year-3 (most recent) diffusion tools on the diffusion data provided at Year 2. 1-st row: patient 2.1 (tracts are overlayed on T1 image), 2-nd row: patient 2.2 (tracts are overlayed on B0 image). Segmentation of the tumor provided by the organizers is also overlayed for patient 2.1. Axial views on (a,d), coronal views on (b,e) and sagittal views on (c,f). These results are to be compared to those obtained at Year 2 and given in figs. 5.5 and 5.6.

Chapter 6

Conclusion, Discussion and Perspectives

Contents

The main objective of the work presented in this manuscript was to achieve a high resolution reconstruction of White Matter (WM) fascicles from low angular resolution Diffusion MRI (dMRI).

The **WM** is one of the two components that constitute the Central Nervous System (CNS), in great part responsible for the transmission of the nervous information. In a normal adult Human brain, it is mainly composed of glial cells and myelinated axons and it is immersed in the Cerebro-Spinal Fluid (CSF). Axons are cigar-shaped cells whose function is precisely the transmission of the nervous information. A myelin sheath generally surrounds them to make the transmission faster. Glial cells are spherically-shaped cells whose goal is to feed and protect the axons. Myelinated axons are grouped into bundles of axons, termed **fascicles**, having locally a common orientation. In a child Human brain or a Human brain suffering from a given brain disorder, some WM fascicles may be degraded and their identification are of great interest for the understanding of brain maturation or for diagnosis/treatment of a brain disorder.

The **reconstruction** of the WM fascicles, often referred to as a "virtual dissection" of the WM, is usually achieved in three steps:

- a *data acquisition* step that consists in collecting measured Diffusion-Weighted (DW) images by the application of a series of magnetic field spatial gradients, referred to as Diffusion-Sensitizing Gradients (DSGs), that are defined by their magnitude through the so-called *b-value*, their direction and their duration;
- a *diffusion modeling* step that consists in inferring the unobserved water diffusion random process from the measured DW images;
- a *tractography* step that consists in performing the actual reconstruction of the WM fascicles assuming that orientations of diffusion identified in the diffusion modeling step correspond to the fascicle orientations.

A fascicle is not observable to the naked eye. To date, researchers have been able to perform reconstructions of simple fascicles without ramifications, referred to as low resolution reconstructions, that are in accordance with previous histological studies [Catani 2002, Jellison 2004, Catani 2008, Thiebaut de Schotten 2011]. However, many WM fascicles present an important number of ramifications (e.g., the Cortico-Spinal Tract (CST)) and/or cross in different parts of the brain. Their reconstruction, qualified as **high resolution**, proves much more difficult [Dell'Acqua 2012]. There are two ways of tackling this issue:

- The *acquisition-based* strategy: this is based on the idea that the more data, the better the understanding of a random process. In a first step, one can thus collect a huge amount of measured DW images in the data acquisition step to achieve the most accurate and precise estimation of the distribution of the Diffusion-Induced Molecular Displacements (DIMD) [Wedeen 2000, Wedeen 2005]. In a second step, one can propose shorter data acquisition steps that lead to similar performances in the diffusion modeling step. For example, [Tuch 1999, Tuch 2002] advice to collect the measured DW images using a few b-values but an important number of DSG directions, which is known as High Angular Resolution Diffusion Imaging (HARDI). [Scherrer 2012] propose the CUbe and SPhere (CUSP) acquisition scheme that yields a collection of measured DW images with multiple b-values at the cost of a single b-value acquisition. Future data acquisition protocols in dMRI could thus be modified to account for these proposals.
- The *theory-based* strategy: this is based on the idea that we often collected measured DW images using a single b-values and few DSG directions, which is referred to as **low angular resolution** dMRI. With this data, "can we do better than a low resolution reconstruction of WM fascicles?" Moving towards this direction, [Merlet 2010, Merlet 2011, Merlet 2012, Deriche 2013] are investigating the use of compressed sensing and [Merlet 2013] propose an interesting dictionary learning approach.

The strategy we adopted in this work is theory-based but quite different with respect to compressed sensing or dictionary learning methods. We focused on proposing a new statistical distribution for characterizing the DIMD. In addition, we also observe that current tractography algorithms do not optimally exploit multiple fascicle orientations identified at the diffusion modeling step. We thus design tractography algorithms in this purpose. We summarize our contributions in section 6.1 and we give a number of perspectives around this work in section 6.2.

6.1 Summary of the contributions

The contributions of this thesis are three-fold: (i) we propose a new Multi-Compartment Model (MCM) to study, from low angular resolution dMRI, the diffusion at the sub-voxel scale (summarized in section 6.1.1), (ii) we provide a means to overcome the model selection issue posed by the use of MCMs, which, in dMRI, boils down to selecting the optimal number of fascicle compartments to be included into the MCM

(summarized in section 6.1.2) and (iii) we successively design two tractography algorithms to efficiently use multiple fascicle orientations when provided by the MCM (summarized in section 6.1.3). All our diffusion tools have been implemented in C++ and the binaries will be available online soon. All the screenshots in the different chapters have been made with the MedINRIA multi-platform medical image processing and visualization software.

6.1.1 Contributions on diffusion modeling

The simplest way of studying the distribution of the DIMD is to assume that it is a centered 3-dimensional Gaussian distribution: this is the basis of Diffusion Tensor Imaging (DTI). When the voxelwise Probability Density Function (PDF) of the distribution of the DIMD, termed the Ensemble Average Propagator (EAP), is a single Gaussian PDF, the Signal Intensity Decay (SID), given by eqs. (1.9) and (1.10), is predicted as monoexponential (eq. (2.1)). However, non monoexponential measured DW signal decays have been observed in the literature that fault the Gaussian assumption. Efforts have been put to understand the causes of such a non monoexponential behavior. At low b-values, it is believed to be mainly due to the presence of multiple non-collinear fascicles and CSF contamination whereas, as the b-value increases, the effects of restricted diffusion in geometrically constrained environments (e.g., the axons) become non negligible and are also believed to be responsible for the non monoexponential measured DW signal decays.

As a result, at low b-values, MCMs based on exponential PDFs have been proposed to explain the non monoexponential behavior. Particularly efficient MCMs that rely on geometrical constraints for the fascicles are the ball-and-stick model [Behrens 2003, Behrens 2007] (in which fascicles are represented as cylinders of zero radius) or the ball-and-zeppelin model [Hosey 2005] (in which fascicles are represented as cylinders with non-zero radius but not explicitly included in the model). At high b-values, MCMs based on exponential PDFs for the extra-axonal space and non monoexponential PDFs for the intra-axonal space (taken from [Neuman 1974] or [**?**] who modeled restricted diffusion within a cylinder), without exchange between the two spaces, were proposed. This is known as Composite Hindered And Restricted ModEl of Diffusion (CHARMED), which was originally proposed in [Assaf 2004] modeling restricted diffusion within axons via Neuman's model. The assumptions behind Neuman's model are often not satisfied in dMRI, which led [Assaf 2005a] to modify the CHARMED by modeling restricted diffusion within axons via van Gelderen's model. In these two versions, the radius of the axons is explicitly included in the model but fixed. [Alexander 2008] propose another version of the CHARMED that allows one to estimate this radius. Subsequently, [Alexander 2010] fixed the diffusivities of the different compartments and estimated both intra-axonal and CSF occupancies in the voxel: this is known as the Minimal Model of White Matter Diffusion (MMWMD). Finally, recently, [Zhang 2012] introduced Neurite Orientation Dispersion and Density Imaging (NODDI), which is an MCM (currently with a single fascicle compartment) in which the non monoexponential measured DW signal

decay is modeled as a spherical deconvolution of a monoexponential SID with a Watson distribution to account for the fascicle orientation dispersion. Table 2.1 summarizes all these MCMs.

Our contributions on diffusion modeling are based on two observations: (i) previously proposed MCMs that account for intra- and extra-axonal spaces all require multiple b-value acquisitions whereas the long held tradition in DTI provided the community with tons of single b-value clinical data and (ii) the b-value threshold above which restricted diffusion is non negligible is unclear and, in any case, rather arbitrary. In section 2.3.3, we thus introduced a novel 3-dimensional statistical distribution whose PDF is based on the convolution between a 2-dimensional Von-Mises & Fisher (vMF) PDF and a 3-dimensional Gaussian PDF and is given by eq. (2.31). It yields an anisotropic distribution of the DIMD within a fascicle that is described by 5 parameters: the polar angle $\theta \in [0, \pi]$ of the fascicle orientation, the azimuth angle $\phi \in [0, 2\pi]$ of the fascicle orientation, the concentration $\kappa \geq 0$ of water molecules along the fascicle orientation (measure of orientation dispersion), the free diffusivity $d > 0$ along the fascicle orientation and the extra-axonal occupancy $\nu \in [0, 1]$ in the fascicle. We encompass this PDF within an MCM with m fascicle compartments and an additional compartment to account for CSF contamination, which is assumed to be well represented by an isotropic centered Gaussian PDF (Section 2.3.4.1). We coin this modeling Diffusion Directions Imaging (DDI), of which we propose a motivated parametrization (Section 2.3.4.3) that allows us to estimate it from single b-value clinical acquisitions. We show that the DDI MCM yields an analytic expression of the SID (Section 2.3.4.4), which (i) is asymptotically monoexponential as the b-value decreases and (ii) can be seen as a ball-and-stick or ball-and-zeppelin MCM under some constraints on the parameters. We evaluated the performances of the proposed constrained DDI MCM in comparison to the other MCMs that can be estimated with single b-value acquisitions, namely the ball-and-stick and ball-and-zeppelin model. We added the MMWMD to the comparison for evaluation of the estimated extra-axonal occupancy. Results (see section 2.4) show that our constrained DDI MCM most of the time outperforms the other models in terms of angular error on the estimated fascicle orientations (which decreases as the b-value, the number of DSG directions and the Signal-to-Noise Ratio (SNR) increase) and outputs unbiased estimates of both the CSF and extra-axonal occupancies in the voxel. The variance of the estimated extra-axonal occupancy is influenced neither by the b-value nor by the number of DSG directions but is greatly reduced at very high SNRs. The MMWMD provides a biased and highly variable estimate of the extra-axonal occupancy except at very high SNRs, b-value and number of DSG directions.

In conclusion, the constrained DDI MCM can be estimated from low angular (and spatial) resolution dMRI. It provides very accurate estimated fascicle orientations together with unbiased estimates of the CSF and extra-axonal occupancies in the voxel. With this modeling, we showed that all these parameters can be extracted correctly even at $b = 1000$ s/mm^2 as far as the SNR is very high.

Publications related to this model are [1] (technical report), [2, 5] (patents),

[6, 7] (peer-reviewed international conferences) and [11] (peer-reviewed journal).

6.1.2 Contributions on model selection

Geometry-based MCMs are adapted to model the EAP because the latter is the contribution of diffusion processes arising from different geometrically constrained environments (e.g., the axons). However, choosing the optimal number of fascicle compartments in a given voxel is a tricky question that may not even have a clear answer. This is a model selection problem. We reviewed in sections 3.1 and 3.2.1 the different approaches that have been adopted so far in the dMRI community. We distinguished in particular two classes of methods. First, brute force approaches output a map with the "optimal" number of fascicles per voxel, where "optimal" can be intended in the sense of minimizing the fitting error [Alexander 2002, Kreher 2005, Scherrer 2012] or, for the most recent works, the generalization error [Scherrer 2013]. Alternatives are averaging approaches. To the best of our knowledge, only the Automatic Relevance Detection (ARD) technique has been employed in the estimation of the ball-and-stick MCM [Behrens 2007]. It requires the estimation of the ball-and-stick MCM with the maximum assumed number of fascicles and performs an implicit model averaging by putting appropriate priors on the fascicle occupancies that shrink them towards 0 with very low variance if not supported by the data.

We proposed instead a solution that consists in explicitly averaging different MCMs (Section 3.2.2). The approach consists in estimating a given geometry-based MCM with $0, 1, \ldots, M$ fascicles and computing the probability of each of them to be the best model to fit the data. At the cost of running time, we obtain a set of weights, termed the *Akaike's weights*, that quantify these probabilities relatively to the size of the candidate set. This approach offers much more flexibility with respect to ARD. Indeed, not only we can quantify the probability of each model but we can derive probabilities that a subset of these models are the "best model" together to fit the data. We compared the ARD approach to ours using the same ball-and-stick MCM to ensure that differences can be interpreted in terms of differences between averaging methods and not between diffusion models (Section 3.2.3). We show that ARD identifies few 3-fascicle regions that are not statistically significant whereas our method robustly identify 3-fascicle regions where expected by previous histological studies (figs. 3.2 and 3.3). We then compared the ball-and-stick MCM and our proposed constrained DDI MCM, both averaged using our approach up to $M = 3$, showing that the DDI MCM further improves the detection and reconstruction of 3-fascicle regions (Figure 3.5).

We finally performed the estimation of our constrained DDI MCM with further model averaging up to $M = 3$ fascicle compartments on both a healthy subject and a patient suffering from a tumor (Section 3.3.2). The diffusion data were acquired with a single b-value of 1000 s/mm^2 and 30 DSG directions for the healthy subject and 20 DSG directions for the patient. We showed that we can provide corrected maps of the usual scalar metrics used by clinicians (e.g., Mean Diffusivity (MD) or Fractional Anisotropy (FA)). The DDI MCM additionally outputs the CSF and

extra-axonal occupancies and an orientation dispersion index in the voxel. These maps are also interested as we showed that the same drop in FA can be characterized by various configurations of these scalars in different part of the peritumoral region. These new scalar maps may thus be better specific biomarkers but this needs further evaluation.

Articles related to this model averaging method are in preparation.

6.1.3 Contributions on tractography

Diffusion modeling, and subsequent model selection/averaging approaches for the case of MCMs, is just one step towards the reconstruction of WM fascicles. The final stage consists in using the geometrical information gathered by the diffusion model to perform the actual reconstruction of the fascicles: this is known as tractography.

In chapter 4, we reviewed (non exhaustively) a great part of tractography algorithms available in the literature, of which we proposed the following classification:

- Deterministic streamline algorithms: these algorithms reconstruct a fascicle in a stepwise fashion from a seed voxel by following locally the orientations identified by the diffusion model;
- Energy-based algorithms: these algorithms reconstruct a globally optimal fascicle from a seed voxel to any other voxel by minimizing appropriate cost functions;
- Locally filtered streamline algorithms: these algorithms regularize streamline fascicles by perturbing the position or orientation at each step and by keeping the more consistent one;
- Probabilistic streamline algorithms: these algorihms generalize locally filtered streamline algorithms by keeping all perturbed positions or orientations but assigning weights to the corresponding fascicle.

We focused on two existing algorithms. The first one is one of the very first introduced tractography algorithm. It is the deterministic streamline algorithm called Fiber Assignement by Continuous Tracking (FACT) introduced by [Mori 1999]. This algorithm was developed to perform DTI tractography and, as such, supports only diffusion models that provide a single fascicle orientation per voxel. In section 4.1, we generalized it to account for multiple fascicle orientations per voxel. We applied it on clinical data with $b = 800$ s/mm^2 and 15 DSG directions using the DT model, the Q-ball model and the DDI MCM. The objectives were two-fold: (i) to evaluate the ability of the models to reconstruct the left and right Hand Motor Tracts (HMTs) and (ii) to study handedness asymmetry. The results showed that the DDI MCM is the only model (among the 3 evaluated) to robustly reconstruct the HMTs and it supports the idea of an asymmetry in the number of tracts between the left and right HMTs for right-handed individuals.

Our multi-modal FACT tractography algorithm however remains a deterministic algorihm. As such, it does not account for the uncertainty on the estimated fascicle orientations. Also, it suffers from a huge computational burden if one wants to explore all the fascicle orientations provided by the diffusion model. The second algorithm we focused on is thus a probabilistic streamline algorithm. We

chose to use the theory of particle filtering and thus started with the works of [Zhang 2009, Pontabry 2013] that introduced the use of particle filters to approximate the distribution of all possible tracts of a given length emanating from a single voxel. Traditional particle filters however have a hard time maintaining the multi-modal distribution of a filtering distribution. To circumvent this issue, we designed an adaptive multi-modal particle filter (Section 4.2). The only difference with the traditional one is that particles are grouped into clusters and resampling is performed within the clusters, which gives better chances to secondary leads. Weighting is thus provided at two levels, the particles and the clusters, with adequate updating rules for compliance with sequential importance sampling. One can see each cluster as a super-particle. The entire chapter 5 is dedicated to the application of this algorithm, coupled with our proposed constrained DDI MCM, for the reconstruction of the CST on patients suffering from tumors at various grades. The data was provided by the organizers of the DTI Tractography challenges on peritumoral WM anatomy for neurosurgical decision-making that were proposed for the last three years at the MICCAI conference. It consisted in low angular and spatial resolution measured DW images (b = 1000 s/mm^2 and 30 (years 1 and 2) and 20 (year 3) DSG directions). Results show that the current tools (constrained DDI MCM and multi-modal particle filter) enables to reconstruct the full extent of the CST away from the tumor and proposes a modified structure of the CST in the peritumoral regions that need further validation.

Publications related to this model are [3, 4, 8, 9, 10] (peer-reviewed international conferences).

6.2 Perspectives around this thesis

Clinical dMRI mainly consists in the application of a single low b-value with few DSG directions. These choices lead to low angular resolution measured DW images that, to date, were analyzed by DTI and subsequent tractography. These tools however have been shown to provide non-specific biomarkers (e.g., MD and FA) and misleading tractography results.

In this manuscript, we developed new tools for the analysis of such low angular resolution data. The objectives were two-fold: (i) to provide more specific biomarkers for brain pathologies and (ii) to achieve high resolution reconstruction of the WM fascicles. Throughout the experiments in the different chapters, we showed that the tools we proposed establish a solid basis towards these goals. Yet, a number of things can now be envisaged with these analysis tools regarding both the diffusion modelingand tractography steps. For instance:

- In diffusion modeling, we proposed the constrained DDI MCM for clinical use.

  – Influence of the b-value: We studied the influence of the b-value on the estimated parameters. However, the b-value depends on (i) the magnitude of the DSG, (ii) the diffusion time between two successively applied DSGs and (iii) the duration of a DSG, through eq. (1.12). Researchers often confound b-value

with magnitude of the DSG but [Scherrer 2012] pointed out that studying the influence of the diffusion time and DSG duration is very important.

– Estimation framework: In this manuscript, we always estimated our MCM by minimizing eq. (2.68), which is a simple least square fitting. This criterion relies on the assumption that the measurement noise is Gaussian, which, in the case of dMRI, is an approximation valid only at high SNRs. [Gudbjartsson 1995] showed that single-coil measured DW images are in fact corrupted with Rice-distributed noise and [Constantinides 1997] showed that multi-coil measured DW images are corrupted with Chi-distributed noise.

– Influence of the SNR: Our experiments led to the conclusion that the DDI MCM can provide valuable information even at $b = 1000$ s/mm^2 and few DSG directions (e.g., 30). The only thing that matters is to keep the SNR as high as possible. We thus advice repetitions of a standard clinical protocol rather than adding b-values and DSG directions if there is more time, in order to improve the SNR. Also, we recommend to reduce the TE as much as possible to achieve a high SNR. Finally, in all our experiments (except for simulated data of course), we filtered the measured DW images with the Rician-adapted non-local means filter introduced by [Wiest-Daesslé 2008]. Some more advanced denoising methods have been proposed recently in the literature [Aja-Fernández 2012, Rajan 2012] and, since achieving high SNR is crucial for an accurate estimation of the DDI MCM, may be employed as well.

– Similarity measure: If we want to use the DDI MCM to study pathologies (e.g., differences between a group of patients and a group of controls), we need to carefully define a similarity measure between two DDI MCMs to be able to register all the subjects on a common reference. Such similarity measures do exist for DTI [Fillard 2007] and multi-tensor models [Taquet 2012b] and need to be defined for DDI.

– Study of pathologies (retrospective and future): once defined a similarity measure between models, the DDI MCM offers the possibility to re-analyze former DTI datasets (see non exhaustive list proposed in section 1.4.5) and, more generally, to analyze any dMRI data. In particular, it would be interesting to further investigate the changes in CSF and extra-axonal occupancies and in orientation dispersion due to a given brain disorder. Indeed, both our experiments and other works (e.g., [Zhang 2012]) suggest that these parameters could be better biomarkers than MD and FA.

- In tractography, we proposed an adaptive multi-modal particle filter to approximate the distribution of all possible tracts of a given length emanating from a single voxel. We made a significant improvement by modifying the traditional particle filter to better maintain the multi-modality of the filtering distribution.

  – Weights of the tracts: Since the algorithm better maintain the multi-modality, it is able to reconstruct multiple tracts over extended volumes and to assign

them non negligible weights. At the moment, we did not study these weights in any way. These corrected weights may be a good starting point for connectiviy studies [Lazar 2010].

– Region Of Interests (ROIs): For all the tractography results presented in this manuscript, ROIs were delineated manually and sometimes yielded the partial reconstruction of other fascicles close to the CST. In these cases, we filtered these tracts. However, some researchers developed automated tools for ROI delineation [Shenton 1995] and atlases also exist to help in manual delineation (e.g., [Catani 2008, Mori 2008]), which could help us obtain even better tractography results.

– Interpolation: Currently, in the tractography process, when we reach a point between voxels, we interpolate by concatenating all the compartments of all the DDI MCMs of the surrounding voxels weighted according to their distance to the current point. This is sub-optimal since many of these compartments probably represent the same fascicles. [Taquet 2012a] introduced recently an interpolation scheme for multi-tensor models, which are a particular class of MCMs. The scheme strongly depends on the assumption that the diffusion in each fascicle is Gaussian so that its application to the DDI MCM is not straightforward. Yet, a similar approach could certainly improve the tractography results.

Appendix A

The distribution behind Diffusion Directions Imaging

A.1 Definitions from probability theory

Let us recall some probability theory basics that are required in Appendix A.2.

• Given two measurable spaces $(E_1, \mathcal{F}_1)$ and $(E_2, \mathcal{F}_2)$, a measurable function $h : E_1 \to E_2$ and a measure $\rho : \mathcal{F}_1 \to [0, +\infty]$, the **pushforward measure** $h \star \rho : \mathcal{F}_2 \to [0, +\infty]$ of ρ induced by h is defined as $(h \star \rho)(B) = \rho\left(h^{-1}(B)\right)$, for any $B \in \mathcal{F}_2$;

• a real p-dimensional **random variable** $\mathbf{x}$ is a measurable function from the probability space $(\Lambda, \mathcal{F}, P)$ to the measurable space $(\mathbb{R}^p, \mathcal{B}_p)$, where $\mathcal{B}_p$ is the Borel σ-algebra of $\mathbb{R}^p$;

• the **probability distribution**, often shorten to **distribution**, of the real p-dimensional random variable $\mathbf{x}$ is the pushforward measure of P induced by $\mathbf{x}$;

• the **Characteristic Function (CF)** $\varphi_{\mathbf{x}} : \mathbb{R}^p \to \mathbb{C}$ of the real p-dimensional random variable $\mathbf{x}$ reads

$$\begin{aligned}\varphi_{\mathbf{x}}(\boldsymbol{t}) &= \int_{\mathbb{R}^p} \exp\{i < \boldsymbol{t}, \boldsymbol{x} >\}\, d(\mathbf{x} \star P)(\boldsymbol{x}) = \int_{\mathbb{R}^p} \exp\{i < \boldsymbol{t}, \boldsymbol{x} >\}\, dP(\mathbf{x}^{-1}(\boldsymbol{x})) \\ &= \int_{\Lambda} \exp\{i < \boldsymbol{t}, \mathbf{x}(\omega) >\}\, \mathrm{d}P(\omega)\ ,\ \text{for any } \boldsymbol{t} \in \mathbb{R}^p\ .\end{aligned}$$

• the **Probability Density Function (PDF)** $p_{\mathbf{x}} : \mathbb{R}^p \to [0, +\infty]$ of the real p-dimensional random variable $\mathbf{x}$ is the Radon-Nikodym derivative of its distribution.

Note that the CF of the real p-dimensional random variable $\mathbf{x}$ always exists whereas its PDF exists if and only if the distribution of $\mathbf{x}$ is absolutely continuous with respect to the Lebesgue measure. In the latter case, the CF and the PDF are related by $\varphi_{\mathbf{x}}(\boldsymbol{t}) = \int_{\mathbb{R}^p} \exp\{i < \boldsymbol{t}, \boldsymbol{x} >\}\, p_{\mathbf{x}}(\boldsymbol{x}) d\boldsymbol{x}$, for any $\boldsymbol{t} \in \mathbb{R}^p$.

A.2 The von Mises & Fisher probability distribution

A.2.1 Definitions

It is a 4-parameter probability distribution. Here is the list of the notations used throughout this section:

• $A_2 = [0, \pi] \times [0, 2\pi[$ is the 2-dimensional space of spherical coordinates; equipped with its corresponding Borel σ-algebra $\mathcal{B}_{A_2}$, it is a measurable space;

• $(\theta_0, \phi_0, \kappa, R) \in A_2 \times [0, +\infty] \times]0, +\infty]$ are the parameters of the vMF distribution $F_{\theta_0, \phi_0, \kappa, R}$;

- Denoting $\boldsymbol{\mu} = (\sin\theta_0\cos\phi_0, \sin\theta_0\sin\phi_0, \cos\theta_0)'$, the vMF distribution can also be denoted by $F_{\boldsymbol{\mu},\kappa,R}$;
- $s_{\boldsymbol{\mu},\kappa} : \mathcal{B}_{A_2} \to [0, +\infty]$ is the vMF measure on the 2-dimensional unit sphere whose density with respect to the Lebesgue measure reads:

$$\begin{aligned} ds_{\boldsymbol{\mu},\kappa}(\theta,\phi) = \frac{\kappa\sin\theta}{4\pi\sinh\kappa}\exp\{\kappa\,(\mu_1\sin\theta\cos\phi \\ +\mu_2\sin\theta\sin\phi + \mu_3\cos\theta)\}\,d\theta d\phi\ , \end{aligned} \tag{A.1}$$

for any $(\theta,\phi) \in A_2$ so that $(A_2, \mathcal{B}_{A_2}, s_{\boldsymbol{\mu},\kappa})$ is a probability space;

- T_R is the real 3D random variable from $(A_2, \mathcal{B}_{A_2}, s_{\boldsymbol{\mu},\kappa})$ to $\left(\mathbb{R}^3, \mathcal{B}_3\right)$ such that

$$T_R(\theta,\phi) = (R\sin\theta\cos\phi, R\sin\theta\sin\phi, R\cos\theta)'\ ,\ \text{for any } (\theta,\phi) \in A_2\,. \tag{A.2}$$

The vMF distribution in $\mathbb{R}^3$ is the pushforward measure $F_{\boldsymbol{\mu},\kappa,R}$ of $s_{\boldsymbol{\mu},\kappa}$ induced by T_R; $\boldsymbol{\mu}$ (unit vector) is the *mean direction*, $\kappa \geq 0$ is the *concentration parameter* which controls the dispersion of the distribution around the mean direction $\boldsymbol{\mu}$ and $R > 0$ is the radius of the sphere on which the distribution has positive value.

A.2.2 Characteristic Function

Let $\mathbf{v}$ be a real 3D random variable following the vMF distribution $F_{\boldsymbol{\mu},\kappa,R}$. Then, for any $\boldsymbol{t} \in \mathbb{R}^3$, its CF is given by:

$$\varphi_{\mathbf{v}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \int_{A_2}\exp\{i < \boldsymbol{t}, T_R(\theta,\phi) >\}\,ds_{\boldsymbol{\mu},\kappa}(\theta,\phi)\,,$$

which, combined with eq. (A.1), becomes:

$$\varphi_{\mathbf{v}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \frac{\kappa}{4\pi\sinh\kappa}\int_{A_2}\exp\{< iR\boldsymbol{t} + \kappa\boldsymbol{\mu}, T_1(\theta,\phi) >\}\sin\theta d\theta d\phi\,. \tag{A.3}$$

The derivation of this integral is carried out in [Dégerine 1979] for the case $R = 1$. It is straightforward to generalize the expression for any $R > 0$:

$$\varphi_{\mathbf{v}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \frac{\kappa}{\sinh\kappa}\sum_{n=0}^{+\infty}\frac{z^n}{(2n+1)!}\,,\ \forall\boldsymbol{t} \in \mathbb{R}^3\,, \tag{A.4}$$

with $z = z(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \kappa^2 - R^2\|\boldsymbol{t}\|^2 + 2i\kappa R\boldsymbol{\mu}'\boldsymbol{t}$. Introducing

$$\begin{aligned} \alpha &= \alpha(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \sqrt{\frac{\text{Re}\ z + |z|}{2}}\,, \\ \beta &= \beta(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \frac{\text{Im}\ z}{\sqrt{2\,(\text{Re}\ z + |z|)}}\,, \\ \Omega &= \left\{\boldsymbol{t} \in \mathbb{R}^3 \text{ s.t. } \|\boldsymbol{t}\| \geq \kappa/R \text{ and } \boldsymbol{t} \perp \boldsymbol{\mu}\right\}\,, \end{aligned} \tag{A.5}$$

in which α and β are well-defined only when $\text{Re}\,(z) + |z| > 0$, eq. (A.4) simplifies to:

$$\varphi_{\mathbf{v}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,R) = \begin{cases} \dfrac{\kappa}{\sinh\kappa}\dfrac{\sin\sqrt{R^2\|\boldsymbol{t}\|^2 - \kappa^2}}{\sqrt{R^2\|\boldsymbol{t}\|^2 - \kappa^2}}\ ,\ \boldsymbol{t} \in \Omega\ , \\ \dfrac{\kappa}{\sinh\kappa}\dfrac{\sinh(\alpha + i\beta)}{\alpha + i\beta}\ ,\ \boldsymbol{t} \notin \Omega\ . \end{cases} \tag{A.6}$$

When $\kappa = 0$, the vMF distribution coincides with the uniform distribution on the 2-dimensional sphere of radius R. Besides, the vMF distribution is not absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^3$ as, for example, A_2 has null Lebesgue measure while $F_{\boldsymbol{\mu},\kappa,R}(A_2) = 1$. Therefore, it does not admit a PDF on $\mathbb{R}^3$.

A.3 Sum of independent vMF & 3D Gaussian variables

A.3.1 Definitions

Let $\mathbf{w} = \mathbf{v} + \mathbf{z}$ be the sum of two independent random variables, where:

- $\mathbf{v}$ follows a **vMF** distribution on the sphere of radius $\sqrt{\nu d}$ ($\nu \in [0,1]$, $d > 0$) with mean direction $+\boldsymbol{\mu}$ ($\|\boldsymbol{\mu}\| = 1$) and concentration parameter $\kappa \geq 0$. It admits a PDF on the 2-sphere of radius $\sqrt{\nu d}$. Its expression is given in [Jupp 1989] for $\nu d = 1$ and can easily be extended on the sphere of radius $\sqrt{\nu d} > 0$ by affine transformation:

$$p_{\mathbf{v}}(\boldsymbol{v}; +\boldsymbol{\mu}, \kappa, d, \nu) = \frac{1}{(\nu d)^{3/2}} \frac{\kappa}{4\pi \sinh \kappa} \exp\left\{ \frac{\kappa}{\sqrt{\nu d}} \boldsymbol{\mu}' \boldsymbol{v} \right\}, \tag{A.7}$$

for any $\boldsymbol{v} \in \mathbb{R}^3$ such that $\|\boldsymbol{v}\|^2 = \nu d$.

- $\mathbf{z}$ follows a zero-mean 3D Gaussian distribution with covariance matrix $\Sigma = \frac{(1-\nu)d}{\kappa+1}(I_3 + \kappa \boldsymbol{\mu}\boldsymbol{\mu}')$. It admits a PDF on $\mathbb{R}^3$. Using the Sherman-Morrison-Woodbury identity [Hager 1989] to invert D, the latter reads:

$$p_{\mathbf{z}}(\boldsymbol{z}; \pm\boldsymbol{\mu}, \kappa, d, \nu) = \frac{\kappa + 1}{(2\pi(1-\nu)d)^{3/2}} \exp\left\{ -\frac{(\kappa+1)\|\boldsymbol{z}\|^2 - \kappa(\boldsymbol{\mu}'\boldsymbol{z})^2}{2(1-\nu)d} \right\}, \tag{A.8}$$

for any $\boldsymbol{z} \in \mathbb{R}^3$.

- $\mathbf{v}$ and $\mathbf{z}$ are statistically independent.

A.3.2 Characteristic Function

The statistical independence of the random variables $\mathbf{v}$ and $\mathbf{z}$ implies that the CF of the random variable $\mathbf{w}$ reads:

$$\varphi_{\mathbf{w}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, d, \nu) = \varphi_{\mathbf{v}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, \sqrt{\nu d}) \varphi_{\mathbf{z}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, d, \nu), \forall \boldsymbol{t} \in \mathbb{R}^3,$$

where $\varphi_{\mathbf{v}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, \sqrt{\nu d})$ is given by eq. (A.6) with $R = \sqrt{\nu d}$ and $\varphi_{\mathbf{z}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, d, \nu)$ is the CF of the distribution of the random variable $\mathbf{z}$, which is given by [Anderson 2003]:

$$\varphi_{\mathbf{z}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, d, \nu) = \exp\left\{ -\frac{(1-\nu)d}{2(\kappa+1)} \left(\|\boldsymbol{t}\|^2 + \kappa(< \boldsymbol{\mu}, \boldsymbol{t} >^2) \right) \right\}.$$

The CF of the random variable $\mathbf{w}$ has thus the following expression:

$$\begin{aligned} \varphi_{\mathbf{w}}(\boldsymbol{t}; \boldsymbol{\mu}, \kappa, d, \nu) = & \exp\left\{ -\frac{(1-\nu)d}{2(\kappa+1)} \left(\|\boldsymbol{t}\|^2 + \kappa < \boldsymbol{\mu}, \boldsymbol{t} >^2 \right) \right\} \\ & \times \frac{\kappa}{\sinh \kappa} \begin{cases} \operatorname{sinc}\left(\sqrt{\nu d \|\boldsymbol{t}\|^2 - \kappa^2} \right), & \text{if } \nu d \|\boldsymbol{t}\|^2 \geq \kappa^2 \text{ and } \boldsymbol{t} \perp \boldsymbol{\mu}, \\ \dfrac{\sinh(\alpha + i\beta)}{\alpha + i\beta}, & \text{otherwise}, \end{cases} \end{aligned} \tag{A.9}$$

where α and β are given by eq. (A.5) with $R = \sqrt{\nu d}$:

$$\begin{aligned}
\alpha &= \sqrt{\frac{\mathrm{Re}\,(z) + |z|}{2}}\,, \\
\beta &= \frac{\mathrm{Im}\,(z)}{\sqrt{2\,(\mathrm{Re}\,(z) + |z|)}}\,, \\
z &= \kappa^2 - \nu d\|\boldsymbol{t}\|^2 + 2i\kappa\sqrt{\nu d} < \boldsymbol{\mu}, \boldsymbol{t} > \,.
\end{aligned}$$

A.3.3 Probability Density Function

Theorem 1. *The random variable* $\mathbf{w}$ *admits a PDF on* $\mathbb{R}^3$*, which is given by:*

$$\begin{aligned}
p_{\mathbf{w}}(\boldsymbol{w}; \boldsymbol{\mu}, \kappa, d, \nu) &= C(\kappa, d, \nu) \exp\left\{-\frac{(\kappa+1)w_\perp^2 + w_{/\!/}^2}{2(1-\nu)d}\right\} \\
&\times \int_{-1}^{1} \exp\left\{\frac{r_\nu \kappa}{2} t^2 + \left(\kappa + \frac{r_\nu w_{/\!/}}{\sqrt{\nu d}}\right) t\right\} I_0\left(\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}} w_\perp \sqrt{1-t^2}\right) dt\,,
\end{aligned}$$

for any $\boldsymbol{w} \in \mathbb{R}^3$*, where:*

$$\begin{aligned}
r_\nu &:= \frac{\nu}{1-\nu}\,, \\
C(\kappa, d, \nu) &:= \frac{\kappa(\kappa+1)\exp\left\{-\frac{r_\nu(\kappa+1)}{2}\right\}}{2\,(2\pi(1-\nu)d)^{3/2}\sinh\kappa}\,, \\
(w_{/\!/}, w_\perp) &:= \left(< \boldsymbol{\mu}, \boldsymbol{w} >, \sqrt{\|\boldsymbol{w}\|^2 - < \boldsymbol{\mu}, \boldsymbol{w} >^2}\right)\,,
\end{aligned}$$

and I_0 *is the* 0*-th order modified Bessel function [Abramowitz 1972], with the convention that* $< \boldsymbol{\mu}, \boldsymbol{w} >= \|\boldsymbol{w}\|$ *for any* $\boldsymbol{w} \in \mathbb{R}^3$ *when* $\kappa = 0$*.*

Proof. Let $F_{\boldsymbol{\mu},\kappa,d,\nu}$, $G_{\boldsymbol{\mu},\kappa,d,\nu}$ and $H_{\boldsymbol{\mu},\kappa,d,\nu}$ be the probability distributions of the real 3-dimensional random variables $\mathbf{v}$, $\mathbf{z}$ and $\mathbf{w}$ respectively. The density of $H_{\boldsymbol{\mu},\kappa,d,\nu}$ with respect to $G_{\boldsymbol{\mu},\kappa,d,\nu}$ reads

$$dH_{\boldsymbol{\mu},\kappa,d,\nu}(\boldsymbol{w}) = \int_{\mathbb{R}^3} dG_{\boldsymbol{\mu},\kappa,d,\nu}(\boldsymbol{w} - \boldsymbol{v}) dF_{\boldsymbol{\mu},\kappa,d,\nu}(\boldsymbol{v})\,, \forall \boldsymbol{w} \in \mathbb{R}^3\,.$$

Now, $G_{\boldsymbol{\mu},\kappa,d,\nu}$ is absolutely continuous with respect to the Lebesgue measure, so does $H_{\boldsymbol{\mu},\kappa,d,\nu}$. The PDF of $\mathbf{w}$ thereby exists and is given by:

$$p_{\mathbf{w}}(\boldsymbol{w}; \boldsymbol{\mu}, \kappa, d, \nu) = \int_{\mathbb{R}^3} p_{\mathbf{z}}(\boldsymbol{w} - \boldsymbol{v}; \boldsymbol{\mu}, \kappa, d, \nu) dF_{\boldsymbol{\mu},\kappa,d,\nu}(\boldsymbol{v})\,, \forall \boldsymbol{w} \in \mathbb{R}^3\,. \tag{A.10}$$

Since $F_{\boldsymbol{\mu},\kappa,d,\nu}$ is a vMF distribution, it is the pushforward measure of $\nu_{\boldsymbol{\mu},\kappa}$ defined in eq. (A.1) induced by the random variable $T_{\sqrt{\nu d}}$ defined as in eq. (A.2), with $R = \sqrt{\nu d}$. We thus have:

$$dF_{\boldsymbol{\mu},\kappa,d,\nu}(\boldsymbol{v}) = ds_{\boldsymbol{\mu},\kappa}\left(T^{-1}_{\sqrt{\nu d}}(\boldsymbol{v})\right)\,, \forall \boldsymbol{v} \in \mathbb{R}^3\,. \tag{A.11}$$

Inserting eqs. (A.1) and (A.11) into eq. (A.10) leads to:

$$p_{\mathbf{w}}(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu) = \frac{\kappa}{4\pi\sinh\kappa}\int_{A_2} p_{\mathbf{z}}(\boldsymbol{w}-\sqrt{\nu d}T_1(\theta,\phi);\boldsymbol{\mu},\kappa,d,\nu) \times \exp\{\kappa < \boldsymbol{\mu}, T_1(\theta,\phi) >\}\sin\theta \mathrm{d}\theta\mathrm{d}\phi\,, \forall \boldsymbol{w}\in\mathbb{R}^3\,. \tag{A.12}$$

The expression of $p_{\mathbf{z}}$ given by eq. (A.8) simplifies eq. (A.12) to:

$$p_{\mathbf{w}}(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu) = \frac{C(\kappa,d,\nu)}{2\pi}\exp\left\{-\frac{(\kappa+1)w_\perp^2 + w_{/\!/}^2}{2(1-\nu)d}\right\} Q(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu)\,, \tag{A.13}$$

where Q is defined as follows:

$$Q(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu) := \int_{A_2}\exp\left\{\frac{r_\nu\kappa}{2} < \boldsymbol{\mu}, T_1(\theta,\phi) >^2 + \kappa\left(1-\frac{r_\nu < \boldsymbol{\mu},\boldsymbol{w} >}{\sqrt{\nu d}}\right) < \boldsymbol{\mu}, T_1(\theta,\phi) > + \frac{r_\nu(\kappa+1)}{\sqrt{\nu d}} < \boldsymbol{w}, T_1(\theta,\phi) >\right\}\sin\theta\mathrm{d}\theta\mathrm{d}\phi\,.$$

Let P be the unitary matrix that rotates $\boldsymbol{\mu}$ to $\boldsymbol{e}_3 = (0,0,1)'$, i.e., $P\boldsymbol{\mu} = \boldsymbol{e}_3$. Defining $\boldsymbol{w}^\star = P\boldsymbol{w}$, we have

$$\begin{aligned} Q(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu) &= \int_{A_2}\exp\left\{\frac{r_\nu\kappa}{2}\cos^2\theta + \kappa\left(1-\frac{r_\nu w_3^\star}{\sqrt{\nu d}}\right)\cos\theta\right\} \\ &\times\exp\left\{\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}\left(w_1^\star\sin\theta\cos\phi + w_2^\star\sin\theta\sin\phi + w_3^\star\cos\theta\right)\right\}\sin\theta\mathrm{d}\theta\mathrm{d}\phi\,, \\ &= \int_0^\pi \exp\left\{\frac{r_\nu\kappa}{2}\cos^2\theta + \left(\kappa+\frac{r_\nu w_3^\star}{\sqrt{\nu d}}\right)\cos\theta\right\} h(\boldsymbol{w},\theta)\sin\theta\mathrm{d}\theta\,, \end{aligned} \tag{A.14}$$

where $h(\boldsymbol{w},\theta) := \int_{\phi_0}^{2\pi+\phi_0}\exp\left\{\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}\sin\theta\sqrt{\left(w_1^\star\right)^2+\left(w_2^\star\right)^2}\sin\phi\right\}\mathrm{d}\phi$ and $\phi_0 = \arctan\frac{w_1^\star}{w_2^\star} + \pi(w_2^\star \le 0)$. Because sin is a 2π-periodic function and thanks to the relation $\int_0^{2\pi} e^{z\sin x}\mathrm{d}x = 2\pi I_0(z)$, for any $z\in\mathbb{R}$, which can be deduced from [Abramowitz 1972, p. 376], we can simplify $h(\boldsymbol{w},\theta)$ to:

$$h(\boldsymbol{w},\theta) = 2\pi I_0\left(\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}\sin\theta\sqrt{\left(w_1^\star\right)^2+\left(w_2^\star\right)^2}\right)\,. \tag{A.15}$$

Substituting eq. (A.15) into eq. (A.14) and observing that $w_3^\star = < \boldsymbol{e}_3, \boldsymbol{w}^\star > = < \boldsymbol{e}_3, P\boldsymbol{w} > = < \boldsymbol{\mu}, \boldsymbol{w} > = w_{/\!/}$ and $\left(w_1^\star\right)^2 + \left(w_2^\star\right)^2 = \|\boldsymbol{w}^\star\|^2 - \left(w_3^\star\right)^2 = \|\boldsymbol{w}\|^2 - < \boldsymbol{\mu},\boldsymbol{w} >^2 = w_\perp^2$ yields the following single-integral representation of Q:

$$\begin{aligned} Q(\boldsymbol{w};\boldsymbol{\mu},\kappa,d,\nu) = 2\pi\int_{-1}^1 &\exp\left\{\frac{r_\nu\kappa}{2}t^2 + \left(\kappa+\frac{r_\nu w_{/\!/}}{\sqrt{\nu d}}\right)t\right\} \\ &\times I_0\left(\frac{r_\nu(\kappa+1)}{\sqrt{\nu d}}w_\perp\sqrt{1-t^2}\right)\mathrm{d}t\,. \end{aligned} \tag{A.16}$$

Inserting eq. (A.16) into eq. (A.13) ends the proof. □

A.3.4 Projections parallel and perpendicular to the mean direction

For any $t \in [-1, 1]$, let us define the following two random variables:

- $\mathrm{n}_t \sim \mathrm{N}_1\left(\sqrt{\nu d}t, (1-\nu)d\right)$ is a univariate Gaussian random variable with mean $\sqrt{\nu d}t$ and variance $(1-\nu)d$,
- $\mathrm{r}_t \sim \mathrm{R}\left(\sqrt{\nu d(1-t^2)}, \sqrt{\frac{(1-\nu)d}{\kappa+1}}\right)$ is a univariate Rician random variable with mean $\sqrt{\nu d(1-t^2)}$ and scale $\sqrt{\frac{(1-\nu)d}{\kappa+1}}$.

Manipulating the expression of the PDF of the distribution of $\mathbf{w}$ given in theorem 1, one can obtain:

$$p_{\mathbf{w}}(\boldsymbol{w}; \boldsymbol{\mu}, \kappa, d, \nu) = \frac{1}{2\pi w_\perp} \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} \exp\{\kappa t\}\, p_{\mathrm{n}_t}(w_{/\!/}) p_{\mathrm{r}_t}(w_\perp) \mathrm{d}t\,, \forall \boldsymbol{w} \in \mathbb{R}^3\,.$$

Expressing $\boldsymbol{w}$ in cylindrical coordinates leads to:

$$p_{\mathbf{w}}(\boldsymbol{w}; \boldsymbol{\mu}, \kappa, d, \nu)\mathrm{d}\boldsymbol{w} = \frac{1}{2\pi} \frac{\kappa}{2\sinh\kappa} \left[\int_{-1}^{1} \exp\{\kappa t\}\, p_{\mathrm{n}_t}(w_{/\!/}) p_{\mathrm{r}_t}(w_\perp) \mathrm{d}t\right] \mathrm{d}w_\perp \mathrm{d}w_{/\!/} \mathrm{d}\beta\,,$$

for any $(w_\perp, w_{/\!/}, \beta) \in \mathbb{R}^+ \times \mathbb{R} \times [0, 2\pi]$. This allows us to state the following

Corollary 1. *The joint PDF of* $(\mathrm{w}_{/\!/}, \mathrm{w}_\perp)$ *is given by:*

$$p_{(\mathrm{w}_{/\!/}, \mathrm{w}_\perp)}(w_{/\!/}, w_\perp) = \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} \exp\{\kappa t\}\, p_{\mathrm{n}_t}(w_{/\!/}) p_{\mathrm{r}_t}(w_\perp)\, dt\,,$$

and, thus, the marginal PDFs of $\mathrm{w}_{/\!/}$ *and* $\mathrm{w}_\perp$ *are respectively given by:*

$$p_{\mathrm{w}_{/\!/}}(w_{/\!/}) = \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} \exp\{\kappa t\}\, p_{\mathrm{n}_t}(w_{/\!/})\, dt\,,$$
$$p_{\mathrm{w}_\perp}(w_\perp) = \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} \exp\{\kappa t\}\, p_{\mathrm{r}_t}(w_\perp)\, dt\,.$$

A.3.4.1 Raw moments of univariate Gaussian and Rice distributions

The Gaussian distribution. Let $\mathrm{n} \sim \mathrm{N}_1(\mu, \sigma^2)$ be a univariate Gaussian random variable with mean μ and variance σ^2. The non-central moments of its distribution are then given, for any $p \geq 0$, by:

$$\mathrm{E}\left[\mathrm{n}^p\right] = \left(\frac{\sigma}{i\sqrt{2}}\right)^p H_p\left(\frac{i\mu}{\sigma\sqrt{2}}\right)\,,$$

where $H_p(\cdot)$ are Hermite's polynomials. In particular, the second and fourth order raw moments are given by:

$$\begin{aligned} \mathrm{E}\left[\mathrm{r}^2\right] &= \sigma^2 + \mu^2\,, \\ \mathrm{E}\left[\mathrm{r}^4\right] &= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4\,. \end{aligned} \tag{A.17}$$

The Rice distribution. Let $\mathrm{r} \sim \mathrm{R}(\nu, \sigma)$ be a univariate Rician random variable with location ν and scale σ. The non-central moments of its distribution are then given, for any $p \geq 0$, by:

$$\mathrm{E}\left[\mathrm{r}^p\right] = \sigma^p 2^{\frac{p}{2}} \Gamma\left(1 + \frac{p}{2}\right) L_{\frac{p}{2}}\left(-\frac{\nu^2}{2\sigma^2}\right),$$

where $L_q(\cdot)$ are Laguerre functions. In particular, the second and fourth order raw moments are given by:

$$\begin{aligned} \mathrm{E}\left[\mathrm{r}^2\right] &= 2\sigma^2 + \nu^2, \\ \mathrm{E}\left[\mathrm{r}^4\right] &= 8\sigma^2 + 8\sigma^2\nu^2 + \nu^4. \end{aligned} \tag{A.18}$$

A.3.4.2 Raw moments of the distribution of $\mathrm{w}_{/\!/}$ and $\mathrm{w}_{\perp}$

Corollary 1 allows one to easily compute all the raw moments of the distributions of the random variables $\mathrm{w}_{/\!/}$ and $\mathrm{w}_{\perp}$, knowing those of univariate Gaussian and Rician distributions. Indeed, for any $p \in \mathbb{N}^\star$, we have:

$$\begin{aligned} \mathrm{E}\left[\mathrm{w}_{/\!/}^p\right] &= \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} \exp\{\kappa t\}\, \mathrm{E}\left[\mathrm{n}_t^p\right] \mathrm{d}t, \\ \mathrm{E}\left[\mathrm{w}_{\perp}^p\right] &= \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} \exp\{\kappa t\}\, \mathrm{E}\left[\mathrm{r}_t^p\right] \mathrm{d}t. \end{aligned}$$

The odd order raw moments are of little interest since these distributions are meant to be encompassed within an antipodally symmetric distribution for which the odd raw moments are zero. On the other hand, the second and fourth order raw moments are of particular interest since they are related to the variance and kurtosis properties of the distribution. Since $\mathrm{n}_t \sim \mathrm{N}_1\left(\sqrt{\nu d}t, (1-\nu)d\right)$ and $\mathrm{r}_t \sim \mathrm{R}\left(\sqrt{\nu d(1-t^2)}, \sqrt{\frac{(1-\nu)d}{\kappa+1}}\right)$, we have:

$$\begin{aligned} \mathrm{E}\left[\mathrm{n}_t^2\right] &= \nu d\left(\frac{1}{r_\nu} + t^2\right), \\ \mathrm{E}\left[\mathrm{n}_t^4\right] &= (\nu d)^2\left(\frac{3}{r_\nu^2} + \frac{6t^2}{r_\nu} + t^4\right), \\ \mathrm{E}\left[\mathrm{r}_t^2\right] &= (1-\nu)d\left(r_\nu + \frac{2}{\kappa+1} - r_\nu t^2\right), \\ \mathrm{E}\left[\mathrm{r}_t^4\right] &= (1-\nu)^2 d^2\left[r_\nu^2 + \frac{8r_\nu}{\kappa+1} + \frac{8}{(\kappa+1)^2} - 2r_\nu\left(r_\nu + \frac{4}{\kappa+1}\right)t^2 + r_\nu^2 t^4\right]. \end{aligned}$$

Defining, for any $p \geq 0$, the following integral:

$$B_p(\kappa) \coloneqq \frac{\kappa}{2\sinh\kappa} \int_{-1}^{1} t^p \exp\{\kappa t\}\, \mathrm{d}t, \tag{A.19}$$

of which we can trivially compute $B_0(\kappa) = 1$, we can express the second and fourth

raw moments of the distributions of the random variables $\mathrm{w}_{/\!/}$ and $\mathrm{w}_{\perp}$ as follows:

$$\mathrm{E}\left[\mathrm{w}_{/\!/}^2\right] = \nu d\left(\frac{1}{r_\nu} + B_2(\kappa)\right),$$
$$\mathrm{E}\left[\mathrm{w}_{/\!/}^4\right] = (\nu d)^2\left(\frac{3}{r_\nu^2} + \frac{6}{r_\nu}B_2(\kappa) + B_4(\kappa)\right),$$
$$\mathrm{E}\left[\mathrm{w}_{\perp}^2\right] = (1-\nu)d\left(r_\nu + \frac{2}{\kappa+1} - r_\nu B_2(\kappa)\right),$$
$$\mathrm{E}\left[\mathrm{w}_{\perp}^4\right] = (1-\nu)^2 d^2\left[r_\nu^2 + \frac{8r_\nu}{\kappa+1} + \frac{8}{(\kappa+1)^2} - 2r_\nu\left(r_\nu + \frac{4}{\kappa+1}\right)B_2(\kappa) + r_\nu^2 B_4(\kappa)\right].$$

Now, observe that eq. (A.19) can be rewritten as follows:

$$B_p(\kappa) = \mathrm{E}\left[\cos^p\theta\right],$$

where θ follows a vMF distribution as defined in [Mardia 1972]. In the same book, the expression of $B_p(\kappa)$ is given as λ_p (p.232, Eq.8.5.24) and reads:

$$B_p(\kappa) = \frac{I_{p+\frac{1}{2}}(\kappa)}{I_{\frac{1}{2}}(\kappa)}, \tag{A.20}$$

where $I_q(\cdot)$ is the q-th order modified Bessel function of the first kind.

Appendix B

The signal predicted by Diffusion Directions Imaging

B.1 DDI-derived Signal Intensity Decay

Theorem 2. *Consider a fascicle represented as a pseudo-cylinder characterized inter-alia by its orientation $\pm\boldsymbol{\mu}$. Assume that the DIMD $\mathbf{y}$ within this fascicle follow a probability distribution whose PDF is an equally weighted mixture of two PDFs as given by theorem 1, one in the direction $+\boldsymbol{\mu}$ and the other in the opposite direction $-\boldsymbol{\mu}$, the remaining parameters κ, d, ν being the same for the two PDFs.*

Under this assumption, the expression of the SID induced by the application of a DSG with b-value b_k and direction $\boldsymbol{g}_j$ is given by:

$$A_{jk} = \exp\left\{-\frac{b_k(1-\nu)d}{\kappa+1}\left(1+\kappa<\boldsymbol{\mu},\boldsymbol{g}_j>^2\right)\right\} \times \frac{\kappa}{\sinh\kappa}\begin{cases} \left|sinc\left(\sqrt{2b_k\nu d-\kappa^2}\right)\right| & , \text{ if } 2b_k\nu d \geq \kappa^2 \text{ and } \boldsymbol{g}_j \perp \boldsymbol{\mu}\,, \\ \dfrac{|\alpha\sinh\alpha\cos\beta+\beta\cosh\alpha\sin\beta|}{\alpha^2+\beta^2} & , \text{ otherwise,} \end{cases}$$

where

$$\alpha = \sqrt{\frac{\mathrm{Re}\,(z)+|z|}{2}} \text{ and } \beta = \frac{\mathrm{Im}\,(z)}{\sqrt{2(\mathrm{Re}\,(z)+|z|)}}\,,$$

with $z = \kappa^2 - 2b_k\nu d + 2i\kappa\sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j >$.

Proof. According to q-space theory, the expression of the SID in terms of the distribution of the DIMD is given by eqs. (1.9) and (1.10) and reads, for a given b-value b_k and a given DSG direction $\boldsymbol{g}_j$:

$$A_{jk} = \left|\varphi_{\mathbf{y}}\left(\sqrt{2b_k}\boldsymbol{g}_j\right)\right|\,, \tag{B.1}$$

where $\varphi_{\mathbf{y}}$ is the CF of the random variable $\mathbf{y}$.

By definition of the DIMD $\mathbf{y}$, the CF of their distribution is given by

$$\varphi_{\mathbf{y}}(\boldsymbol{t};\pm\boldsymbol{\mu},\kappa,d,\nu) = \frac{1}{2}\varphi_{\mathbf{w}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,d,\nu) + \frac{1}{2}\varphi_{\mathbf{w}}(\boldsymbol{t};-\boldsymbol{\mu},\kappa,d,\nu)\,,$$

where $\varphi_{\mathbf{w}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,d,\nu)$ is the CF of the distribution of a random variable $\mathbf{w}$ with PDF as stated in theorem 1 and is given by eq. (A.9). Moreover, it follows from eq. (A.5)

that $\alpha(\boldsymbol{t};-\boldsymbol{\mu},\kappa,R)=\alpha(\boldsymbol{t};\boldsymbol{\mu},\kappa,R)=\alpha$ and $\beta(\boldsymbol{t};-\boldsymbol{\mu},\kappa,R)=-\beta(\boldsymbol{t};\boldsymbol{\mu},\kappa,R)=-\beta$. We thus obtain:

$$\begin{aligned}\varphi_{\mathbf{y}}(\boldsymbol{t};\boldsymbol{\mu},\kappa,d,\nu) &= \exp\left\{-\frac{(1-\nu)d}{2(\kappa+1)}\left(\|\boldsymbol{t}\|^2+\kappa<\boldsymbol{\mu},\boldsymbol{t}>^2\right)\right\}\\ &\times\frac{\kappa}{\sinh\kappa}\begin{cases}\operatorname{sinc}\left(\sqrt{\nu d\|\boldsymbol{t}\|^2-\kappa^2}\right), & \text{if } \nu d\|\boldsymbol{t}\|^2\geq\kappa^2 \text{ and } \boldsymbol{t}\perp\boldsymbol{\mu},\\ \frac{1}{2}\left(\frac{\sinh(\alpha+i\beta)}{\alpha+i\beta}+\frac{\sinh(\alpha-i\beta)}{\alpha-i\beta}\right), & \text{otherwise},\end{cases}\end{aligned} \tag{B.2}$$

Finally, simple derivations yield:

$$\begin{aligned}\frac{1}{2}\left(\frac{\sinh(\alpha+i\beta)}{\alpha+i\beta}+\frac{\sinh(\alpha-i\beta)}{\alpha-i\beta}\right) &= \operatorname{Re}\left(\frac{\sinh(\alpha+i\beta)}{\alpha+i\beta}\right)\\ &= \frac{\alpha\sinh\alpha\cos\beta+\beta\cosh\alpha\sin\beta}{\alpha^2+\beta^2},\end{aligned}$$

and inserting eq. (B.2) into eq. (B.1) ends the proof. □

B.1.1 SID measured along particular directions

Along the orientation of the fascicle. When the DSG is applied in a direction $\boldsymbol{g}_j$ parallel to the orientation $\pm\boldsymbol{\mu}$ of the fascicle, α and β take a simpler form. Indeed, from $z=\kappa^2-2b_k\nu d+2i\kappa\sqrt{2b_k\nu d}<\boldsymbol{\mu},\boldsymbol{g}_j>$ with $\left|<\boldsymbol{\mu},\boldsymbol{g}_j>\right|=1$, it is trivial to see that $|z|=\kappa^2+2b_k\nu d$, which leads to:

$$\alpha=\kappa \text{ and } \beta=\sqrt{2b_k\nu d}\cdot\operatorname{sgn}<\boldsymbol{\mu},\boldsymbol{g}_j>,$$

and the expression of the SID given in theorem 2 reads:

$$A_{jk}=\exp\{-b_k(1-\nu)d\}\left|\cos\sqrt{2b_k\nu d}\right|\frac{\left|1+\sqrt{2b_k\nu d}\frac{\coth\kappa}{\kappa}\tan\sqrt{2b_k\nu d}\right|}{1+\frac{2b_k\nu d}{\kappa^2}}.$$

Perpendicularly to the orientation of the fascicle. When the DSG is applied in a direction $\boldsymbol{g}_j$ perpendicular to the orientation $\pm\boldsymbol{\mu}$ of the fascicle, α and β also take a simpler form. Indeed, from $z=\kappa^2-2b_k\nu d+2i\kappa\sqrt{2b_k\nu d}<\boldsymbol{\mu},\boldsymbol{g}_j>$ with $<\boldsymbol{\mu},\boldsymbol{g}_j>=0$, it is trivial to see that $|z|=\left|\kappa^2-2b_k\nu d\right|$.

As a consequence, for low b-values $b_k<\frac{\kappa^2}{2\nu d}$, we obtain $\alpha=\sqrt{\kappa^2-2b_k\nu d}$ and $\beta=0$, which yields the following simplified SID:

$$A_{jk}=\exp\left\{-\frac{b_k(1-\nu)d}{\kappa+1}\right\}\left(1-\frac{2b_k\nu d}{\kappa^2}\right)^{-1/2}\frac{\sinh\left(\kappa\sqrt{1-\frac{2b_k\nu d}{\kappa^2}}\right)}{\sinh\kappa}.$$

Differently, for high b-values $b_k\geq\frac{\kappa^2}{2\nu d}$, we use the first form given in theorem 2 which yields:

$$A_{jk}=\exp\left\{-\frac{b_k(1-\nu)d}{\kappa+1}\right\}\frac{\kappa}{\sinh\kappa}\left|\operatorname{sinc}\left(\sqrt{2b_k\nu d-\kappa^2}\right)\right|.$$

B.1.2 SID in lowly orientationally dispersed media

Low orientational dispersion is characterized by $\kappa \to \infty$. In order to understand how does the SID described by theorem 2 behave at such a regime, we can rearrange the expressions of α and β as follows. First, from $z = \kappa^2 - 2b_k\nu d + 2i\kappa\sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j >$, it is trivial to see that $|z|^2 \sim \kappa^4$ when κ is large. Thus, we get $\mathrm{Re}\,(z) + |z| \sim 2\kappa^2$ for large κ. As a consequence, we obtain the following limiting behaviors for α and β:

$$\alpha \xrightarrow{\kappa\to\infty} \kappa \text{ and } \beta \xrightarrow{\kappa\to\infty} \sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j >, \tag{B.3}$$

and the expression of the SID given in theorem 2 reads, for large κ:

$$A_{jk} \sim \exp\left\{-b_k(1-\nu)d < \boldsymbol{\mu}, \boldsymbol{g}_j >^2\right\} |\cos\beta| \frac{\left|1 + \beta\frac{\coth\kappa}{\kappa}\tan\beta\right|}{1 + \frac{\beta^2}{\kappa^2}}.$$

The expression can be further simplified to:

$$A_{jk} \xrightarrow{\kappa\to\infty} \exp\left\{-b_k(1-\nu)d < \boldsymbol{\mu}, \boldsymbol{g}_j >^2\right\} \left|\cos\left(\sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j >\right)\right|,$$

where β has been replaced by its limit given by eq. (B.3).

B.1.3 SID in highly orientationally dispersed media

High orientational dispersion is characterized by $\kappa \to 0$. In order to understand how does the SID described by theorem 2 behave at such a regime, we can rearrange the expressions of α and β as follows. First, from $z = \kappa^2 - 2b_k\nu d + 2i\kappa\sqrt{2b_k\nu d} < \boldsymbol{\mu}, \boldsymbol{g}_j >$, we can derive:

$$|z|^2 = (2b_k\nu d)^2 \left(1 + \left(2 < \boldsymbol{\mu}, \boldsymbol{g}_j >^2 - 1\right)\frac{\kappa^2}{b_k\nu d} + \left(\frac{\kappa^2}{2b_k\nu d}\right)^2\right),$$

which leads to the following approximations for $\kappa \to 0$:

$$|z| = 2b_k\nu d + \left(2 < \boldsymbol{\mu}, \boldsymbol{g}_j >^2 - 1\right)\kappa^2 + o(\kappa^2) \text{ and thus } \mathrm{Re}\,(z) + |z| = 2 < \boldsymbol{\mu}, \boldsymbol{g}_j >^2 \kappa^2 + o(\kappa^2).$$

We thus obtain the following limiting behaviors for α and β:

$$\alpha \xrightarrow{\kappa\to 0} 0 \text{ and } \beta \xrightarrow{\kappa\to 0} \mathrm{sgn} < \boldsymbol{\mu}, \boldsymbol{g}_j > \sqrt{2b_k\nu d}, \tag{B.4}$$

and the expression of the SID given in theorem 2 reads:

$$A_{jk} = \exp\left\{-b_k(1-\nu)d\right\} \begin{cases} \left|\mathrm{sinc}\left(\sqrt{2b_k\nu d}\right)\right| & , \text{ if } \boldsymbol{g}_j \perp \boldsymbol{\mu}, \\ \frac{|\beta| \cdot |\sin\beta|}{\beta^2} & , \text{ otherwise}. \end{cases}$$

Now observe that $\frac{|\beta|\cdot|\sin\beta|}{\beta^2} = |\mathrm{sinc}\,(\beta)| = \left|\mathrm{sinc}\left(\sqrt{2b_k\nu d}\right)\right|$ thanks to eq. (B.4), which yields:

$$A_{jk} \xrightarrow{\kappa\to 0} \exp\left\{-b_k(1-\nu)d\right\} \left|\mathrm{sinc}\left(\sqrt{2b_k\nu d}\right)\right|.$$

B.2 Mean Diffusivity

Let $\mathbf{y}$ be a zero-mean 3-dimensional random vector and $\boldsymbol{\mu}$ be a direction on the 2-sphere $\mathbb{S}^2$. First of all, observe that:

$$\forall \boldsymbol{c} \in \mathbb{R}^3 \bigcap \mathbb{S}^2 \,,\, \exists (a, \boldsymbol{c}_\perp) \in [-1,1] \times (\mathrm{span}\{\boldsymbol{\mu}\})^\perp \bigcap \mathbb{S}^2 : \boldsymbol{c} = a\boldsymbol{\mu} + \sqrt{1-a^2}\boldsymbol{c}_\perp \,.$$

If we decompose $\mathbf{y}$ as follows:

$$\mathbf{y} = \mathrm{y}_{/\!/}\boldsymbol{\mu} + \mathrm{y}_\perp \mathbf{u}\,, \text{ with } \mathbf{u} = \frac{\mathbf{y} - \mathrm{y}_{/\!/}\boldsymbol{\mu}}{\mathrm{y}_\perp} \in (\mathrm{span}\{\boldsymbol{\mu}\})^\perp \bigcap \mathbb{S}^2 \,,$$

where $\mathrm{y}_{/\!/} =< \boldsymbol{\mu}, \mathbf{y} >$ and $\mathrm{y}_\perp = \sqrt{\|\mathbf{y}\|^2 - < \boldsymbol{\mu}, \mathbf{y} >^2}$, we can compute the apparent diffusivity in a direction $\boldsymbol{c} \in \mathbb{R}^3 \bigcap \mathbb{S}^2$ as:

$$D(\boldsymbol{c}) = \mathrm{V}\left[< \boldsymbol{c}, \mathbf{y} >\right] = \mathrm{E}\left[< \boldsymbol{c}, \mathbf{y} >^2\right] = a^2 \mathrm{E}\left[\mathrm{y}_{/\!/}^2\right] + (1-a^2)\mathrm{E}\left[\mathrm{y}_\perp^2 < \boldsymbol{c}_\perp, \mathbf{u} >^2\right] \,,$$

and the expression of the MD simplifies to:

$$\begin{aligned}
\mathrm{MD} &= \frac{1}{4\pi} \int_{\boldsymbol{c} \in \mathbb{R}^3 \bigcap \mathbb{S}^2} D(\boldsymbol{c}) dS(\boldsymbol{c}) \\
&= \frac{1}{4\pi} \int_{-1}^{1} \left(\int_{\boldsymbol{c}_\perp \in (\mathrm{span}\{\boldsymbol{\mu}\})^\perp \bigcap \mathbb{S}^2} D(a, \boldsymbol{c}_\perp) dS(\boldsymbol{c}_\perp) \right) da \\
&= \frac{1}{4\pi} \int_{-1}^{1} \Bigg[a^2 \mathrm{E}\left[\mathrm{y}_{/\!/}^2\right] \left(\int_{\boldsymbol{c}_\perp \in (\mathrm{span}\{\boldsymbol{\mu}\})^\perp \bigcap \mathbb{S}^2} dS(\boldsymbol{c}_\perp) \right) \\
&\qquad + (1-a^2) \left(\int_{\boldsymbol{c}_\perp \in (\mathrm{span}\{\boldsymbol{\mu}\})^\perp \bigcap \mathbb{S}^2} \mathrm{E}\left[\mathrm{y}_\perp^2 < \boldsymbol{c}_\perp, \mathbf{u} >^2\right] dS(\boldsymbol{c}_\perp) \right) \Bigg] da \\
&= \frac{1}{2} \int_{-1}^{1} \left[a^2 D_{/\!/} + (1-a^2) D_\perp \right] da = \frac{D_{/\!/} + 2D_\perp}{3} \,.
\end{aligned}$$

List of Acronyms

Acronyms related to brain anatomy

CC	Corpus Callosum	10
CNS	Central Nervous System	
CSF	Cerebro-Spinal Fluid	xxvii
CST	Cortico-Spinal Tract	xxvii
GM	Gray Matter	8
HMT	Hand Motor Tract	117
PNS	Peripheral Nervous System	4
WM	White Matter	

Acronyms related to MRI acquisitions

aMRI	Anatomical MRI	122
BOLD	Blood Oxygen Level Dependent	122
CURVE	CUbe Rays to Vertices and Edges	23
CUSP	CUbe and SPhere	23
dMRI	Diffusion MRI	xxv
DSG	Diffusion-Sensitizing Gradient	xxvi
DW	Diffusion-Weighted	xxvi
EPI	Echo Planar Imaging	122
FLAIR	FLuid Attenuated Inversion Recovery	2
fMRI	Functional MRI	122
HARDI	High Angular Resolution Diffusion Imaging	37
MRI	Magnetic Resonance Imaging	iii
NPA	Narrow Pulse Approximation	43
PGSE	Pulse Gradient Spin Echo	15
RF	Radio Frequency	14
SNR	Signal-to-Noise Ratio	23
TE	Echo Time	15
TFE	Turbo Field Echo	122
TR	Replication Time	15

Acronyms related to brain pathologies

AD Alzheimer's Disease . . . 26
ADHD Attention-Deficit/Hyperactivity Disorder . . . 40
ADNI Alzheimer's Disease Neuroimaging Initiative . . . 30
ALS Amyotrophic Lateral Sclerosis . . . 27
CJD Creutzfeldt-Jakob Disease . . . 2
MCI Mild Cognitive Impairment . . . 26
MS Multiple Sclerosis . . . 25
PPMI Parkinson's Progression Markers Initiative . . . 30
TBI Traumatic Brain Injury . . . 28

Acronyms related to diffusion models

ADC Apparent Diffusion Coefficient . . . 20
CHARMED Composite Hindered And Restricted ModEl of Diffusion . . . 41
DDI Diffusion Directions Imaging . . . iv
DIMD Diffusion-Induced Molecular Displacements . . . xxv
DKI Diffusional Kurtosis Imaging . . . 38
dODF Diffusion Orientation Distribution Function . . . 32
DOT Diffusion Orientation Transform . . . 33
DPI Diffusion Propagator Imaging . . . 32
DSI Diffusion Spectrum Imaging . . . 21
DT Diffusion Tensor . . . iii
DTI Diffusion Tensor Imaging . . . xxv
EAP Ensemble Average Propagator . . . 18
FA Fractional Anisotropy . . . xxviii
fODF Fiber Orientation Distribution Function . . . 32
KT Kurtosis Tensor . . . 38
LI Lattice Index . . . 20
MCM Multi-Compartment Model . . . iv
MD Mean Diffusivity . . . 20
MK Mean Kurtosis . . . 39
MMWMD Minimal Model of White Matter Diffusion . . . 45
NODDI Neurite Orientation Dispersion and Density Imaging . . . iii

OD Orientation Dispersion . . . 59
PAS Persistent Angular Structure . . . 115
QBI Q-Ball Imaging . . . 32
RA Relative Anisotropy . . . 20
SID Signal Intensity Decay . . . v
VR Volume Ratio . . . 20

Acronyms related to tractography

FACT Fiber Assignement by Continuous Tracking . . . iv
HJB Hamilton-Jacobi-Bellman . . . 111
ICBM International Consortium for Brain Mapping . . . 30
ROI Region Of Interest . . . 28
TEND TENsor Deflection . . . 110

Acronyms related to mathematical concepts

AIC Akaike Information Criterion . . . xxviii
ARD Automatic Relevance Detection . . . xxviii
BIC Bayesian Information Criterion . . . 88
BOBYQA Bound Optimization BY Quadratic Approximation . . . 148
CF Characteristic Function . . . 173
COBYLA Constrained Optimization BY Linear Approximations . . . 74
MCMC Markov Chain Monte Carlo . . . xxviii
NEWUOA NEW Unconstrained Optimization Algorithm . . . 119
PDF Probability Density Function . . . 17
RMSE Root Mean Squared Error
SH Spherical Harmonics . . . 33
vMF Von-Mises & Fisher . . . v

Other acronyms

CT Computed Tomography . . . 2
LONI Laboratory Of NeuroImaging . . . 30
MICCAI Medical Image Computing and Computer-Assisted Intervention . . . 35
NIH National Institute of Health . . . viii
PING Pediatric Imaging, Neurocognition and Genetics . . . 30

List of Figures

List of Tables

Bibliography

[Abragam 1961] Anatole Abragam. The principles of nuclear magnetism, volume 32. Oxford University Press, 1961. (Cited on page 14.)

[Abramowitz 1972] M. Abramowitz and I.A. Stegun. Handbook of mathematical functions with formulas, graphs and mathematical tables. Dover Publications, 9th édition, 1972. (Cited on pages xvi, 54, 57, 176 and 177.)

[Aganj 2010] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil and N. Harel. *Reconstruction of the Orientation Distribution Function in single- and multiple-shell Q-Ball Imaging within constant solid angle.* Magnetic Resonance in Medicine, vol. 64, no. 2, pages 554–66, 2010. (Cited on page 131.)

[Aja-Fernández 2012] Santiago Aja-Fernández, Véronique Brion and Antonio Tristán-Vega. *Effective noise estimation and filtering from correlated multiple-coil MR data.* Magnetic Resonance Imaging, 2012. (Cited on page 170.)

[Akaike 1974] Hirotugu Akaike. *A new look at the statistical model identification.* Transactions on Automatic Control, vol. 19, no. 6, pages 716–723, 1974. (Cited on page 92.)

[Alexander 2001] a L Alexander, K M Hasan, M Lazar, J S Tsuruda and D L Parker. *Analysis of partial volume effects in diffusion-tensor MRI.* Magnetic Resonance in Medicine, vol. 45, no. 5, pages 770–80, 2001. (Cited on pages 31 and 114.)

[Alexander 2002] D.C. Alexander, G.J. Barker and S.R. Arridge. *Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data.* Magnetic Resonance Medicine, vol. 48, no. 2, pages 331–40, 2002. (Cited on pages xviii, 31, 50, 88, 89 and 167.)

[Alexander 2006] D.C. Alexander. An introduction to computational diffusion MRI: the diffusion tensor and beyond, chapitre 5, pages 83–106. Visualization and Processing of Tensor Fields. Springer Berlin Heidelberg, 2006. (Cited on page 37.)

[Alexander 2008] D.C. Alexander. *A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features.* Magnetic Resonance in Medicine, vol. 60, no. 2, pages 439–48, 2008. (Cited on pages 45, 49, 80 and 165.)

[Alexander 2010] D.C. Alexander, P.L. Hubbard, M.G. Hall, E.A. Moore, M. Ptito, G.J.M. Parker and T.B. Dyrby. *Orientationally invariant indices of axon diameter and density from diffusion MRI.* NeuroImage, vol. 52, no. 4, pages 1374–89, 2010. (Cited on pages 45, 46, 49, 55, 56, 69, 80, 155 and 165.)

[Anderson 2003] T.W. Anderson. An introduction to multivariate statistical analysis. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley-Interscience, 2003. (Cited on page 175.)

[Arfanakis 2002] Konstantinos Arfanakis, Bruce P Hermann, Baxter P Rogers, John D Carew, Michael Seidenberg and Mary E Meyerand. *Diffusion tensor MRI in temporal lobe epilepsy.* Magnetic Resonance Imaging, vol. 20, no. 7, pages 511–9, 2002. (Cited on page 27.)

[Arsigny 2006] Vincent Arsigny, Pierre Fillard, Xavier Pennec and Nicholas Ayache. *Log-Euclidean metrics for fast and simple calculus on diffusion tensors.* Magnetic Resonance in Medicine, vol. 56, no. 2, pages 411–21, 2006. (Cited on page 119.)

[Assaf 2000] Y. Assaf and Y. Cohen. *Assignment of the water slow-diffusing component in the central nervous system using q-space diffusion MRS: implications for fiber tract imaging.* Magnetic Resonance in Medicine, vol. 43, no. 2, pages 191–9, 2000. (Cited on page 43.)

[Assaf 2002a] Y. Assaf, D. Ben Bashat, J. Chapman, S. Peled, I.E. Biton, M. Kafri, Y. Segev, T. Hendler, A.D. Korczyn, M. Graif and Y. Cohen. *High b-value q-space analyzed diffusion-weighted MRI: Application to multiple sclerosis.* Magnetic Resonance in Medicine, vol. 47, no. 1, pages 115–126, 2002. (Cited on page 44.)

[Assaf 2002b] Y. Assaf, M. Kafri, H. Shinar, J. Chapman, A.D. Korczyn, G. Navon and Y. Cohen. *Changes in axonal morphology in experimental autoimmune neuritis as studied by high b-value q-space (1)H and (2)H DQF diffusion magnetic resonance spectroscopy.* Magnetic Resonance in Medicine, vol. 48, no. 1, pages 71–81, 2002. (Cited on page 44.)

[Assaf 2002c] Y. Assaf, O. Mayzel-Oreg, A. Gigi, D. Ben-Bashat, M. Mordohovitch, R. Verchovsky, I.I. Reider-Groswasser, T. Hendler, M. Graif, Y. Cohen and A.D. Korczyn. *High b value q-space-analyzed diffusion MRI in vascular dementia: a preliminary study.* Journal of the Neurological Sciences, vol. 203-204, pages 235–9, 2002. (Cited on page 44.)

[Assaf 2003] Bassam a Assaf, Feroze B Mohamed, Karine J Abou-Khaled, J Michael Williams, May S Yazeji, John Haselgrove and Scott H Faro. *Diffusion tensor imaging of the hippocampal formation in temporal lobe epilepsy.* American Journal of Neuroradiology, vol. 24, no. 9, pages 1857–62, 2003. (Cited on page 27.)

[Assaf 2004] Y. Assaf, R.Z. Freidlin, G.K. Rohde and P.J. Basser. *New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter.* Magnetic Resonance in Medicine, vol. 52, no. 5, pages 965–78, 2004. (Cited on pages xiv, 41, 44, 45, 46, 55, 80 and 165.)

[Assaf 2005a] Y. Assaf and P.J. Basser. *Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain.* NeuroImage, vol. 27, no. 1, pages 48–58, 2005. (Cited on pages 31, 35, 45, 46, 47, 50, 51, 55, 59, 80 and 165.)

[Assaf 2005b] Y. Assaf, R.Z. Freidlin and P.J. Basser. *The measurement of the axon diameter distribution in white matter using diffusion MR methods.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 13, page 842, 2005. (Cited on pages xiv, 45 and 47.)

[Assaf 2008a] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel and P.J. Basser. *AxCaliber: a method for measuring axon diameter distribution from diffusion MRI.* Magnetic Resonance in Medicine, vol. 59, no. 6, pages 1347–54, 2008. (Cited on pages xiv, 45 and 46.)

[Assaf 2008b] Yaniv Assaf and Ofer Pasternak. *Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review.* Journal of Molecular Neuroscience, vol. 34, no. 1, pages 51–61, 2008. (Cited on page 31.)

[Assemlal 2011] H.E. Assemlal, D. Tschumperlé, L. Brun and K. Siddiqi. *Recent advances in diffusion MRI modeling: Angular and radial reconstruction.* Medical Image Analysis, 2011. (Cited on pages 33 and 118.)

[Banerjee 2006] Arindam Banerjee, IS Dhillon, J Ghosh and S Sra. *Clustering on the unit hypersphere using von Mises-Fisher distributions.* Journal of Machine Learning Research, vol. 6, pages 1345–1382, 2006. (Cited on page 129.)

[Barazany 2009] D. Barazany, P.J. Basser and Y. Assaf. *In vivo measurement of axon diameter distribution in the corpus callosum of rat brain.* Brain, vol. 132, no. 5, pages 1210–20, 2009. (Cited on pages xiv, 45 and 46.)

[Basser 1994a] P.J. Basser, J. Mattiello and D. Le Bihan. *Estimation of the effective self-diffusion tensor from the NMR spin echo.* Journal of Magnetic Resonance. Series B, vol. 103, no. 3, pages 247–54, 1994. (Cited on page 22.)

[Basser 1994b] P.J. Basser, J. Mattiello and D. Le Bihan. *MR diffusion tensor spectroscopy and imaging.* Biophysical Journal, vol. 66, no. 1, pages 259–67, 1994. (Cited on pages xii, 22, 23 and 107.)

[Basser 1996a] P J Basser and C Pierpaoli. *Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI.* Journal of Magnetic Resonance. Series B, vol. 111, no. 3, pages 209–19, 1996. (Cited on pages xii, 20 and 59.)

[Basser 1996b] Peter J Basser, James H Mattiello and Denis LeBihan. *Method and system for measuring the diffusion tensor and for diffusion tensor imaging*, 1996. US Patent 5,539,310. (Cited on pages 22 and 24.)

[Basser 2000] P J Basser, S Pajevic, C Pierpaoli, J Duda and a Aldroubi. *In vivo fiber tractography using DT-MRI data.* Magnetic Resonance in Medicine, vol. 44, no. 4, pages 625–32, 2000. (Cited on pages xix and 109.)

[Basser 2002] Peter J Basser and Derek K Jones. *Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 456–67, 2002. (Cited on pages 38 and 117.)

[Basser 2009] P.J. Basser and E. Ozarslan. *Introduction to diffusion MR.* In H. Johansen-Berg and T.E.J. Behrens, editeurs, Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy., pages 205–236. 2009. (Cited on pages 11, 13, 17, 189 and 190.)

[Batchelor 2002] P Batchelor. *Optimisation of direction schemes for diffusion tensor imaging.* In Proceedings of the workshop on diffusion MRI: biophysical issues (What can we measure), 2002. (Cited on page 23.)

[Beaulieu 2009] C. Beaulieu. *The biological basis of diffusion anisotropy.* In H. Johansen-Berg and T.E.J. Behrens, editeurs, Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy., pages 105–126. Academic Press, 2009. (Cited on page 31.)

[Behrens 2003] T.E.J. Behrens, M.W. Woolrich, M. Jenkinson, H. Johansen-Berg, R.G. Nunes, S. Clare, P.M. Matthews, J.M. Brady and S.M. Smith. *Characterization and propagation of uncertainty in diffusion-weighted MR imaging.* Magnetic Resonance in Medicine, vol. 50, no. 5, pages 1077–88, 2003. (Cited on pages 41, 42, 59, 72, 80, 90, 114, 157 and 165.)

[Behrens 2007] T.E.J. Behrens, H. Johansen-Berg, S. Jbabdi, M.F.S. Rushworth and M.W. Woolrich. *Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?* NeuroImage, vol. 34, no. 1, pages 144–55, 2007. (Cited on pages xiii, xviii, 42, 50, 89, 90, 91, 93, 96, 113, 114, 118, 165 and 167.)

[Bennett 2003] Kevin M Bennett, Kathleen M Schmainda, Raoqiong Tong Bennett, Daniel B Rowe, Hanbing Lu and James S Hyde. *Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model.* Magnetic Resonance in Medicine, vol. 50, no. 4, pages 727–34, 2003. (Cited on page 50.)

[Berman 2008] Jeffrey I Berman, SungWon Chung, Pratik Mukherjee, Christopher P Hess, Eric T Han and Roland G Henry. *Probabilistic streamline q-ball tractography using the residual bootstrap.* NeuroImage, vol. 39, no. 1, pages 215–22, 2008. (Cited on pages xix and 115.)

[Bhalerao 2007] Abhir Bhalerao and Carl-Fredrik Westin. *Hyperspherical von Mises-Fisher mixture (HvMF) modelling of high angular resolution diffusion*

MRI. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 10, pages 236–43, 2007. (Cited on page 129.)

[Björnemo 2002] M. Björnemo, A. Brun, R. Kikinis and C.F. Westin. *Regularized stochastic white matter tractography using diffusion tensor MRI.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 2488, pages 435–442, 2002. (Cited on pages xix and 116.)

[Bloch 1946] Felix Bloch, WW Hansen and Martin Packard. *Nuclear induction.* Physical Review, vol. 70, no. 7-8, pages 460–474, 1946. (Cited on pages 2 and 18.)

[Bodini 2009] B. Bodini and O. Ciccarelli. *Diffusion MRI in neurological disorders.* In H. Johansen-Berg and T.E.J. Behrens, editeurs, Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy., pages 175–204. Academic Press, 2009. (Cited on page 31.)

[Bokkers 2012] Reinoud P H Bokkers, Daymara a Hernandez, José G Merino, Raymond V Mirasol, Matthias J van Osch, Jeroen Hendrikse, Steven Warach and Lawrence L Latour. *Whole-brain arterial spin labeling perfusion MRI in patients with acute stroke.* Stroke. A journal of Cerebral Circulation, vol. 43, no. 5, pages 1290–4, 2012. (Cited on pages 3 and 189.)

[Boss 1965] B.D. Boss and E.O. Stejskal. *Anisotropic Diffusion in Hydrated Vermiculite.* Journal of Chemical Physics, vol. 43, no. 3, page 1068, 1965. (Cited on page 43.)

[Bossy 1990] Jean Bossy. Neuro-anatomie, anatomie clinique. Springer-Verlag, Paris, 1990. (Cited on page 6.)

[Brown 1828] Robert Brown. *XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies.* The Philosophical Magazine, or Annals of Chemistry, Mathematics, Astronomy, Natural History and General Science, vol. 4, no. 21, pages 161–173, 1828. (Cited on page 13.)

[Brun 2002] A Brun, M Bjornemo, R Kikinis and CF Westin. *White matter tractography using sequential importance sampling.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 10, 2002. (Cited on pages xix and 116.)

[Buchsbaum 2006] Monte S Buchsbaum, Joseph Friedman, Bradley R Buchsbaum, King-Wai Chu, Erin a Hazlett, Randall Newmark, Jason S Schneiderman, Yuliya Torosjan, Cheuk Tang, Patrick R Hof, Daniel Stewart, Kenneth L

Davis and Jack Gorman. *Diffusion tensor imaging in schizophrenia.* Biological Psychiatry, vol. 60, no. 11, pages 1181–7, 2006. (Cited on page 27.)

[Budde 2007] Matthew D Budde, Joong Hee Kim, Hsiao-Fang Liang, Robert E Schmidt, John H Russell, Anne H Cross and Sheng-Kwei Song. *Toward accurate diagnosis of white matter pathology using diffusion tensor imaging.* Magnetic Resonance in Medicine, vol. 57, no. 4, pages 688–95, 2007. (Cited on page 32.)

[Budde 2011] Matthew D Budde, Lindsay Janes, Eric Gold, Lisa Christine Turtzo and Joseph a Frank. *The contribution of gliosis to diffusion tensor anisotropy and tractography following traumatic brain injury: validation in the rat using Fourier analysis of stained tissue sections.* Brain: A Journal of Neurology, vol. 134, no. 8, pages 2248–60, 2011. (Cited on page 32.)

[Bürgel 2006] U. Bürgel, K. Amunts, L. Hoemke, H. Mohlberg, J.M. Gilsbach and K. Zilles. *White matter fiber tracts of the human brain: three-dimensional mapping at microscopic resolution, topography and intersubject variability.* NeuroImage, vol. 29, no. 4, pages 1092–105, 2006. (Cited on page 47.)

[Burnham 2002] K P Burnham and D R Anderson. Model selection and multimodel inference: a practical information-theoretic approach. Springer, 2002. (Cited on pages 92 and 93.)

[Caan 2010] Matthan W a Caan, H Ganesh Khedoe, Dirk H J Poot, Arjan J den Dekker, Silvia D Olabarriaga, Kees a Grimbergen, Lucas J van Vliet and Frans M Vos. *Estimation of diffusion properties in crossing fiber bundles.* Transactions on Medical Imaging, vol. 29, no. 8, pages 1504–15, 2010. (Cited on pages 31, 50 and 51.)

[Callaghan 1991] P.T. Callaghan. Principles of Nuclear Magnetic Resonance microscopy. Oxford University Press, 1991. (Cited on pages x, 17 and 43.)

[Campbell 2004] JSW Campbell, K Siddiqi and GB Pike. *Full-brain q-ball imaging in a clinically acceptable time: Application to white matter fibre tractography.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 11, page 448, 2004. (Cited on page 34.)

[Canales-Rodríguez 2009] Erick Jorge Canales-Rodríguez, Lester Melie-García and Yasser Iturria-Medina. *Mathematical description of q-space in spherical coordinates: exact q-ball imaging.* Magnetic Resonance in Medicine, vol. 61, no. 6, pages 1350–67, 2009. (Cited on page 33.)

[Canales-Rodríguez 2010] Erick Jorge Canales-Rodríguez, Ching-Po Lin, Yasser Iturria-Medina, Chun-Hung Yeh, Kuan-Hung Cho and Lester Melie-García. *Diffusion orientation transform revisited.* NeuroImage, vol. 49, no. 2, pages 1326–39, 2010. (Cited on page 33.)

[Carr 1954] H. Carr and E. Purcell. *Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments.* Physical Review, vol. 94, no. 3, pages 630–638, 1954. (Cited on page 15.)

[Catani 2002] Marco Catani, Robert J. Howard, Sinisa Pajevic and Derek K. Jones. *Virtual in Vivo Interactive Dissection of White Matter Fasciculi in the Human Brain.* NeuroImage, vol. 17, no. 1, pages 77–94, 2002. (Cited on pages x, 9, 28, 92, 117, 142 and 164.)

[Catani 2006] Marco Catani. *Diffusion tensor magnetic resonance imaging tractography in cognitive disorders.* Current Opinion in Neurology, vol. 19, no. 6, pages 599–606, 2006. (Cited on page 29.)

[Catani 2008] Marco Catani and Michel Thiebaut de Schotten. *A diffusion tensor imaging tractography atlas for virtual in vivo dissections.* Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, vol. 44, no. 8, pages 1105–32, 2008. (Cited on pages x, 9, 28, 92, 117, 142, 164 and 171.)

[Chang 2002] Shih Chin Chang, Ping Hong Lai, Wei Liang Chen, Hsu Huei Weng, Jih Tsun Ho, Jyh Seng Wang, Chuan Yu Chang, Huay Ben Pan and Chien Fang Yang. *Diffusion-weighted MRI features of brain abscess and cystic or necrotic brain tumors: comparison with conventional MRI.* Clinical Imaging, vol. 26, no. 4, pages 227–36, 2002. (Cited on page 2.)

[Cheng 2010] J. Cheng, A. Ghosh, T. Jiang and R. Deriche. *Model-free and analytical EAP reconstruction via spherical polar Fourier diffusion MRI.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 13, pages 590–7, 2010. (Cited on page 33.)

[Cheung 2009] Matthew M Cheung, Edward S Hui, Kevin C Chan, Joseph a Helpern, Liqun Qi and Ed X Wu. *Does diffusion kurtosis imaging lead to better neural tissue characterization? A rodent brain maturation study.* NeuroImage, vol. 45, no. 2, pages 386–92, 2009. (Cited on pages xiii, xiv, 31, 40, 43, 50 and 51.)

[Chua 2008] Terence C Chua, Wei Wen, Melissa J Slavin and Perminder S Sachdev. *Diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease: a review.* Current Opinion in Neurology, vol. 21, no. 1, pages 83–92, 2008. (Cited on page 26.)

[Ciccarelli 2008] Olga Ciccarelli, Marco Catani, Heidi Johansen-Berg, Chris Clark and Alan Thompson. *Diffusion-based tractography in neurological disorders: concepts, applications, and future developments.* Lancet Neurology, vol. 7, no. 8, pages 715–27, 2008. (Cited on pages xii, xiii, 29 and 32.)

[Ciofolo 2009] Cybèle Ciofolo and Christian Barillot. *Atlas-based segmentation of 3D cerebral structures with competitive level sets and fuzzy control.* Medical

Image Analysis, vol. 13, no. 3, pages 456–470, 2009. (Cited on pages 8 and 189.)

[Constantinides 1997] C.D. Constantinides, E. Atalar and E.R. McVeign. *Signal-to-noise measurements in magnitude images from NMR phased arrays.* Magnetic Resonance in Medicine, vol. 38, no. 5, pages 852–857, 1997. (Cited on page 170.)

[Conturo 1996] Thomas E Conturo, Robert C McKinstry, Erbil Akbudak and Bruce H Robinson. *Encoding of anisotropic diffusion with tetrahedral gradients: a general mathematical diffusion formalism and experimental results.* Magnetic Resonance in Medicine, vol. 35, no. 3, pages 399–412, 1996. (Cited on page 23.)

[Conturo 1999] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton and M. E. Raichle. *Tracking neuronal fiber pathways in the living human brain.* In Proceedings of the National Academy of Sciences, volume 96, pages 10422–10427, 1999. (Cited on pages xix, 109 and 110.)

[Cook 2004] P.A. Cook, D.C. Alexander and G.J.M. Parker. *Modelling noise-induced fibre-orientation error in diffusion-tensor MRI.* In Proceedings of the International Symposium on Biomedical Imaging, volume 2, pages 332–335, 2004. (Cited on page 115.)

[Cook 2006] P.A. Cook, Y. Bai, K.K. Seunarine, M.G. Hall, G.J.M. Parker and D.C. Alexander. *Camino : Open-Source Diffusion-MRI Reconstruction and Processing.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 14, page 2759, 2006. (Cited on page 73.)

[Correia 2009] Marta Morgado Correia, Thomas a Carpenter and Guy B Williams. *Looking for the optimal DTI acquisition scheme given a maximum scan time: are more b-values a waste of time?* Magnetic Resonance Imaging, vol. 27, no. 2, pages 163–75, 2009. (Cited on page 22.)

[Damadian 1974] Raymond V Damadian. *Apparatus and method for detecting cancer in tissue*, 1974. US Patent 3,789,832. (Cited on pages viii and 2.)

[De Keyser 2008] Jacques De Keyser, Jop P Mostert and Marcus W Koch. *Dysfunctional astrocytes as key players in the pathogenesis of central nervous system disorders.* Journal of the Neurological Sciences, vol. 267, no. 1-2, pages 3–16, 2008. (Cited on pages 7 and 189.)

[De Oliveira 1987] AR De Oliveira. *Un répertoire des accidents radiologiques 1945–1985.* Radioprotection, vol. 22, no. 2, pages 89–135, 1987. (Cited on page 2.)

[De Santis 2011] S. De Santis, Y. Assaf, C.J. Evans and D.K. Jones. *Improved precision in the CHARMED model of white matter through sampling scheme*

optimization and model parsimony testing. In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 19, page 3928, 2011. (Cited on page 46.)

[De Santis 2012] S. De Santis, Y. Assaf and D.K. Jones. *Using the biophysical CHARMED model to elucidate the underpinnings of contrast in diffusional kurtosis analysis of diffusion-weighted MRI.* Magma, vol. 25, no. 4, pages 267–76, 2012. (Cited on page 46.)

[Dégerine 1979] S. Dégerine. *Lois de von Mises et lois liées.* Annales de l'Institut Henri Poincaré, vol. 15, no. 1, pages 63–77, 1979. (Cited on page 174.)

[Dell'Acqua 2012] Flavio Dell'Acqua and Marco Catani. *Structural human brain networks: hot topics in diffusion tractography.* Current Opinion in Neurology, vol. 25, no. 4, pages 375–83, 2012. (Cited on pages 34 and 164.)

[Demiralp 2011] C. Demiralp and D.H. Laidlaw. *Generalizing diffusion tensor model using probabilistic inference in Markov random fields.* In MICCAI CDMRI Workshop, 2011. (Cited on page xviii.)

[Deriche 2009] R. Deriche, J. Calder and M. Descoteaux. *Optimal real-time Q-ball imaging using regularized Kalman filtering with incremental orientation sets.* Medical Image Analysis, vol. 13, no. 4, pages 564–79, 2009. (Cited on page 33.)

[Deriche 2013] R. Deriche. *Continuous diffusion signal, EAP and ODF estimation via Compressive Sensing in diffusion MRI.* Medical Image Analysis, 2013. (Cited on pages 33 and 164.)

[Descoteaux 2007] M. Descoteaux, E. Angelino, S. Fitzgibbons and R. Deriche. *Regularized, fast, and robust analytical Q-Ball Imaging.* Magnetic Resonance in Medicine, vol. 58, no. 3, pages 497–510, 2007. (Cited on pages 33 and 119.)

[Descoteaux 2009] M. Descoteaux, R. Deriche, T.R. Knösche and A. Anwander. *Deterministic and probabilistic tractography based on complex fibre orientation distributions.* Transactions on Medical Imaging, vol. 28, no. 2, pages 269–86, 2009. (Cited on pages 33 and 119.)

[Descoteaux 2011] M. Descoteaux, R. Deriche, D. Le Bihan, J.F. Mangin and C. Poupon. *Multiple q-shell diffusion propagator imaging.* Medical Image Analysis, vol. 15, no. 4, pages 603–21, 2011. (Cited on page 32.)

[Douaud 2011] Gwenaëlle Douaud, Saâd Jbabdi, Timothy E J Behrens, Ricarda a Menke, Achim Gass, Andreas U Monsch, Anil Rao, Brandon Whitcher, Gordon Kindlmann, Paul M Matthews and Stephen Smith. *DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease.* NeuroImage, vol. 55, no. 3, pages 880–90, 2011. (Cited on page 26.)

[Doucet 2000] Arnaud Doucet, Simon Godsill and C. Andrieu. *On sequential Monte Carlo sampling methods for Bayesian filtering.* Statistics and Computing, vol. 10, no. 3, pages 197–208, 2000. (Cited on pages xxi, 116, 128 and 129.)

[Duvernoy 1999] Henri M Duvernoy and B Parratte. The human brain: surface, three-dimensional sectional anatomy with mri, and blood supply. Springer Verlag Wien, 1999. (Cited on page 5.)

[Efron 1983] Bradley Efron. *Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation.* Journal of the American Statistical Association, vol. 78, no. 382, pages 316–331, 1983. (Cited on pages 88 and 89.)

[Efron 1997] Bradley Efron and Robert Tibshirani. *Improvements on Cross-Validation: The 632+ Bootstrap Method.* Journal of the American Statistical Association, vol. 92, no. 438, pages 548–560, 1997. (Cited on page 89.)

[Einstein 1905] Albert Einstein. *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen.* Annalen der Physik, vol. 322, no. 8, pages 549–560, 1905. (Cited on pages xi, 13, 17 and 19.)

[Ennis 2006] Daniel B Ennis and Gordon Kindlmann. *Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images.* Magnetic Resonance in Medicine, vol. 55, no. 1, pages 136–46, 2006. (Cited on page 31.)

[Falangola 2007] MF Falangola, A Ramani, A Di Martino, JH Jensen, JS Babb, C Hu, K Szulc, FX Castellanos and JA Helpern. *Age-Related Patterns of Change in Brain Microstructure by Diffusional Kurtosis Imaging.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 19, page 667, 2007. (Cited on page 40.)

[Fick 1855a] A. Fick. *Concerns diffusion and concentration gradient.* Annalen der Physik, vol. 170, no. 59, 1855. (Cited on page 13.)

[Fick 1855b] Adolf Fick. *Ueber diffusion.* Annalen der Physik, vol. 170, no. 1, pages 59–86, 1855. (Cited on page 13.)

[Fieremans 2011] Els Fieremans, Jens H Jensen and Joseph a Helpern. *White matter characterization with diffusional kurtosis imaging.* NeuroImage, vol. 58, no. 1, pages 177–88, 2011. (Cited on page 46.)

[Fillard 2007] P. Fillard, X. Pennec, V. Arsigny and N. Ayache. *Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics.* Transactions on Medical Imaging, vol. 26, no. 11, pages 1472–82, 2007. (Cited on pages xii, 23, 119 and 170.)

[Fillard 2009] Pierre Fillard, Cyril Poupon and Jean-François Mangin. *A novel global tractography algorithm based on an adaptive spin glass model.* In Proceedings of the International Conference on Medical Image Computing and

Computer-Assisted Intervention, volume 12, pages 927–34, 2009. (Cited on pages xix and 112.)

[Fletcher 2007] P Thomas Fletcher, Ran Tao, Won-Ki Jeong and Ross T Whitaker. *A volumetric approach to quantifying region-to-region white matter connectivity in diffusion tensor MRI.* In Proceedings of the International Conference on Information Processing in Medical Imaging, volume 20, pages 346–58, 2007. (Cited on page 111.)

[Frank 2002] Lawrence R Frank. *Characterization of anisotropy in high angular resolution diffusion-weighted MRI.* Magnetic Resonance in Medicine, vol. 47, no. 6, pages 1083–99, 2002. (Cited on page 33.)

[Friman 2006] Ola Friman, Gunnar Farnebäck and Carl Fredrik Westin. *A Bayesian approach for stochastic white matter tractography.* Transactions on Medical Imaging, vol. 25, no. 8, pages 965–78, 2006. (Cited on page 115.)

[Garcia 2010] Vincent Garcia, Olivier Commowick and Grégoire Malandain. *A Robust and Efficient Block Matching Framework for Non Linear Registration of Thoracic CT Images.* In Grand Challenge in Medical Image Analysis Workshop of the International Conference on Medical Image Computing and Computer-Assisted Intervention, 2010. (Cited on page 144.)

[Gasser 1986] Theo Gasser, Lothar Sroka and Christine Jennen-Steinmetz. *Residual variance and residual pattern in nonlinear regression.* Biometrika, vol. 73, no. 3, pages 625–633, 1986. (Cited on page 130.)

[Ghosh 2010] A. Ghosh and R. Deriche. *Fast and closed-form Ensemble-Average-Propagator approximation from 4th-order diffusion tensor.* In Proceedings of the International Symposium on Biomedical Imaging, pages 1105–8, 2010. (Cited on page 33.)

[Ghosh 2013a] A. Ghosh, T. Milne and R. Deriche. *Constrained diffusion kurtosis imaging using ternary quartics & MLE.* Magnetic Resonance in Medicine, pages 1–11, 2013. (Cited on page 39.)

[Ghosh 2013b] A. Ghosh, E. Tsigaridas, B. Mourrain and R. Deriche. *A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI.* Medical Image Analysis, vol. 17, no. 5, pages 503–14, 2013. (Cited on page 32.)

[Gössl 2002] C Gössl, L Fahrmeir, B Pütz, L M Auer and D P Auer. *Fiber tracking from DTI using linear state space models: detectability of the pyramidal tract.* NeuroImage, vol. 16, no. 2, pages 378–88, 2002. (Cited on page 113.)

[Gudbjartsson 1995] HáKon Gudbjartsson and Samuel Patz. *The Rician distribution of noisy MRI data.* Magnetic Resonance in Medicine, vol. 34, no. 6, pages 910–914, 1995. (Cited on pages 22, 39, 49, 74, 114, 130 and 170.)

[Hager 1989] W.W. Hager. *Updating the inverse of a matrix.* SIAM Review, vol. 31, no. 2, pages 221–39, 1989. (Cited on pages 53 and 175.)

[Hahn 1950] Erwin L Hahn. *Spin echoes.* Physical Review, vol. 80, no. 4, page 580, 1950. (Cited on page 14.)

[Hall 2009] M.G. Hall and D.C. Alexander. *Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI.* Transactions on Medical Imaging, vol. 28, no. 9, pages 1354–64, 2009. (Cited on page 73.)

[Han 2010] Bong Soo Han, Ji Heon Hong, Cheolpyo Hong, Sang Seok Yeo, Dong Hoon Lee, Hee Kyung Cho and Sung Ho Jang. *Location of the corticospinal tract at the corona radiata in human brain.* Brain Research, vol. 1326, pages 75–80, 2010. (Cited on page 117.)

[Han 2013] Xiaoning Han, Michael Chen, Fushun Wang, Martha Windrem, Su Wang, Steven Shanz, Qiwu Xu, Nancy Ann Oberheim, Lane Bekar, Sarah Betstadt, Alcino J Silva, Takahiro Takano, Steven a Goldman and Maiken Nedergaard. *Forebrain engraftment by human glial progenitor cells enhances synaptic plasticity and learning in adult mice.* Cell Stem Cell, vol. 12, no. 3, pages 342–53, 2013. (Cited on page 7.)

[Hasan 2001] K M Hasan, D L Parker and a L Alexander. *Comparison of gradient encoding schemes for diffusion-tensor MRI.* Journal of Magnetic Resonance Imaging, vol. 13, no. 5, pages 769–80, 2001. (Cited on page 23.)

[Hasan 2010] K.M. Hasan, I.S. Walimuni, H. Abid and K.R. Hahn. *A review of diffusion tensor magnetic resonance imaging computational methods and software tools.* Computers in Biology and Medicine, pages 1–11, 2010. (Cited on pages 23 and 38.)

[Helpern 2007] JA Helpern, MF Falangola, A Di Martino, A Ramani, JS Babb, C Hu, JH Jensen and FX Castellanos. *Alterations in Brain Microstructure in ADHD by Diffusional Kurtosis Imaging.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 15, page 1580, 2007. (Cited on page 40.)

[Holodny 2005] Andrei I Holodny, Devang M Gor, Richard Watts, Philip H Gutin and Aziz M Ulug. *Diffusion-tensor MR tractography of somatotopic organization of corticospinal tracts in the internal capsule: initial anatomic results in contradistinction to prior reports.* Radiology, vol. 234, no. 3, pages 649–53, 2005. (Cited on page 117.)

[Hong 2010] Ji Heon Hong, Su Min Son and Sung Ho Jang. *Somatotopic location of corticospinal tract at pons in human brain: a diffusion tensor tractography study.* NeuroImage, vol. 51, no. 3, pages 952–5, 2010. (Cited on page 117.)

[Horsfield 2002] Mark a Horsfield and Derek K Jones. *Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - a review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 570–7, 2002. (Cited on pages xii, 25, 26 and 27.)

[Hosey 2005] T. Hosey, G. Williams and R. Ansorge. *Inference of multiple fiber orientations in high angular resolution diffusion imaging.* Magnetic Resonance in Medicine, vol. 54, no. 6, pages 1480–9, 2005. (Cited on pages xiv, 42, 43, 72, 80 and 165.)

[Hounsfield 1973] Godfrey N Hounsfield. *Computerized transverse axial scanning (tomography): Part 1. Description of system.* British Journal of Radiology, vol. 46, no. 552, pages 1016–1022, 1973. (Cited on page 2.)

[Hui 2008] Edward S Hui, Matthew M Cheung, Liqun Qi and Ed X Wu. *Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis.* NeuroImage, vol. 42, no. 1, pages 122–34, 2008. (Cited on pages 39, 40 and 60.)

[Iwata 2008] N K Iwata, S Aoki, S Okabe, N Arai, Y Terao, S Kwak, O Abe, I Kanazawa, S Tsuji and Y Ugawa. *Evaluation of corticospinal tracts in ALS with diffusion tensor MRI and brainstem stimulation.* Neurology, vol. 70, no. 7, pages 528–32, 2008. (Cited on page 118.)

[Jackowski 2005] Marcel Jackowski, Chiu Yen Kao, Maolin Qiu, R Todd Constable and Lawrence H Staib. *White matter tractography by anisotropic wavefront evolution and diffusion tensor imaging.* Medical Image Analysis, vol. 9, no. 5, pages 427–40, 2005. (Cited on pages xix and 111.)

[Jakob 2012] Wenzel Jakob. *Numerically stable sampling of the von Mises Fisher distribution on the 2-sphere (and other tricks).*, 2012. (Cited on pages 116 and 132.)

[Jbabdi 2007] S Jbabdi, M W Woolrich, J L R Andersson and T E J Behrens. *A Bayesian framework for global tractography.* NeuroImage, vol. 37, no. 1, pages 116–29, 2007. (Cited on page 114.)

[Jbabdi 2012] Saad Jbabdi, Stamatios N Sotiropoulos, Alexander M Savio, Manuel Graña and Timothy E J Behrens. *Model-based analysis of multishell diffusion MR data for tractography: how to get over fitting problems.* Magnetic Resonance in Medicine, vol. 68, no. 6, pages 1846–55, 2012. (Cited on pages 50 and 59.)

[Jellison 2004] Brian J Jellison, Aaron S Field, Joshua Medow, Mariana Lazar, M Shariar Salamat and Andrew L Alexander. *Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns.* AJNR. American Journal of Neuroradiology, vol. 25,

no. 3, pages 356–69, 2004. (Cited on pages x, xi, 9, 10, 11, 12, 92, 117, 142, 164 and 189.)

[Jenkinson 2012] Mark Jenkinson, Christian F Beckmann, Timothy E J Behrens, Mark W Woolrich and Stephen M Smith. *FSL.* NeuroImage, vol. 62, no. 2, pages 782–90, 2012. (Cited on page 96.)

[Jensen 2003] JH Jensen and JA Helpern. *Quantifying Non-Gaussian Water Diffusion by Means of Pulsed-Field-Gradient MRI.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 859, page 2154, 2003. (Cited on page 40.)

[Jensen 2005] Jens H Jensen, Joseph a Helpern, Anita Ramani, Hanzhang Lu and Kyle Kaczynski. *Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging.* Magnetic Resonance in Medicine, vol. 53, no. 6, pages 1432–40, 2005. (Cited on pages 38, 39, 40 and 50.)

[Jensen 2010] Jens H Jensen and Joseph A Helpern. *MRI quantification of non-Gaussian water diffusion by kurtosis analysis.* NMR in Biomedicine, vol. 23, no. 7, pages 698–710, 2010. (Cited on pages 39, 40, 58 and 60.)

[Jespersen 2007] S.N. Jespersen, C.D. Kroenke, L. Ø stergaard, J.J.H. Ackerman and D.A. Yablonskiy. *Modeling dendrite density from magnetic resonance diffusion measurements.* NeuroImage, vol. 34, no. 4, pages 1473–86, 2007. (Cited on pages 47 and 48.)

[Jespersen 2010] S.N. Jespersen, C.R. Bjarkam, J.R. Nyengaard, M.M. Chakravarty, B. Hansen, T. Vosegaard, L. Ø stergaard, D.A. Yablonskiy, N.C. Nielsen and P. Vestergaard-Poulsen. *Neurite density from magnetic resonance diffusion measurements at ultrahigh field: comparison with light microscopy and electron microscopy.* NeuroImage, vol. 49, no. 1, pages 205–16, 2010. (Cited on page 48.)

[Jespersen 2012] S.N. Jespersen, L.A. Leigland, A. Cornea and C.D. Kroenke. *Determination of axonal and dendritic orientation distributions within the developing cerebral cortex by diffusion tensor imaging.* Transactions on Medical Imaging, vol. 31, no. 1, pages 16–32, 2012. (Cited on page 48.)

[Jeurissen 2011] Ben Jeurissen, Alexander Leemans, Derek K Jones, Jacques-Donald Tournier and Jan Sijbers. *Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution.* Human Brain Mapping, vol. 32, no. 3, pages 461–79, 2011. (Cited on page 115.)

[Jian 2007] B. Jian and B.C. Vemuri. *A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI.* Transactions on Medical Imaging, vol. 26, no. 11, pages 1464–71, 2007. (Cited on page 33.)

[Johnson 2013] S.G. Johnson. *The NLopt nonlinear-optimization package.* http://ab-initio.mit.edu/nlopt, 2013. (Cited on page 74.)

[Jones 2003] D.K. Jones. *Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI.* Magnetic Resonance in Medicine, vol. 49, pages 7–12, 2003. (Cited on page 114.)

[Jones 2004] Derek K Jones. *The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study.* Magnetic Resonance in Medicine, vol. 51, no. 4, pages 807–15, 2004. (Cited on pages xii, 22 and 23.)

[Jones 2006] Derek K Jones, Marco Catani, Carlo Pierpaoli, Suzanne J C Reeves, Sukhwinder S Shergill, Michael O'Sullivan, Pasha Golesworthy, Phillip McGuire, Mark a Horsfield, Andrew Simmons, Steven C R Williams and Robert J Howard. *Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia.* Human Brain Mapping, vol. 27, no. 3, pages 230–8, 2006. (Cited on page 29.)

[Jones 2007] D.K. Jones. *How many shells? - investigating a long held tradition in DT-MRI.* In Proceedings of the Interntional Society for Magnetic Resonance in Medicine, 2007. (Cited on page 22.)

[Jupp 1989] P.E. Jupp and K.V. Mardia. *A unified view of the theory of directional Statistics, 1975-1988.* International Statistical Review, vol. 57, no. 3, pages 261–94, 1989. (Cited on pages xx, 53, 127 and 175.)

[Kaden 2007] E. Kaden, T.R. Knösche and A. Anwander. *Parametric spherical deconvolution: inferring anatomical connectivity using diffusion MR imaging.* NeuroImage, vol. 37, no. 2, pages 474–88, 2007. (Cited on pages 47, 50 and 115.)

[Kanaan 2005] Richard a a Kanaan, Jin-Suh Kim, Walter E Kaufmann, Godfrey D Pearlson, Gareth J Barker and Philip K McGuire. *Diffusion tensor imaging in schizophrenia.* Biological Psychiatry, vol. 58, no. 12, pages 921–9, 2005. (Cited on page 27.)

[Kastler 2011] Bruno Kastler, Daniel Vetter, Zoltan Patay and Philippe Germain. Comprendre l'irm: manuel d'auto-apprentissage, volume 6. Elsevier Masson, 2011. (Cited on page 2.)

[Kim 2008] Yun-Hee Kim, Dae-Shik Kim, Ji Heon Hong, Chang Hyun Park, Ning Hua, Kevin C Bickart, Woo Mok Byun and Sung Ho Jang. *Corticospinal tract location in internal capsule of human brain: diffusion tensor tractography and functional MRI study.* Neuroreport, vol. 19, no. 8, pages 817–20, 2008. (Cited on page 117.)

[Kono 2001] K Kono, Y Inoue, K Nakayama, M Shakudo, M Morino, K Ohata, K Wakasa and R Yamada. *The role of diffusion-weighted imaging in patients with brain tumors.* AJNR. American Journal of Neuroradiology, vol. 22, no. 6, pages 1081–8, 2001. (Cited on page 2.)

[Kreher 2005] B W Kreher, J F Schneider, I Mader, E Martin, J Hennig and K a Il'yasov. *Multitensor approach for analysis and tracking of complex fiber configurations.* Magnetic Resonance in Medicine, vol. 54, no. 5, pages 1216–25, 2005. (Cited on pages xviii, 88 and 167.)

[Kreher 2008] B W Kreher, I Mader and V G Kiselev. *Gibbs tracking: a novel approach for the reconstruction of neuronal pathways.* Magnetic Resonance in Medicine, vol. 60, no. 4, pages 953–63, 2008. (Cited on pages xix and 112.)

[Kubicki 2007] Marek Kubicki, Robert McCarley, Carl-Fredrik Westin, Hae-Jeong Park, Stephan Maier, Ron Kikinis, Ferenc a Jolesz and Martha E Shenton. *A review of diffusion tensor imaging studies in schizophrenia.* Journal of Psychiatric Research, vol. 41, no. 1-2, pages 15–30, 2007. (Cited on pages 27 and 29.)

[Kunz 2011] N. Kunz, H. Zhang, K.R. O'Brien, Y. Assaf, D.C. Alexander, F. Lazeyras and P.S. Huppi. *Probing micro-structural information using the CHARMED model in the non-myelinated human newborn brain at 3T.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 19, page 2489, 2011. (Cited on pages xiii, xiv, 35, 43, 46 and 55.)

[Kwong 1991] KK Kwong, RC McKinstry, D Chien, AP Crawley, JD Pearlman and BR Rosen. *CSF-suppressed quantitative single-shot diffusion imaging.* Magnetic Resonance in Medicine, vol. 21, no. 1, pages 157–163, 1991. (Cited on page 31.)

[Larsson 1992] H.B.W. Larsson, C. Thomsen, J. Frederiksen, M. Stubgaard and O. Henriksen. *In vivo magnetic resonance diffusion measurement in the brain of patients with multiple sclerosis.* Magnetic Resonance Imaging, vol. 10, no. 1, pages 7–12, 1992. (Cited on page 2.)

[Latour 2002] Lawrence L Latour and Steven Warach. *Cerebral spinal fluid contamination of the measurement of the apparent diffusion coefficient of water in acute stroke.* Magnetic Resonance in Medicine, vol. 48, no. 3, pages 478–86, 2002. (Cited on page 51.)

[Lauterbur 1973] Paul C Lauterbur. *Image formation by induced local interactions: examples employing nuclear magnetic resonance.* Nature, vol. 242, no. 5394, pages 190–191, 1973. (Cited on pages viii and 2.)

[Lazar 2003] Mariana Lazar, David M Weinstein, Jay S Tsuruda, Khader M Hasan, Konstantinos Arfanakis, M Elizabeth Meyerand, Benham Badie, Howard a

Rowley, Victor Haughton, Aaron Field and Andrew L Alexander. *White matter tractography using diffusion tensor deflection.* Human Brain Mapping, vol. 18, no. 4, pages 306–21, 2003. (Cited on page 110.)

[Lazar 2010] Mariana Lazar. *Mapping brain anatomical connectivity using white matter tractography.* NMR in Biomedicine, vol. 23, no. 7, pages 821–35, 2010. (Cited on page 170.)

[Le Bihan 1985] Denis Le Bihan and E. Breton. *Imagerie de diffusion in-vivo par résonance magnétique nucléaire.* Comptes-Rendus de l'Académie des Sciences, vol. 93, no. 5, pages 27–34, 1985. (Cited on pages viii and 2.)

[Le Bihan 1991] D. Le Bihan. *Molecular diffusion nuclear magnetic resonance imaging.* Magnetic Resonance Quarterly, vol. 7, no. 1, pages 1–30, 1991. (Cited on pages xii and 19.)

[Le Bihan 2001] D. Le Bihan, J.F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko and H. Chabriat. *Diffusion tensor imaging: concepts and applications.* Journal of Magnetic Resonance Imaging, vol. 13, no. 4, pages 534–46, 2001. (Cited on pages 24, 25, 26, 38 and 195.)

[Le Bihan 2003] D. Le Bihan. *Looking into the functional architecture of the brain with diffusion MRI.* Nature reviews. Neuroscience, vol. 4, no. 6, pages 469–80, 2003. (Cited on pages 2 and 117.)

[Le Bihan 2006] Denis Le Bihan. *Looking into the functional architecture of the brain with diffusion MRI.* International Congress Series, vol. 1290, pages 1–24, 2006. (Cited on page 19.)

[Lenglet 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 and P.M. Thompson. *Mathematical methods for diffusion MRI processing.* NeuroImage, vol. 45, no. 1 Suppl, pages S111–22, 2009. (Cited on pages 118 and 119.)

[Levenberg 1944] K. Levenberg. *A method for the solution of certain non-linear problems in least squares.* Quarterly Journal of Applied Mathematics, vol. 2, no. 2, pages 164–8, 1944. (Cited on page 75.)

[Lifshits 2009] Shlomi Lifshits, Arie Tamir and Yaniv Assaf. *Combinatorial fiber-tracking of the human brain.* NeuroImage, vol. 48, no. 3, pages 532–40, 2009. (Cited on page 112.)

[Lim 2002] K O Lim and J a Helpern. *Neuropsychiatric applications of DTI - a review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 587–93, 2002. (Cited on page 27.)

[Lin 2007] F Lin, C Yu and T Jiang. *Diffusion Tensor Tractography-Based Group Mapping of the Pyramidal Tract in Relapsing-Remitting Multiple Sclerosis Patients.* AJNR. American Journal of Neuroradiology, vol. 28, pages 278–82, 2007. (Cited on page 118.)

[Liu 2004] C. Liu, R. Bammer, B. Acar and M.E. Moseley. *Characterizing non-Gaussian diffusion by using generalized diffusion tensors.* Magnetic Resonance in Medicine, vol. 51, no. 5, pages 924–37, 2004. (Cited on pages 33 and 38.)

[Liu 2013] M. Liu, B.C. Vemuri and R. Deriche. *A robust variational approach for simultaneous smoothing and estimation of DTI.* NeuroImage, vol. 67, pages 33–41, 2013. (Cited on page 24.)

[Lori 2002] N F Lori, E Akbudak, J S Shimony, T S Cull, a Z Snyder, R K Guillory and T E Conturo. *Diffusion tensor fiber tracking of human brain connectivity: aquisition methods, reliability analysis and biological results.* NMR in Biomedicine, vol. 15, no. 7-8, pages 494–515, 2002. (Cited on page 110.)

[Lu 2006] Hanzhang Lu, Jens H Jensen, Anita Ramani and Joseph a Helpern. *Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging.* NMR in Biomedicine, vol. 19, no. 2, pages 236–47, 2006. (Cited on pages 39 and 40.)

[Ludeman 2008] N a Ludeman, J I Berman, Y W Wu, R J Jeremy, J Kornak, a I Bartha, a J Barkovich, D M Ferriero, R G Henry and O a Glenn. *Diffusion tensor imaging of the pyramidal tracts in infants with motor dysfunction.* Neurology, vol. 71, no. 21, pages 1676–82, 2008. (Cited on page 118.)

[Maes 1997] F Maes, a Collignon, D Vandermeulen, G Marchal and P Suetens. *Multimodality image registration by maximization of mutual information.* Transactions on Medical Imaging, vol. 16, no. 2, pages 187–98, 1997. (Cited on page 123.)

[Maier 2004] Stephan E Maier, Sridhar Vajapeyam, Hatsuho Mamata, Carl-Fredrik Westin, Ferenc a Jolesz and Robert V Mulkern. *Biexponential diffusion tensor analysis of human brain diffusion data.* Magnetic Resonance in Medicine, vol. 51, no. 2, pages 321–30, 2004. (Cited on page 50.)

[Malcolm 2010] J.G. Malcolm, O. Michailovich, S. Bouix, C.F. Westin, M.E. Shenton and Y. Rathi. *A filtered approach to neural tractography using the Watson directional function.* Medical Image Analysis, vol. 14, no. 1, pages 58–69, 2010. (Cited on page 113.)

[Mangin 2002] J-F Mangin, C Poupon, Y Cointepas, D Rivière, D Papadopoulos-Orfanos, C a Clark, J Régis and D Le Bihan. *A framework based on spin glass models for the inference of anatomical connectivity from diffusion-weighted*

MR data - a technical review. NMR in Biomedicine, vol. 15, no. 7-8, pages 481–92, 2002. (Cited on pages xix and 112.)

[Mansfield 1977] Peter Mansfield. *Multi-planar image formation using NMR spin echoes.* Journal of Physics C: Solid State Physics, vol. 10, no. 3, page L55, 1977. (Cited on pages viii and 2.)

[Mardia 1972] K.V. Mardia. Statistics of directional data. Academic Press, 1972. (Cited on pages 47 and 180.)

[Mattiello 1994] J Mattiello, PJ Basser and D LeBihan. *Analytical expressions for the b matrix in NMR diffusion imaging and spectroscopy.* Journal of Magnetic Resonance. Series B, vol. 108, pages 131–141, 1994. (Cited on page 22.)

[Mayzel-Oreg 2007] O. Mayzel-Oreg, Y. Assaf, A. Gigi, D. Ben-Bashat, R. Verchovsky, M. Mordohovitch, M. Graif, T. Hendler, A.D. Korczyn and Y. Cohen. *High b-value diffusion imaging of dementia: application to vascular dementia and alzheimer disease.* Journal of the Neurological Sciences, vol. 257, no. 1-2, pages 105–13, 2007. (Cited on page 44.)

[Merboldt 1985] Klaus-Dietmar Merboldt, Wolfgang Hanicke and Jens Frahm. *Self-diffusion NMR imaging using stimulated echoes.* Journal of Magnetic Resonance, vol. 64, no. 3, pages 479–486, 1985. (Cited on pages viii and 2.)

[Merlet 2010] S. Merlet and R. Deriche. *Compressed sensing for accelerated EAP recovery in diffusion MRI.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pages 14–21, 2010. (Cited on pages 33 and 164.)

[Merlet 2011] S. Merlet, J. Cheng, A. Ghosh and R. Deriche. *Spherical Polar Fourier EAP and ODF reconstruction via compressed sensing in diffusion MRI.* In Proceedings of the International Symposium on Biomedical Imaging (ISBI), pages 365–371, 2011. (Cited on pages 33 and 164.)

[Merlet 2012] S. Merlet, M. Paquette, R. Deriche and M. Descoteaux. *Ensemble average propagator reconstruction via compressed sensing: discrete or continuous bases?* In Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), volume 20, page 2277, 2012. (Cited on pages 33 and 164.)

[Merlet 2013] S. Merlet, E. Caruyer, A. Ghosh and R. Deriche. *A computational diffusion MRI and parametric dictionary learning framework for modeling the diffusion signal and its features.* Medical Image Analysis, vol. 17, no. 7, pages 830–843, 2013. (Cited on pages 33 and 164.)

[Metzler-Baddeley 2012] Claudia Metzler-Baddeley, Michael J O'Sullivan, Sonya Bells, Ofer Pasternak and Derek K Jones. *How and how not to correct*

for CSF-contamination in diffusion MRI. NeuroImage, vol. 59, no. 2, pages 1394–403, 2012. (Cited on pages 34 and 51.)

[Minati 2007] Ludovico Minati, Domenico Aquino, Stefano Rampoldi, Sergio Papa, Marina Grisoli, Maria Grazia Bruzzone and Elio Maccagnano. *Biexponential and diffusional kurtosis imaging, and generalised diffusion-tensor imaging (GDTI) with rank-4 tensors: a study in a group of healthy subjects.* Magma, vol. 20, no. 5-6, pages 241–53, 2007. (Cited on page 40.)

[Monk 2006] P.N. Monk and P.J. Shaw. *ALS: life and death in a bad neighborhood.* Nature Medicine, vol. 12, no. 8, pages 885–7, 2006. (Cited on pages 8 and 189.)

[Mori 1999] S. Mori, B. J. Crain, V. P. Chacko and P. C. van Zijl. *Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging.* Annals of Neurology, vol. 45, no. 2, pages 265–269, 1999. (Cited on pages xix, 109, 120, 145, 148 and 168.)

[Mori 2002] Susumu Mori and Peter C M van Zijl. *Fiber tracking: principles and strategies - a technical review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 468–80, 2002. (Cited on pages 117 and 118.)

[Mori 2005] Susumu Mori, Setsu Wakana, Peter CM Van Zijl and LM Nagae-Poetscher. Mri atlas of human white matter. Am Soc Neuroradiology, 2005. (Cited on page 117.)

[Mori 2008] Susumu Mori, Kenichi Oishi, Hangyi Jiang, Li Jiang, Xin Li, Kazi Akhter, Kegang Hua, Andreia V Faria, Asif Mahmood, Roger Woods, Arthur W Toga, G Bruce Pike, Pedro Rosa Neto, Alan Evans, Jiangyang Zhang, Hao Huang, Michael I Miller, Peter van Zijl and John Mazziotta. *Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template.* NeuroImage, vol. 40, no. 2, pages 570–82, 2008. (Cited on page 171.)

[Moseley 1990] ME Moseley, Y Cohen, J Kucharczyk, J Mintorovitch, HS Asgari, MF Wendland, J Tsuruda and D Norman. *Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system.* Radiology, vol. 176, no. 2, pages 439–445, 1990. (Cited on pages viii, 2 and 11.)

[Moseley 2002] Michael Moseley. *Diffusion tensor imaging and aging - a review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 553–60, 2002. (Cited on page 25.)

[Mukherjee 2008a] P Mukherjee, J I Berman, S W Chung, C P Hess and R G Henry. *Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings.* AJNR. American Journal of Neuroradiology, vol. 29, no. 4, pages 632–41, 2008. (Cited on page 118.)

[Mukherjee 2008b] P Mukherjee, S W Chung, J I Berman, C P Hess and R G Henry. *Diffusion tensor MR imaging and fiber tractography: technical considerations.* AJNR. American Journal of Neuroradiology, vol. 29, no. 5, pages 843–52, 2008. (Cited on page 118.)

[Mulkern 2000] R V Mulkern, H P Zengingonul, R L Robertson, P Bogner, K H Zou, H Gudbjartsson, C R Guttmann, D Holtzman, W Kyriakos, F a Jolesz and S E Maier. *Multi-component apparent diffusion coefficients in human brain: relationship to spin-lattice relaxation.* Magnetic Resonance in Medicine, vol. 44, no. 2, pages 292–300, 2000. (Cited on page 50.)

[Mulkern 2009] Robert V Mulkern, Steven J Haker and Stephan E Maier. *On high b diffusion imaging in the human brain: ruminations and experimental insights.* Magnetic Resonance Imaging, vol. 27, no. 8, pages 1151–62, 2009. (Cited on page 50.)

[Neil 2002] J Neil, J Miller, P Mukherjee and P S Hüppi. *Diffusion tensor imaging of normal and injured developing human brain - a technical review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 543–52, 2002. (Cited on page 25.)

[Netsch 2004] Thomas Netsch and Arianne van Muiswinkel. *Quantitative evaluation of image-based distortion correction in diffusion tensor imaging.* Transactions on Medical Imaging, vol. 23, no. 7, pages 789–98, 2004. (Cited on page 122.)

[Neuman 1974] C.H. Neuman. *Spin echo of spins diffusing in a bounded medium.* Journal of Chemical Physics, vol. 60, no. 11, page 4508, 1974. (Cited on pages xiv, 45, 46, 55 and 165.)

[Niogi 2010] Sumit N Niogi and Pratik Mukherjee. *Diffusion tensor imaging of mild traumatic brain injury.* The Journal of Head Trauma Rehabilitation, vol. 25, no. 4, pages 241–255, 2010. (Cited on pages 28 and 31.)

[Norris 2001] David G Norris. *The effects of microscopic tissue parameters on the diffusion weighted magnetic resonance imaging experiment.* NMR in Biomedicine, vol. 14, pages 77–93, 2001. (Cited on page 50.)

[O'Donnell 2002] Lauren O'Donnell, Steven Haker and Carl-fredrik Westin. *New Approaches to Estimation of White Matter Connectivity in Diffusion Tensor MRI : Elliptic PDEs and Geodesics in a Tensor-Warped Space.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 459–466, 2002. (Cited on pages xix and 111.)

[Oldfield 1971] Richard C Oldfield. *The assessment and analysis of handedness: the Edinburgh inventory.* Neuropsychologia, vol. 9, no. 1, pages 97–113, 1971. (Cited on page 122.)

[Ourselin 2000] S. Ourselin, A. Roche, S. Prima and N. Ayache. *Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 557–66, 2000. (Cited on page 144.)

[Ozarslan 2003] E. Ozarslan and T.H. Mareci. *Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging.* Magnetic Resonance in Medicine, vol. 50, no. 5, pages 955–65, 2003. (Cited on page 33.)

[Ozarslan 2005] Evren Ozarslan, Baba C Vemuri and Thomas H Mareci. *Generalized scalar measures for diffusion MRI using trace, variance, and entropy.* Magnetic Resonance in Medicine, vol. 53, no. 4, pages 866–76, 2005. (Cited on page 34.)

[Ozarslan 2006] Evren Ozarslan, Timothy M Shepherd, Baba C Vemuri, Stephen J Blackband and Thomas H Mareci. *Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT).* NeuroImage, vol. 31, no. 3, pages 1086–103, 2006. (Cited on pages 31, 33, 50 and 51.)

[Panagiotaki 2012] Eleftheria Panagiotaki, Torben Schneider, Bernard Siow, Matt G Hall, Mark F Lythgoe and Daniel C Alexander. *Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison.* NeuroImage, vol. 59, no. 3, pages 2241–54, 2012. (Cited on pages 45, 74 and 75.)

[Papadakis 1999] N G Papadakis, D Xing, C L Huang, L D Hall and T a Carpenter. *A comparative study of acquisition schemes for diffusion tensor imaging using MRI.* Journal of Magnetic Resonance, vol. 137, no. 1, pages 67–82, 1999. (Cited on page 22.)

[Papadakis 2000] N G Papadakis, C D Murrills, L D Hall, C L Huang and T Adrian Carpenter. *Minimal gradient encoding for robust estimation of diffusion anisotropy.* Magnetic Resonance Imaging, vol. 18, no. 6, pages 671–9, 2000. (Cited on page 22.)

[Parker 2002] Geoffrey J M Parker, Claudia A M Wheeler-Kingshott and Gareth J Barker. *Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging.* Transactions on Medical Imaging, vol. 21, no. 5, pages 505–12, 2002. (Cited on pages xix and 111.)

[Parker 2003a] Geoff J M Parker and Daniel C Alexander. *Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information.* In Proceedings of the International Conference on Information Processing in Medical Imaging, volume 18, pages 684–95, 2003. (Cited on pages xix and 115.)

[Parker 2003b] Geoffrey J M Parker, Hamied a Haroon and Claudia a M Wheeler-Kingshott. *A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements.* Journal of Magnetic Resonance Imaging, vol. 18, no. 2, pages 242–54, 2003. (Cited on pages xix and 115.)

[Parker 2005] Geoffrey J M Parker and Daniel C Alexander. *Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue.* Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, vol. 360, no. 1457, pages 893–902, 2005. (Cited on pages xix and 115.)

[Pasternak 2009] Ofer Pasternak, Nir Sochen, Yaniv Gur, Nathan Intrator and Yaniv Assaf. *Free water elimination and mapping from diffusion MRI.* Magnetic Resonance in Medicine, vol. 62, no. 3, pages 717–30, 2009. (Cited on page 34.)

[Peled 2009] Sharon Peled, Stephen Whalen, Ferenc a Jolesz and Alexandra J Golby. *High b-value apparent diffusion-weighted images from CURVE-ball DTI.* Journal of Magnetic Resonance Imaging, vol. 30, no. 1, pages 243–8, 2009. (Cited on page 23.)

[Penfield 1937] W. Penfield and E. Boldrey. *Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation.* Brain, vol. 60, no. 4, pages 389–443, 1937. (Cited on page 142.)

[Pfefferbaum 2000] Adolf Pfefferbaum, Edith V Sullivan, Maj Hedehus, Elfar Adalsteinsson, Kelvin O Lim and Michael Moseley. *In vivo detection and functional correlates of white matter microstructural disruption in chronic alcoholism.* Alcoholism: Clinical and Experimental Research, vol. 24, no. 8, pages 1214–1221, 2000. (Cited on page 31.)

[Pfeuffer 1998] J Pfeuffer, U Flögel, W Dreher and D Leibfritz. *Restricted diffusion and exchange of intracellular water: theoretical modelling and diffusion time dependence of 1H NMR measurements on perfused glial cells.* NMR in Biomedicine, vol. 11, no. 1, pages 19–31, 1998. (Cited on pages ix and 41.)

[Pichon 2005] E Pichon, CF Westin and AR Tannenbaum. *A Hamilton-Jacobi-Bellman Approach to High Angular Resolution Diffusion Tractography.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 8, pages 180–187, 2005. (Cited on page 111.)

[Pierpaoli 1996a] C. Pierpaoli and P.J. Basser. *Toward a quantitative assessment of diffusion anisotropy.* Magnetic Resonance Medicine, vol. 36, no. 6, pages 893–906, 1996. (Cited on pages xiii, 20 and 31.)

[Pierpaoli 1996b] C. Pierpaoli, P. Jezzard, P.J. Basser, A. Barnett and G. Di Chiro. *Diffusion tensor MR imaging of the human brain.* Radiology, vol. 201, no. 3, pages 637–48, 1996. (Cited on pages xvi, 23, 24, 38 and 62.)

[Pipe 2009] J. Pipe. *Pulse sequences for diffusion-weighted MRI.* In H. Johansen-Berg and T.E.J. Behrens, editeurs, Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy., pages 205–236. Academic Press, 2009. (Cited on pages 11, 15, 16 and 190.)

[Pontabry 2013] J. Pontabry, F. Rousseau, E. Oubel, C. Studholme, M. Koob and J.L. Dietemann. *Probabilistic tractography using Q-ball imaging and particle filtering: application to adult and in-utero fetal brain studies.* Medical Image Analysis, vol. 17, no. 3, pages 297–310, 2013. (Cited on pages xix, xx, 116, 117, 126, 127, 129, 131, 138 and 168.)

[Posada 2004] David Posada and Thomas R Buckley. *Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests.* Systematic Biology, vol. 53, no. 5, pages 793–808, 2004. (Cited on pages xviii, 92 and 93.)

[Poupon 2001] C. Poupon, J.F. Mangin, C.A. Clark, V. Frouin, J. Régis, D. Le Bihan and I. Bloch. *Towards inference of human brain connectivity from MR diffusion tensor data.* Med. Image Anal., vol. 5, no. 1, pages 1–15, 2001. (Cited on page xviii.)

[Powell 1994] M.J.D. Powell. *A direct search optimization method that models the objective and constraint functions by linear interpolation.* In Susana Gomez and Jean-Pierre Hennart, editeurs, Advances in Optimization and Numerical Analysis, pages 51–67. Springer Netherlands, 1994. (Cited on pages 74, 150 and 161.)

[Powell 2006] M.J.D. Powell. *The NEWUOA software for unconstrained optimization without derivatives.* In G. Pillo and M. Roma, editeurs, Large-Scale Nonlinear Optimization, volume 83 of *Nonconvex Optimization and Its Applications*, pages 255–297. Kluwer Academic Publishers, 2006. (Cited on pages 119, 145 and 161.)

[Powell 2009] M.J.D. Powell. *The BOBYQA algorithm for bound constrained optimization without derivatives.* Rapport technique, University of Cambridge, 2009. (Cited on pages 148 and 161.)

[Prados 2006] Emmanuel Prados, Stefano Soatto and Christophe Lenglet. *Control Theory and Fast Marching Techniques for Brain Connectivity Mapping.* In Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, pages 1076–1083, 2006. (Cited on pages xix and 111.)

[Purcell 1946] Edward M Purcell, HC Torrey and Robert V Pound. *Resonance absorption by nuclear magnetic moments in a solid.* Physical Review, vol. 69, no. 1-2, page 37, 1946. (Cited on page 2.)

[Raftery 1992] Adrian E Raftery. *Bayesian Model Selection in Structural Equation Models.* Rapport technique, 1992. (Cited on page 88.)

[Rajan 2012] J. Rajan, J. Veraart, J. Van Audekerke, M. Verhoye and J. Sijbers. *Nonlocal maximum likelihood estimation method for denoising multiple-coil magnetic resonance images.* Magnetic Resonance Imaging, 2012. (Cited on page 170.)

[Ramani 2007] A Ramani, JH Jensen, KU Szulc, O Ali, C Hu, H Lu, JD Brodie and JA Helpern. *Assessment of abnormalities in the cerebral microstructure of schizophrenia patients : a diffusional kurtosis imaging study.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 15, page 648, 2007. (Cited on page 40.)

[Reisert 2011] Marco Reisert, Irina Mader, Constantin Anastasopoulos, Matthias Weigel, Susanne Schnell and Valerij Kiselev. *Global fiber reconstruction becomes practical.* NeuroImage, vol. 54, no. 2, pages 955–62, 2011. (Cited on pages xix and 112.)

[Rockafellar 1998] R Tyrrell Rockafellar and Roger J-B Wets. Variational analysis: Grundlehren der mathematischen wissenschaften, volume 317. Springer, 1998. (Cited on page 129.)

[Rontgen 1896] WC Rontgen. *On a new kind of rays.* Nature, vol. 53, no. 1369, pages 274–276, 1896. (Cited on page 1.)

[Rowe 2013] MC Rowe, H Zhang, N Oxtoby and DC Alexander. *Beyond Crossing Fibers: Tractography Exploiting Sub-voxel Fibre Dispersion and Neighbourhood Structure.* pages 402–413, 2013. (Cited on page 113.)

[Salat 2009] D.H. Salat, S.Y. Lee, P. Yu, B. Setty, H.D. Rosas and P.E. Grant. *DTI in development and aging.* In H. Johansen-Berg and T.E.J. Behrens, editeurs, Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy., pages 205–236. Academic Press, 2009. (Cited on page 31.)

[Salvador 2005] Raymond Salvador, Alonso Peña, David K Menon, T Adrian Carpenter, John D Pickard and Ed T Bullmore. *Formal characterization and extension of the linearized diffusion tensor model.* Human Brain Mapping, vol. 24, no. 2, pages 144–55, 2005. (Cited on page 130.)

[Savadjiev 2010] Peter Savadjiev, Yogesh Rathi, James G Malcolm, Martha E Shenton and Carl-Fredrik Westin. *A geometry-based particle filtering approach to white matter tractography.* In Proceedings of the International Conference on

Medical Image Computing and Computer-Assisted Intervention, volume 13, pages 233–40, 2010. (Cited on page 113.)

[Scherrer 2010] B. Scherrer and S.K. Warfield. *Why multiple b-values are required for multi-tensor models. Evaluation with a constrained log-Euclidean model.* In Proceedings of the International Symposium on Biomedical Imaging, pages 1389–92, 2010. (Cited on pages xiv, 118 and 156.)

[Scherrer 2012] B. Scherrer and S.K. Warfield. *Parametric representation of multiple white matter fascicles from cube and sphere diffusion MRI.* PloS one, vol. 7, no. 11, page e48232, 2012. (Cited on pages xviii, 17, 23, 34, 50, 51, 59, 75, 88, 164, 167 and 169.)

[Scherrer 2013] Benoit Scherrer, Maxime Taquet and Simon K Warfield. *Reliable Selection of the Number of Fascicles in Diffusion Images by Estimation of the Generalization Error.* In Proceedings of the International Conference on Information Processing in Medical Imaging, volume 7917, pages 742–753, 2013. (Cited on pages xviii, 88, 89 and 167.)

[Schultz 2010] Thomas Schultz, Carl-Fredrik Westin and Gordon Kindlmann. *Multi-diffusion-tensor fitting via spherical deconvolution: a unifying framework.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 13, pages 674–81, 2010. (Cited on pages xviii and 88.)

[Schwarcz 2004] Attila Schwarcz, Peter Bogner, Philippe Meric, Jean-Loup Correze, Zoltan Berente, József Pál, Ferenc Gallyas, Tamas Doczi, Brigitte Gillet and Jean-Claude Beloeil. *The existence of biexponential signal decay in magnetic resonance diffusion-weighted imaging appears to be independent of compartmentalization.* Magnetic Resonance in Medicine, vol. 51, no. 2, pages 278–85, 2004. (Cited on page 50.)

[Schwarz 1978] G. Schwarz. *Estimating the Dimension of a Model.* The Annals of Statistics, vol. 6, no. 2, pages 461–464, 1978. (Cited on page 93.)

[Sehy 2004] Jonathan V Sehy, Lin Zhao, Junqian Xu, Heidi J Rayala, Joseph J H Ackerman and Jeffrey J Neil. *Effects of physiologic challenge on the ADC of intracellular water in the Xenopus oocyte.* Magnetic Resonance in Medicine, vol. 52, no. 2, pages 239–47, 2004. (Cited on page 50.)

[Seo 2012] Jeong Pyo Seo, Pyung-Hun Chang and Sung Ho Jang. *Anatomical location of the corticospinal tract according to somatotopies in the centrum semiovale.* Neuroscience Letters, vol. 523, no. 2, pages 111–4, 2012. (Cited on page 117.)

[Sexton 2009] Claire E Sexton, Clare E Mackay and Klaus P Ebmeier. *A systematic review of diffusion tensor imaging studies in affective disorders.* Biological Psychiatry, vol. 66, no. 9, pages 814–23, 2009. (Cited on pages xii, 27 and 29.)

[Shenton 1995] M.E. Shenton, R Kikinis, W. McCarley, P. Saiviroonporn, H.H. Hokama, A Robatino, D. Metcalf, C.G. Wible, C.M. Portas, D.V. Iosifescu, R. Donnino, J.M. Goldstein and F.A. Jolesz. *Harvard brain atlas: a teaching and visualization tool.* In Proceedings of the International Conference on Biomedical Visualization, volume 10, pages 10–17, 1995. (Cited on page 171.)

[Shenton 2012] M E Shenton, H M Hamoda, J S Schneiderman, S Bouix, O Pasternak, Y Rathi, M-a Vu, M P Purohit, K Helmer, I Koerte, a P Lin, C-F Westin, R Kikinis, M Kubicki, R a Stern and R Zafonte. *A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury.* Brain Imaging and Behavior, vol. 6, no. 2, pages 137–92, 2012. (Cited on pages xiii, 28 and 31.)

[Shiga 2004] Y. Shiga, K. Miyazawa, S. Sato, R. Fukushima, S. Shibuya, Y. Sato, H. Konno, K. Doh-ura, S. Mugikura, H. Tamura, S. Higano, S. Takahashi and Y. Itoyama. *Diffusion-weighted MRI abnormalities as an early diagnostic marker for Creutzfeldt-Jakob disease.* Neurology, vol. 63, no. 3, pages 443–9, 2004. (Cited on page 2.)

[Smith 2002] Stephen M Smith. *Fast robust automated brain extraction.* Human Brain Mapping, vol. 17, no. 3, pages 143–55, 2002. (Cited on page 120.)

[Song 2002] Sheng-Kwei Song, Shu-Wei Sun, Michael J. Ramsbottom, Chen Chang, John Russell and Anne H. Cross. *Dysmyelination Revealed through MRI as Increased Radial (but Unchanged Axial) Diffusion of Water.* NeuroImage, vol. 17, no. 3, pages 1429–1436, 2002. (Cited on page 32.)

[Song 2003] Sheng-Kwei Song, Shu-Wei Sun, Won-Kyu Ju, Shiow-Jiuan Lin, Anne H Cross and Arthur H Neufeld. *Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia.* NeuroImage, vol. 20, no. 3, pages 1714–1722, 2003. (Cited on page 32.)

[Sotak 2002] Christopher H Sotak. *The role of diffusion tensor imaging in the evaluation of ischemic brain injury - a review.* NMR in Biomedicine, vol. 15, no. 7-8, pages 561–9, 2002. (Cited on page 25.)

[Sotiropoulos 2012] S.N. Sotiropoulos, T.E.J. Behrens and S. Jbabdi. *Ball and rackets: Inferring fiber fanning from diffusion-weighted MRI.* NeuroImage, vol. 60, no. 2, pages 1412–25, 2012. (Cited on pages 47 and 50.)

[Sperl 2012] J.I. Sperl, M.I. Menzel, E.T. Tan, K. Khare, K.F. King, C.J. Hardy and L. Marinelli. *Effects of compressed sensing reconstruction on kurtosis tensor fitting in diffusion spectrum imaging.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 20, page 1888, 2012. (Cited on page 39.)

[Staempfli 2006] P Staempfli, T Jaermann, G R Crelier, S Kollias, a Valavanis and P Boesiger. *Resolving fiber crossing using advanced fast marching tractography based on diffusion tensor imaging.* NeuroImage, vol. 30, no. 1, pages 110–20, 2006. (Cited on pages xix and 111.)

[Stanisz 1997] G.J. Stanisz, A. Szafer, G.A. Wright and R.M. Henkelman. *An analytical model of restricted diffusion in bovine optic nerve.* Magnetic Resonance in Medicine, vol. 37, no. 1, pages 103–11, 1997. (Cited on pages 41, 45, 55 and 69.)

[Stejskal 1965a] E.O. Stejskal. *Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow.* Journal of Chemical Physics, vol. 43, no. 10, pages 3597–603, 1965. (Cited on page 19.)

[Stejskal 1965b] E.O. Stejskal and J.E. Tanner. *Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient.* Journal of Chemical Physics, vol. 42, no. 1, page 288, 1965. (Cited on pages xii, xv, 15, 19, 20 and 51.)

[Stieltjes 2001] B Stieltjes, W E Kaufmann, P C van Zijl, K Fredericksen, G D Pearlson, M Solaiyappan and S Mori. *Diffusion tensor imaging and axonal tracking in the human brainstem.* NeuroImage, vol. 14, no. 3, pages 723–35, 2001. (Cited on page 110.)

[Sundgren 2004] P C Sundgren, Q Dong, D Gómez-Hassan, S K Mukherji, P Maly and R Welsh. *Diffusion tensor imaging of the brain: review of clinical applications.* Neuroradiology, vol. 46, no. 5, pages 339–50, 2004. (Cited on pages xii, 25, 26, 27 and 28.)

[Taquet 2012a] M. Taquet, B. Scherrer, C. Benjamin, S. Prabhu, B. Macq and S.K. Warfield. *Interpolating multi-fiber models by Gaussian mixture simplification.* In Proceedings of the International Symposium on Biomedical Imaging (ISBI), pages 928–931, 2012. (Cited on page 171.)

[Taquet 2012b] M. Taquet, B. Scherrer, O. Commowick, J. Peters, M. Sahin, B. Macq and S.K. Warfield. *Registration and Analysis of White Matter Group Differences with a Multi-Fiber Model.* In Nicholas Ayache, Hervé Delingette, Polina Golland and Kensaku Mori, editeurs, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), volume 7512 of *Lecture Notes in Computer Science*, pages 313–320, 2012. (Cited on page 170.)

[Tariq 2012] M. Tariq, T. Schneider, D.C. Alexander, C.A.M. Wheeler-Kingshott and H. Zhang. *Scan-rescan reproducibility of neurite microstructure estimates using NODDI.* Medical Image Understanding and Analysis, pages 255–61, 2012. (Cited on page 50.)

[Taylor 1985] DG Taylor and MC Bushell. *The spatial mapping of translational diffusion coefficients by the NMR imaging technique.* Physics in Medicine and Biology, vol. 30, no. 4, page 345, 1985. (Cited on pages viii and 2.)

[Thiebaut de Schotten 2011] Michel Thiebaut de Schotten, Dominic H Ffytche, Alberto Bizzi, Flavio Dell'Acqua, Matthew Allin, Muriel Walshe, Robin Murray, Steven C Williams, Declan G M Murphy and Marco Catani. *Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography.* NeuroImage, vol. 54, no. 1, pages 49–59, 2011. (Cited on pages x, 9, 28, 92, 117, 142 and 164.)

[Toga 2003] Arthur W Toga and Paul M Thompson. *Mapping brain asymmetry.* Nature Reviews. Neuroscience, vol. 4, no. 1, pages 37–48, 2003. (Cited on pages 118 and 126.)

[Torrey 1956] Henry C Torrey. *Bloch equations with diffusion terms.* Physical Review, vol. 104, no. 3, page 563, 1956. (Cited on page 18.)

[Tournier 2003] J.-Donald Tournier, Fernando Calamante, David G Gadian and Alan Connelly. *Diffusion-weighted magnetic resonance imaging fibre tracking using a front evolution algorithm.* NeuroImage, vol. 20, no. 1, pages 276–288, 2003. (Cited on page 111.)

[Tournier 2004] J.D. Tournier, F. Calamante, D.G. Gadian and A. Connelly. *Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution.* NeuroImage, vol. 23, no. 3, pages 1176–85, 2004. (Cited on page 33.)

[Tournier 2007] J.D. Tournier, F. Calamante and A. Connelly. *Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution.* NeuroImage, vol. 35, no. 4, pages 1459–72, 2007. (Cited on page 33.)

[Tournier 2008] J.D. Tournier, C.H. Yeh, F. Calamante, K.H. Cho, A. Connelly and C.P. Lin. *Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data.* NeuroImage, vol. 42, no. 2, pages 617–25, 2008. (Cited on page 115.)

[Tuch 1999] D.S. Tuch, R.M. Weisskoff, J.W. Belliveau and V.J. Wedeen. *High Angular Resolution Diffusion Imaging of the human brain.* In Proceedings of the International Society for Magnetic Resonance in Medicine, page 321, 1999. (Cited on pages 66 and 164.)

[Tuch 2002] D.S. Tuch, T.G. Reese, M.R. Wiegell, N. Makris, J.W. Belliveau and V.J. Wedeen. *High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity.* Magnetic Resonance in Medicine, vol. 48, pages 577–82, 2002. (Cited on pages 44, 59, 66 and 164.)

[Tuch 2003] David S Tuch, Timothy G Reese, Mette R Wiegell and Van J Wedeen. *Diffusion MRI of complex neural architecture.* Neuron, vol. 40, no. 5, pages 885–95, 2003. (Cited on page 34.)

[Tuch 2004] D.S. Tuch. *Q-Ball Imaging.* Magnetic Resonance in Medicine, vol. 52, no. 6, pages 1358–72, 2004. (Cited on pages 32, 112, 116 and 119.)

[Ulrich 1984] G Ulrich. *Computer Generation of Distributions on the m-Sphere.* Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 33, no. 2, pages 158–163, 1984. (Cited on pages 116 and 132.)

[van Gelderen 1994] P. van Gelderen, D. DesPres, P.C. van Zijl and C.T. Moonen. *Evaluation of restricted diffusion in cylinders. Phosphocreatine in rabbit leg muscle.* Journal of Magnetic Resonance. Series B, vol. 103, no. 3, pages 255–60, 1994. (Cited on page xiv.)

[Veraart 2011a] J. Veraart, D.H.J. Poot, W. Van Hecke, I. Blockx, A. Van der Linden, M. Verhoye and J. Sijbers. *More accurate estimation of diffusion tensor parameters using diffusion Kurtosis imaging.* Magnetic Resonance in Medicine, vol. 65, no. 1, pages 138–45, 2011. (Cited on page 40.)

[Veraart 2011b] J. Veraart, W. Van Hecke and J. Sijbers. *Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model.* Magnetic Resonance in Medicine, vol. 66, no. 3, pages 678–86, 2011. (Cited on pages xvi, 39, 40, 62 and 103.)

[Vermaak 2003] Jaco Vermaak, Arnaud Doucet and Patrick Pérez. *Maintaining multimodality through mixture tracking.* In Proceedings of the International Conference on Computer Vision, volume 2, pages 1110–6, 2003. (Cited on pages xx, xxi, xxii, 126, 127, 128, 129 and 130.)

[Wedeen 2000] V.J. Wedeen, T.G. Reese, D.S. Tuch, M.R. Weigel, J.G. Dou, R.M. Weiskoff and D. Chessler. *Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI.* In Proceedings of the International Society for Magnetic Resonance in Medicine, volume 8, page 82, 2000. (Cited on pages 21, 32 and 164.)

[Wedeen 2005] V.J. Wedeen, P. Hagmann, W.Y.I. Tseng, T.G. Reese and R.M. Weisskoff. *Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging.* Magnetic Resonance in Medicine, vol. 54, no. 6, pages 1377–86, 2005. (Cited on pages 21, 32 and 164.)

[Weinstein 1999] D. Weinstein, G. Kindlmann and E. Lundberg. *Tensorlines: advection-diffusion based propagation through diffusion tensor fields.* In Proceedings of the Conference on Visualization, volume 3, pages 249–530, 1999. (Cited on page 110.)

[Westin 2002] CF Westin, S E Maier, H Mamata, a Nabavi, F a Jolesz and R Kikinis. *Processing and visualization for diffusion tensor MRI.* Medical Image Analysis, vol. 6, no. 2, pages 93–108, 2002. (Cited on pages xii, 21, 110, 111 and 116.)

[Wiest-Daesslé 2008] N. Wiest-Daesslé, S. Prima, P. Coupé, S.P. Morrissey and C. Barillot. *Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: applications to DT-MRI.* In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 171–9, 2008. (Cited on pages 96, 122, 144 and 170.)

[Wood 1994] Andrew T.A Wood. *Simulation of the von Mises Fisher distribution.* Communications in Statistics. Simulation and Computation, vol. 23, no. 1, pages 157–164, 1994. (Cited on page 116.)

[Wu 2009] Xi Wu, Qing Xu, Lei Xu, Jiliu Zhou, Adam W Anderson and Zhaohua Ding. *Genetic white matter fiber tractography with global optimization.* Journal of Neuroscience Methods, vol. 184, no. 2, pages 375–9, 2009. (Cited on page 112.)

[Wu 2010] Ed X Wu and Matthew M Cheung. *MR diffusion kurtosis imaging for neural tissue characterization.* NMR in Biomedicine, vol. 23, no. 7, pages 836–48, 2010. (Cited on page 60.)

[Xue 1999] R Xue, P C van Zijl, B J Crain, M Solaiyappan and S Mori. *In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging.* Magnetic Resonance in Medicine, vol. 42, no. 6, pages 1123–7, 1999. (Cited on pages xix and 109.)

[Yablonskiy 2003] Dmitriy a Yablonskiy, G Larry Bretthorst and Joseph J H Ackerman. *Statistical model for diffusion attenuated MR signal.* Magnetic Resonance in Medicine, vol. 50, no. 4, pages 664–9, 2003. (Cited on pages 38 and 50.)

[Yousry 1997] T.A. Yousry, U.D. Schmid, H. Alkadhi, D. Schmidt, A. Peraud, A. Buettner and P. Winkler. *Localization of the motor hand area to a knob on the precentral gyrus. A new landmark.* Brain: A Journal of Neurology, vol. 120, pages 141–57, 1997. (Cited on page 117.)

[Yu 2009] Chunshui Yu, Chaozhe Zhu, Yujin Zhang, Hai Chen, Wen Qin, Moli Wang and Kuncheng Li. *A longitudinal diffusion tensor imaging study on Wallerian degeneration of corticospinal tract after motor pathway stroke.* NeuroImage, vol. 47, no. 2, pages 451–8, 2009. (Cited on page 118.)

[Zhan 2006] Wang Zhan and Yihong Yang. *How accurately can the diffusion profiles indicate multiple fiber orientations? A study on general fiber crossings in diffusion MRI.* Journal of Magnetic Resonance, vol. 183, no. 2, pages 193–202, 2006. (Cited on page 32.)

[Zhang 2004] Song Zhang, Mark E Bastin, David H Laidlaw, Saurabh Sinha, Paul a Armitage and Thomas S Deisboeck. *Visualization and analysis of white matter structural asymmetry in diffusion tensor MRI data.* Magnetic Resonance in Medicine, vol. 51, no. 1, pages 140–7, 2004. (Cited on page 110.)

[Zhang 2009] Fan Zhang, Edwin R Hancock, Casey Goodlett and Guido Gerig. *Probabilistic white matter fiber tracking using particle filtering and von Mises-Fisher sampling.* Medical Image Analysis, vol. 13, no. 1, pages 5–18, 2009. (Cited on pages xix, xx, 116, 117, 126, 127, 129, 130, 131, 139 and 168.)

[Zhang 2011] H. Zhang, P.L. Hubbard, G.J.M. Parker and D.C. Alexander. *Axon diameter mapping in the presence of orientation dispersion with diffusion MRI.* NeuroImage, vol. 56, no. 3, pages 1301–15, 2011. (Cited on pages 47, 48, 49 and 50.)

[Zhang 2012] Hui Zhang, Torben Schneider, Claudia a Wheeler-Kingshott and Daniel C Alexander. *NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain.* NeuroImage, vol. 61, no. 4, pages 1000–16, 2012. (Cited on pages xiv, 41, 48, 49, 50, 51, 55, 56, 59, 80, 96, 165 and 170.)

Publications

Thesis-related papers

[1] **Aymeric Stamm**, Patrick Pérez, and Christian Barillot. Diffusion Directions Imaging. Technical report, INRIA, France, 2011. RR-7683. (Cited on page 166.)

[2] **Aymeric Stamm**, Christian Barillot, and Patrick Pérez. Method and MRI device to detect a direction of at least one fiber in a body, November 17 2011. US Patent App. 13/299,309. (Cited on page 166.)

[3] Nicolas Wiest-Daesslé, Olivier Commowick, **Aymeric Stamm**, Patrick Pérez, Christian Barillot, Romuald Seizeur, and Sylvain Prima. Comparison of 3 diffusion models to track the hand motor fibers within the corticospinal tract using functional, anatomical and diffusion MRI. In *MICCAI 2011 Workshop on Computational Diffusion MRI*, pages 150–157, Toronto (Canada), 2011. (Cited on page 169.)

[4] Olivier Commowick, **Aymeric Stamm**, Romuald Seizeur, Patrick Pérez, Christian Barillot, Sylvain Prima, and Nicolas Wiest-Daesslé. Multifiber deterministic streamline tractography of the corticospinal tract based on a new diffusion model. In *MICCAI 2011 DTI Tractography Challenge Workshop*, pages 18–24, Toronto (Canada), 2011. (Cited on page 169.)

[5] **Aymeric Stamm**, Christian Barillot, and Patrick Pérez. Diffusion MRI for detecting a direction of at least one fibre in a body, May 30 2012. EP Patent 2,458,397. (Cited on pages xv, 52 and 166.)

[6] **Aymeric Stamm**, Patrick Pérez, and Christian Barillot. A new multi-fiber model for low angular resolution diffusion MRI. In *Proceedings of the International Symposium on Biomedical Imaging (ISBI)*, pages 936–39, Barcelona (Spain), 2012. (Cited on pages xv, 52 and 166.)

[7] **Aymeric Stamm**, Patrick Pérez, and Christian Barillot. A new multi-directional fiber model for low angular resolution diffusion imaging. In *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*, volume 20, page 695, Melbourne (Australia), 2012. Magna Cum Laude Award. (Cited on page 166.)

[8] **Aymeric Stamm**, Olivier Commowick, Patrick Pérez, and Christian Barillot. Multifiber Deterministic Streamline Tractography of the Corticospinal Tract Based on a New Diffusion Model: Part II. In *MICCAI 2012 DTI Tractography Challenge Workshop*, pages 12–20, Nice (France), 2012. (Cited on page 169.)

[9] **Aymeric Stamm**, Olivier Commowick, Christian Barillot, and Patrick Pérez. Adaptive Multi-modal Particle Filtering for Probabilistic White Matter Tractography. In *Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI)*, pages 594–606, Asilomar (California, USA), 2013. (Cited on page 169.)

[10] **Aymeric Stamm**, Olivier Commowick, Patrick Pérez, and Christian Barillot. Tracking the corticospinal tract from low spatial and angular resolution diffusion MRI. In *MICCAI 2013 DTI Tractography Challenge Workshop*, Nagoya (Japan), 2013. (Cited on page 169.)

[11] **Aymeric Stamm**, Olivier Commowick, Patrick Pérez, and Christian Barillot. Diffusion Directions Imaging. *Plos One*, 2013. To be submitted. (Cited on page 166.)

[12] **Aymeric Stamm**, Patrick Pérez, and Christian Barillot. Fast identification of optimal fascicle configurations from standard clinical diffusion MRI using Akaike information criterion. Submitted to the *International Symposium on Biomedical Imaging (ISBI)* to be held in Beijing (China) in 2014.

Statistical analyses of clinical MRI studies

[1] Hélène Raoult, Jan Petr, Elise Bannier, **Aymeric Stamm**, Jean-Yves Gauvrit, Christian Barillot, and Jean-Christophe Ferré. Arterial spin labeling for motor activation mapping at 3T with a 32-channel coil: reproducibility and spatial accuracy in comparison with BOLD fMRI. *Neuroimage*, 58(1):157–167, 2011.

[2] Hélène Raoult, Jan Petr, Jean-Yves Gauvrit, **Aymeric Stamm**, Elise Bannier, Christian Barillot, and Jean-Christophe Ferré. Reproducibility of Arterial Spin Labeling for Motor Activation Mapping: Evaluation of a short 3 min duration sequence. In *European Congress of Radiology (ECR)*, Vienna (Austria), 2011.

[3] Hélène Raoult, Jean-Christophe Ferré, Jan Petr, Elise Bannier, **Aymeric Stamm**, Christian Barillot, and Jean-Yves Gauvrit. Functional arterial spin labeling: Optimal sequence duration for motor activation mapping in clinical practice. *Journal of Magnetic Resonance Imaging*, 36(6):1435–1444, 2012.

[4] Hélène Raoult, François Eugène, Jean-Christophe Ferré, Jean-Christophe Gentric, Thomas Ronzière, **Aymeric Stamm**, and Jean-Yves Gauvrit. Prognostic factors for outcomes after mechanical thrombectomy with solitaire stent. *Journal of Neuroradiology*, 2013.

[5] Aurore Esquevin, Hélène Raoult, Jean-Christophe Ferré, Thomas Ronzière, **Aymeric Stamm**, Mathieu Perennes, Abdelouahab Bellou, and Jean-Yves

Gauvrit. Systematic combined noncontrast CT-CT angiography in the management of unexplained nontraumatic coma. *The American Journal of Emergency Medicine*, 31(3):494–8, 2013.

[6] Aurore Esquevin, Hélène Raoult, Jean-Christophe Ferré, Thomas Ronzière, **Aymeric Stamm**, Mathieu Perennes, Abdelouahab Bellou, and Jean-Yves Gauvrit. Arterial Spin Labeling in semantic dementia: hypoperfusion's detection and diagnostic accuracy. Could it make as well as FDG-PET?. In *Proceedings of the Radiological Society of North America (RSNA)*, Chicago (Illinois, USA), 2013.

Fundamental statistics for large p small n data

[1] Piercesare Secchi, **Aymeric Stamm**, and Simone Vantini. Large p Small n: Inference for the Mean. In *Proceedings of the Scientific Meeting of the Italian Statistical Society (SIS)*, Padua (Italy), 2010.

[2] Piercesare Secchi, **Aymeric Stamm**, and Simone Vantini. A generalization of Hotelling's theorem for large p small n data. In *Proceedings of the International Conference on Complex Models and Computational Intensive Methods for Estimation and Prediction (SCo)*, Padua (Italy), 2011.

[3] Piercesare Secchi, **Aymeric Stamm**, and Simone Vantini. Inference for the mean of large p small n data: a finite-sample high-dimensional generalization of Hotelling's Theorem. *Electronic Journal of Statistics (EJS)*, 2013. Accepted manuscript. In press.

Consulting in Statistics

[1] Olivier Commowick and **Aymeric Stamm**. Non-Local Robust Detection of DTI White Matter Differences with Small Databases. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 476–484, Nice (France), 2012.

[2] Emmanuel Vallée, Olivier Commowick, Camille Maumet, **Aymeric Stamm**, Elisabeth Le Rumeur, Catherine Allaire, Jean-Christophe Ferré, Clément de Guibert, and Christian Barillot. Statistical analysis of white matter integrity for the clinical study of specific language impairment in children. In *MICCAI 2013 Workshop on Computational Diffusion MRI*, Nagoya (Japan), 2013.

Diffusion Directions Imaging: High resolution reconstruction of white matter fascicles from low angular resolution diffusion MRI

Abstract: The objective of this thesis is to provide a complete pipeline that achieves an accurate reconstruction of the white matter fascicles using clinical diffusion images characterized by a low angular resolution. This involves (i) a diffusion model inferred in each voxel from the diffusion images and (ii) a tractography algorithm fed with these local models to perform the actual reconstruction of fascicles.

Our contribution in *diffusion modeling* is a new statistical distribution, the properties of which are extensively studied. We model the diffusion as a mixture of such distributions, for which we design a model selection tool that estimates the number of mixture components. We show that the model can be accurately estimated from single shell low angular resolution diffusion images and that it provides specific biomarkers for studying tumors.

Our contribution in *tractography* is an algorithm that approximates the distribution of fascicles emanating from a seed voxel. We achieve that by means of a particle filter better adapted to multi-modal distributions than the traditional filters. To demonstrate the clinical applicability of our tools, we participated to all three editions of the MICCAI DTI Tractography challenge aiming at reconstructing the cortico-spinal tract from single-shell low angular and low spatial resolution diffusion images. Results show that our pipeline provides a reconstruction of the full extent of the CST.

Keywords: Multi-fiber model, non-Gaussian diffusion, particle filter, tractography

---

Diffusion Directions Imaging: Reconstruction haute résolution des faisceaux de matière blanche par IRM de diffusion basse résolution angulaire

Abstract: L'objectif de cette thèse est de fournir une chaine de traitement complète pour la reconstruction des faisceaux de la matière blanche à partir d'images pondérées en diffusion caractérisées par une faible résolution angulaire. Cela implique (i) d'inférer en chaque voxel un modèle de diffusion à partir des images de diffusion et (ii) d'accomplir une "tractographie", i.e., la reconstruction des faisceaux à partir de ces modèles locaux.

Notre contribution en *modélisation de la diffusion* est une nouvelle distribution statistique dont les propriétés sont étudiées en détail. Nous modélisons le phénomène de diffusion par un mélange de telles distributions incluant un outil de sélection de modèle destiné à estimer le nombre de composantes du mélange. Nous montrons que le modèle peut être correctement estimé à partir d'images de diffusion "single-shell" à faible résolution angulaire et qu'il fournit des biomarqueurs spécifiques pour l'étude des tumeurs.

Notre contribution en *tractographie* est un algorithme qui approxime la distribution des faisceaux émanant d'un voxel donné. Pour cela, nous élaborons un filtre particulaire mieux adapté aux distributions multi-modales que les filtres traditionnels. Pour démontrer l'applicabilité de nos outils en usage clinique, nous avons participé aux trois éditions du MICCAI DTI Tractography challenge visant à reconstruire le faisceau cortico-spinal à partir d'images de diffusion "single-shell" à faibles résolutions angulaire et spatiale. Les résultats montrent que nos outils permettent de reconstruire toute l'étendue de ce faisceau.

Mots-clé: Modèle multi-fibres, diffusion non-Gaussienne, filtre particulaire, tractographie