UNIVERSITÉ PARIS-SUD 11

ÉCOLE DOCTORALE STITS

NEUROSPIN, CEA

DISCIPLINE: PHYSICS

PhD THESIS

defended on the 30/04/2013

by

Véronique BRION

Towards real-time diffusion imaging: noise correction and inference of the human brain connectivity

Thesis supervisor:	Fabrice POUPON	Researcher, NeuroSpin, France
Thesis director:	Cyril POUPON	Researcher, NeuroSpin, France

Jury composition:

Jury president:	Santiago AJA-FERNÁNDEZ	Associate Professor, ETSI, Spain
Reviewers:	Habib BENALI	Research Director, UPMC, France
	Jan SIJBERS	Professor, Vision Lab, Belgium
Examiners:	Jean-Christophe GINEFRI	Associate Professor, IR4M, France
	Denis LE BIHAN	Research Director, NeuroSpin, France
	Cyril POUPON	Researcher, NeuroSpin, France
	Fabrice POUPON	Researcher, NeuroSpin, France

UNIVERSITÉ PARIS-SUD 11

ÉCOLE DOCTORALE STITS

NEUROSPIN, CEA

DISCIPLINE: PHYSIQUE

THESE DE DOCTORAT

soutenue le 30/04/2013

par

Véronique BRION

Imagerie de diffusion en temps-réel: correction du bruit et inférence de la connectivité cérébrale

Superviseur:	Fabrice POUPON	Ingénieur chercheur, NeuroSpin, France
Directeur de thèse:	Cyril POUPON	Ingénieur chercheur, NeuroSpin, France

Composition du jury:

Président du jury:	Santiago AJA-FERNÁNDEZ	Maître de Conférence, ETSI, Espagne
Rapporteurs:	Habib BENALI	Directeur de recherche, UPMC, France
	Jan SIJBERS	Professeur, Vision Lab, Belgique
Examinateurs:	Jean-Christophe GINEFRI	Maître de Conférence, IR4M, France
	Denis LE BIHAN	Directeur de recherche, NeuroSpin, France
	Cyril POUPON	Ingénieur chercheur, NeuroSpin, France
	Fabrice POUPON	Ingénieur chercheur, NeuroSpin, France

Acknowledgments

First, I would like to warmly thank Dr. Fabrice Poupon, my advisor, for offering me to work on his great real-time project: thank you for everything I have learned with you and for your patience, for the calm and confidence present in our team. It really was nice to work with you.

Huge thanks also to Dr. Cyril Poupon, my PhD director, for his constant commitment and his enthusiasm that were contagious and kept me motivated even in the hard times of the thesis: I was very happy to work with you and to learn a lot from you on many fields. The kindness and generosity of both of you are impressive and I am glad to have been part of your team. Thanks for the many missions all around the world!

I would like to thank Pr. Habib Benali and Pr. Jan Sijbers for their true interest to this PhD work through their nice and interesting reviews. I am grateful to them and to the jury president Dr. Santiago Aja-Fernández, the examiners Dr. Jean-Christophe Ginefri and Dr. Denis Le Bihan as well, for their questions and the discussions during the day of my defense.

The real-time project would not have been the same without Olivier Riff: you were such a wonderful colleague, full of kindness and always present to help on the coding and logic. Many people, including me, still miss you at NeuroSpin!

A lot of thanks to Pr. Maxime Descoteaux for welcoming me in Sherbrooke and working together on the parallel Kalman Filter: Canada in autumn is simply beautiful and I have very nice memories of the poutine eaten with you and your great students, as well as the beautiful landscapes of the Laurentides with Alan, Denis, Fabrice and Olivier. Thank you Maxime for all your kind advises and support all along the thesis, as well as for your enthusiasm.

I am also grateful to Dr. Santiago Aja-Fernández and Dr. Antonio Tristán-Vega for the help and advises on the LMMSE filter and for our discussions per mail. I enjoyed very much our collaboration. Thanks as well to Pr. Plataniotis for the mail exchange on the parallel Kalman filter. I also want to thank Dr. Jeny Rajan for our discussions through the different conferences in Leipzig and Barcelona: I wish you all the best in India!

Thanks to the Japanese team for the nice work interactions: Mami Iima and Dr. Shunro Fujiwara, as well as Dr. Denis Le Bihan.

Now, I would like to acknowledge many wonderful colleagues for their support and good mood who contributed to lighten my time at NeuroSpin and make it the place of nice conversations and work.

I am very grateful to you, Linda, for being such an incredible friend: I enjoyed a lot our discussions together; I had the privilege to know you from the beginning of my work at NeuroSpin and I feel so lucky for that. Thank you for your courage and integrity which are models for me.

Huge thanks to Alice, Clarisse and Delphine for your friendship and your cute and courageous way towards life and research: you rock and I was glad to have you as colleagues. Special thanks to Clarisse and Delphine for the help on the Connectomist software.

Thanks to you, Jimmy and Sophie, for your generous friendship and your support until the day of my defense. Thanks to Benoît for your generosity and for all times you drove me to Massy-Palaiseau RER station. Thank you Alfredo, Anne-Laure, Dang, François, Hang, Jennifer, Jing-Rebecca and Pamela for our nice talks together and your support.

Special thanks to the funny and nice Prince-Pepito-Coca-Schweppes-madeleines-take-as-you-want Glunch team with Denis Fournier, Greg (même si on n'appelait pas encore ces moments des "glunch"), Jérémy, Marie-France, Laurence, Laurent, Yann. Merci également à Bernadette, Dominique, Elisabeth, Emmanuelle, Johanna (mmmhh les sarmale !!!), Maryline, Nathalie, ainsi que les hôtesses d'accueil Elodie, Hélène et Virginie : merci pour les moments amicaux avec vous tous ! Merci aussi à l'équipe des manip radio Chantal, Séverine, Séverine et merci à Laurence et Véronique.

Thanks to Alan, Alexandre, Alexis Amadon, Alexis Barbot, Anne, Antonio, Aurélien, Benjamin, Benoît Larrat, Céline, Christine Bakkhous, Christine Delmaire, Clara, Denis Rivière, Dominique Geffroy, Edith, Edouard, Fawzi, Françoise, Gary, Ileana, Irina, Jeff, Josselin, Joy, Julien, Karla, Lucie, Luisa, Martijn, Nicolas Boulant, Nicolas Zilber, Olivier Reynaud, Pamela, Pauline, Samuel, Sébastien, Soizic, Solveig, Thomas, Urielle, Virgile...

Thanks to the Irtelis group with Dr. Sylvie Esterlin. I had really good times forgetting about my thesis work and focusing on interesting subjects mainly for the professional project.

Finally, a lot of thanks to my friends from Grenoble for your precious presence throughout these three and a half years: Aurélie (merci pour ta bonne humeur exotique !), Blandine, Élodie, Fabien, Hugo, Isabelle (tes encouragements jusqu'au bout), Jérémy (merci pour les coups de fil pleins d'énergie et de confiance), Laëtitia (merci pour les "entrevues de la vie la vraie" en pleine période de stress !), Marie (merci pour les mails de l'espace et la force Bob Morane), Olivier, Solveig, Vincent, Yannick. J'espère bien profiter de notre amitié encore et toujours !

Un grand merci aussi à mes amis du lycée qui ont contribué à mon bon moral au fil des années de thèse : Camille, Caroline, Christophe, François, Gwenaëlle, Laëtitia, Laure.

Merci à Mme Michèle Pascal pour les séances de tennis décrassantes du week-end !

Merci à vous, Mme Nadine Fagot, pour votre présence et soutien à ma soutenance.

A special thank to Mme Claude Parker for your wonderful friendship. Your courage is a model and a force for many of us.

Pour finir, je souhaiterais remercier ma famille : Anne, Dorothée, Elisabeth, Haris, Laurent, Maman, Mamie de Nantes, Mamie de Reims, Marie, Olivier, Papa (merci à vous trois pour la relecture des premiers chapitres !), Papy de Nantes, Papy de Reims, Pénélope. Merci pour votre soutien et confiance sans faille ainsi que pour vos conseils sur la thèse et sur la recherche d'emploi.

Je dédie cette thèse à mes sœurs et mes parents.

Contents

List of Figures

List of Tables

List of symbols

Let us point out that the vectors are written in bold in the manuscript text, and sometimes with an arrow above them in some figures.

2D	two dimensional
ACS	auto-calibration signal
aQBI	analytical QBI
CC	corpus callosum
cdODF	corrected dODF
cGFA	corrected GFA
CPU	central processing unit
CSF	cerebrospinal fluid
DCT	discrete cosine transform
dMRI	diffusion MRI
dODF	diffusion orientation distribution function
DSI	diffusion spectrum imaging
DTI	diffusion tensor imaging
DW	diffusion weighted
FA	fractional anisotropy
FFT	fast Fourier transform
FID	free induction decay
fODF	fiber orientation distribution function
FOV	field of view
FRT	Funk-Radon transform
GFA	generalized fractional anisotropy
GM	grey matter
GRAPPA	generalized autocalibrating partially parallel acquisition
HARDI	high angular resolution diffusion imaging
LMMSE	linear minimum mean square error
LUT	lookup table
ML	maximum likelihood
MMSE	minimum mean square error
MSE	mean square error

MRI	magnetic resonance imaging
nc-χ	noncentral χ
NLM	non local means
NMR	nuclear magnetic resonance
ODF	orientation distribution function
OLS	ordinary least squares
PCA	principal component analysis
PDE	partial differential equation
PDF	probability density function
PGSE	pulse gradient spin echo
PKF	parallel Kalman filter
QBI	Q-Ball imaging
RBW	read bandwidth
RF	radio-frequency
ROI	region of interest
RT	real time
rtdMRI	real-time dMRI
sa-aQBI	aQBI with the solid angle
SD	spherical deconvolution
SENSE	sensitivity-encoding for fast MRI
SH	spherical harmonics
SMASH	simultaneous acquisition of spatial harmonics
SNR	signal-to-noise ratio
SoS	sum of squares
WLS	weighted least squares
WM	white matter

Abstract

Most magnetic resonance imaging (MRI) system manufacturers propose a huge set of software applications to post-process the reconstructed MRI data a posteriori, but few of them can run in real-time during the ongoing scan. To our knowledge, apart from solutions dedicated to functional MRI allowing limited relatively simple experiments or for interventional MRI generally developed to perform anatomical scans during surgery, no tool has been developed in the field of diffusion-weighted MRI (dMRI). However, because dMRI scans are extremely sensitive to lots of hardware or subject-based perturbations inducing corrupted data, it can be interesting to investigate the possibility of processing dMRI data directly during the ongoing scan and this thesis is dedicated to this challenging topic. The organization of the methodological developments achieved in this thesis is fourfold. First, a high performance real-time software platform was developed and connected to the 3T MRI system of the NeuroSpin center offering the computational ressources mandatory to allow post-processings in a dozen of seconds (corresponding to the average repetition time needed to acquire a full diffusion-weighted volume). The second objective, in fact the major contribution of this thesis, aimed at providing solutions to the challenging problem of denoising dMRI data in real-time. Indeed, the diffusion-weighted signal may be corrupted by a significant level of noise strongly reducing the angular resolution of advanced local angular diffusion models requiring the use of high b-values yielding a low signal-to-noise ratio (SNR). Because the noise distribution corrupting the magnitude of the MRI signal is not Gaussian anymore but Rician or noncentral χ, averaging cannot be used as a strategy to denoise the signal. After making a detailed review of the literature, we extended the linear minimum mean square error (LMMSE) estimator and adapted it to our real-time framework. We compared its efficiency to the standard Gaussian filtering approach more difficult to implement in most scanners as it requires a modification of the reconstruction pipeline to insert the filter immediately after the demodulation and translation of the acquired complex signal in the Fourier space. We proved that our real-time LMMSE denoising method implemented with a Kalman filter was more efficient in practice than this conventional Gaussian filtering method, since it allowed to preserve anatomical details by taking into account the underlying structure encoded in the MRI signal. We also developed a parallel Kalman filter to deal with any noise distribution and we showed that its efficiency was quite comparable to the non parallel Kalman filter approach. The third objective of this thesis was to demonstrate the feasibility of computing estimates of

the field of the orientation distribution functions such as the diffusion tensor model, the analytical Q-ball model in an incremental way using a Kalman framework, and in less than a repetition time. We proved that it could be achieved using our high performance computing environment, thus opening the way to get estimations of any quantitative diffusion map incrementally during an ongoing scan. Last, we addressed the feasibility of performing tractography in real-time in order to infer the structural connectivity online. We hope that this set of methodological developments will help improving the diagnosis of brain pathologies, in particular when a quick diagnosis has to be performed in case of emergency to check the integrity of white matter fiber bundles.

Résumé (français)

Chapitre 1: Introduction

L'IRM de diffusion (IRMd), introduite par [LeBihan and Breton (1985); Merboldt et al. (1985); Taylor and Bushell (1985)], est une modalité d'imagerie très utilisée aujourd'hui. Ses domaines d'applications vont du diagnostic précoce d'accidents ischémiques jusqu'à l'inférence de la connectivité anatomique cérébrale. L'IRMd repose sur la mesure du phénomène physique de diffusion des molécules d'eau contenues dans le cerveau, cette mesure permettant d'en déduire les directions des faisceaux de fibres de la substance blanche à l'intérieur du cerveau. Durant les vingt dernières années, de nombreuses méthodes ont été développées pour représenter au mieux la carte de ces faisceaux reliant des régions cérébrales entre elles. Avec l'amélioration de la précision de ces cartes représentant les autoroutes cérébrales, l'enthousiasme s'est intensifié dans la communauté et de nouvelles investigations se sont multipliées pour mieux étudier la microarchitecture cérébrale.

Récemment, une nouvelle avancée pour l'IRMd, ainsi que pour l'IRM fonctionnelle (IRMf), a émergée avec l'apparition de l'IRM dite temps-réel (TR). L'IRM TR est un nouveau concept permettant de réaliser les habituels post-traitement des images par IRM en direct, c'est-à-dire en même temps que l'acquisition a lieu avec le patient dans le scanner IRM. L'IRMf TR, introduite par [Cox et al. (1995)], a ouvert de nouvelles portes dans plusieurs domaines. Cette technique a par exemple permis de communiquer avec des patients complètement paralysés, dont les médecins pensaient qu'ils étaient dans un état végétatif [Birbaumer et al. (1999)]. L'IRMf TR a apporté de nouvelles possibilités pour générer des interfaces cerveau-machine [Weiskopf et al. (2004); deCharms (2007)], jusqu'à pouvoir réduire des douleurs chroniques chez des patients grâce à un entraînement dit de neurofeedback (réponse rétroactive à un stimulus cérébral) [deCharms (2008)]. En parallèle de ces dévelopements permettant de lire les processus de fonctionnement cérébral en TR, l'apparition de l'IRMd TR (IRMdtr) [Poupon et al. (2008b)] a ouvert la voie de la lecture de l'architecture structurale du cerveau en TR, pendant l'examen du patient. Dans le cadre de cette thèse, nous nous focalisons sur l'IRMdtr.

L'IRMdtr permet de voir en TR le résultat du calcul des cartes d'IRMd qui se raffinent chaque fois qu'une nouvelle mesure de la diffusion de l'eau est effectuée. Les cartes habituelles d'IRMd —comme celles de l'anisotropie fractionnelle, du coefficient de diffusion apparent, des functions de distributions des orientations de la diffusion ou des fibres—

peuvent être évaluées presque instantanément alors que la séquence d'IRMd est en cours d'exécution. Les possibilités offertes par ce traitement TR sont puissantes. En effet, les résultats obtenus en direct permettent de juger efficacement de la qualité des mesures effectuées sur le patient et éventuellement de corriger les artéfacts dus à des problèmes inattendus tels que le mouvement du patient ou de détecter un problème technique... Enfin, ce système TR offre la possibilité d'insérer un traitement incrémental supplémentaire sur les données acquises, accélérant le processus de traitement des données et permettant d'arrêter l'acquisition dès que les estimations ont produit des résultats suffisamment informatifs en vue du diagnostic.

L'imagerie de diffusion est caractérisée par une décroissance exponentielle du signal qui génère un signal résiduel fortement bruité. Alors que ce bruit reste gaussien sur les deux canaux d'acquisition du signal complexe, il devient ricien ou χ non centré, lorsque l'on considère l'amplitude du signal des données IRM reconstruites. La tendance actuelle avec les modèles d'imagerie de diffusion à haute résolution angulaire (HARDI) et d'imagerie de diffusion hybride (HYDI) est d'augmenter la pondération en diffusion dans le signal, ce qui intensifie le niveau de bruit dans le signal. De plus, il peut être intéressant de diminuer la taille du voxel pour augmenter la résolution spatiale des images, ce qui diminue également le rapport signal à bruit (SNR). Par conséquent, une méthode de débruitage en TR devient essentielle pour utiliser l'IRMdtr de manière optimale.

Dans cette thèse, nous nous sommes intéressés à la possibilité de corriger le bruit en TR, ce qui peut être très difficile si l'on considère le laps de temps disponible d'une douzaine de secondes entre l'acquisition de deux volumes. Tout d'abord, nous nous sommes focalisés sur une des techniques de débruitage offline (c'est-à-dire effectué en post-traitement). Cette technique, qui s'adapte bien à nos contraintes de TR, est l'estimateur linéaire minimisant l'erreur quadratique moyenne (LMMSE) qui a été d'abord dédié au bruit ricien et que nous avons étendu à la correction de bruit χ non centré. Puis, nous avons développé une première méthode de correction en TR fondée sur l'association de ce LMMSE étendu avec un filtre de Kalman connectés ensemble avec l'intégration d'une boucle de feedback (rétroaction) pour contrôler le surlissage induit par le LMMSE. En plus de cette première méthode, nous avons développé une autre technique de débruitage TR reposant sur un outil unique: un filtre de Kalman parallèle prenant directement en compte le caractère non gaussien du bruit. Chaque technique a été appliquée en TR sur les signaux d'amplitude reconstruits par la technique dite "sum of squares" (SoS) (par somme des carrés). Nous avons également élaboré dans cette thèse une méthode de correction du bruit en TR qui s'applique directement sur les signaux avant la reconstruction SoS. Cette dernière technique a l'avantage de pouvoir considérer un bruit gaussien à moyenne nulle. L'ensemble de ce nouvel environnement TR a été testé sur un imageur IRM Tim Trio à 3T et pourrait être adapté à tout autre système clinique d'IRM. Nous avons également abordé la question de la faisabilité du suivi de fibres en TR permettant de voir les fibres se construire de manière incrémentale à chaque nouvelle acquisition. Nous avons étudié l'impact de la correction du bruit en TR pour cette application de tractographie.

Cette thèse est organisée de la manière suivante: les chapitres 2, 3 et 4 sont regroupés dans la partie sur l'état de l'art de l'IRMdtr. Ils couvrent les bases de l'IRM (chapitre 2), celles de l'IRMd (chapitre 3) pour converger vers l'IRMdtr (chapitre 4). La partie suivante contient nos contributions sur la correction du bruit en TR dédiée l'IRMdtr. Le chapitre 5 contient un important état de l'art sur la description du bruit en IRM, ainsi que des contributions sur une méthode d'estimation de la variance du bruit sans utiliser le fond de l'image et la proposition d'une nouvelle mesure de la qualité d'une image d'IRMd. Le chapitre 6 présente la problématique de la correction du bruit en TR et apporte une solution reposant sur un LMMSE étendu. Puis, le chapitre 7 propose trois méthodes de correction du bruit en TR. La première repose sur la combinaison d'un LMMSE et d'un filtre de Kalman incorporant une boucle de feedback. La deuxième repose sur un débruitage effectué sur les données de l'espace **k**, avant la reconstruction SoS. La dernière, qui est appliquée comme la première sur les signaux d'amplitude, est fondée sur un filtre de Kalman parallèle. Enfin, la dernière partie contient le chapitre 8 qui propose de sonder la possibilité de réaliser un suivi de fibres en TR. Ce dernier chapitre met en valeur l'impact de la correction du bruit en TR sur l'estimation des fibres réalisée en TR.

Chapitre 2: Principes de base de l'imagerie par résonance magnétique nucléaire (IRM)

L'imagerie par résonance magnétique nucléaire (IRM) a été introduite par [Lauterbur (1973)] et [Mansfield (1977)]. Cette modalité permet de générer des images du corps de manière non invasive. L'IRM est fondée sur la réponse des protons d'hydrogène à des perturbations magnétiques imposées. En effet, comme le proton possède un moment magnétique, il est sensible au magnétisme ambiant. Cette sensibilité est à la base de la modalité d'IRM.

La figure 1 schématise un scanner IRM. Celui-ci est constitué par une bobine supraconductrice qui génère un champ magnétique $\mathbf{B_0}$ homogène dans un certain champ de vue situé autour de l'isocentre du champ magnétique. Une antenne permet de perturber l'équilibre magnétique généré par la bobine supraconductrice. Elle permet de déclencher le phénomène de résonance magnétique. Elle peut être aussi utilisée pour réceptionner le signal d'IRM causé par la réponse des protons à la perturbation magnétique imposée. L'antenne de réception est en général adaptée à la géométrie de l'objet d'intérêt, ici la tête. Enfin, les bobines de gradients permettent d'encoder spatialement le signal reçu et de créer l'image.

La résonance magnétique et les phénomènes de relaxation.

On définit le vecteur d'aimantation $\mathbf{A}$ comme étant la somme des moments magnétiques de chaque spin de l'objet étudié divisé par le volume de l'objet. Lorsque le champ $\mathbf{B_0}$ est créé par la bobine supraconductrice, l'aimantation $\mathbf{A}$ s'aligne le long de ce champ et

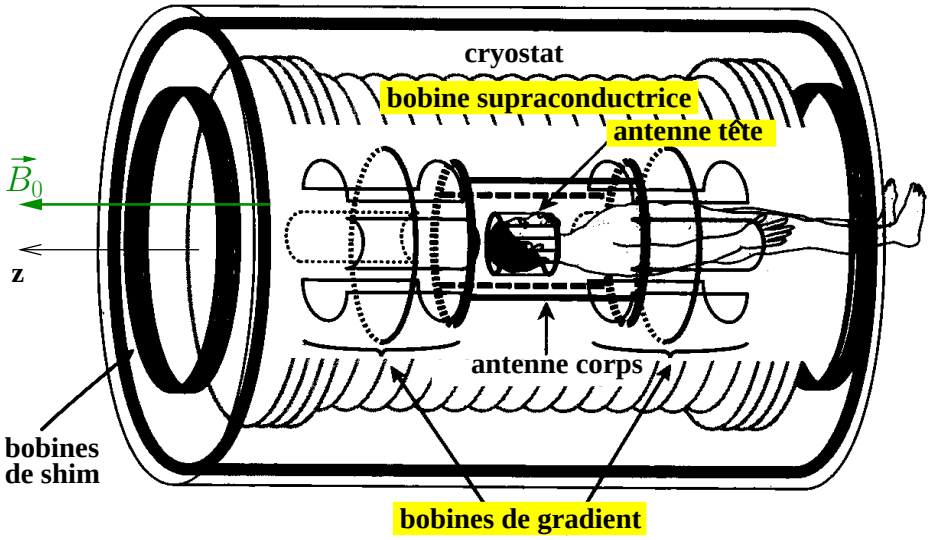


Figure 1: Structure d'un scanner IRM. Fig. adaptée de [Kastler et al. (2006)].

est orientée comme $\mathbf{B_0}$. Cet équilibre peut être perturbé par une antenne RF qui crée un champ tournant $\mathbf{B_1}$, comme indiqué sur la figure 2. Pour cela, il faut que la vitesse angulaire ω_{exc} de $\mathbf{B_1}$ respecte la condition de résonance: $\omega_{exc} = \gamma \mathbf{B_0}$ ([Bloch (1946)], [Purcell et al. (1946)]). La figure 2 montre comment l'impulsion RF permet de faire basculer l'aimantation (à 90° sur la figure). Au bout d'un certain temps, la relaxation de la composante transversale de $\mathbf{A}$ a lieu, suivie par la relaxation de la composante longitudinale de $\mathbf{A}$.

L'évolution de l'aimantation peut être décrite par l'équation de Bloch dans le référentiel du laboratoire (Oxyz) [Bloch (1946)]:

$$\boxed{\frac{d\mathbf{A}}{dt} = \gamma\left(\mathbf{A} \times \mathbf{B}\right) - \frac{A_x\mathbf{x} + A_y\mathbf{y}}{T_2} - \frac{A_z - A_m}{T_1}\mathbf{z}}, \tag{1}$$

où $\times$ est le produit vectoriel. T_1 et T_2 sont les temps de relaxation des composantes respectivement longitudinales et transversales de l'aimantation et $A_m = A_z(\infty)$.

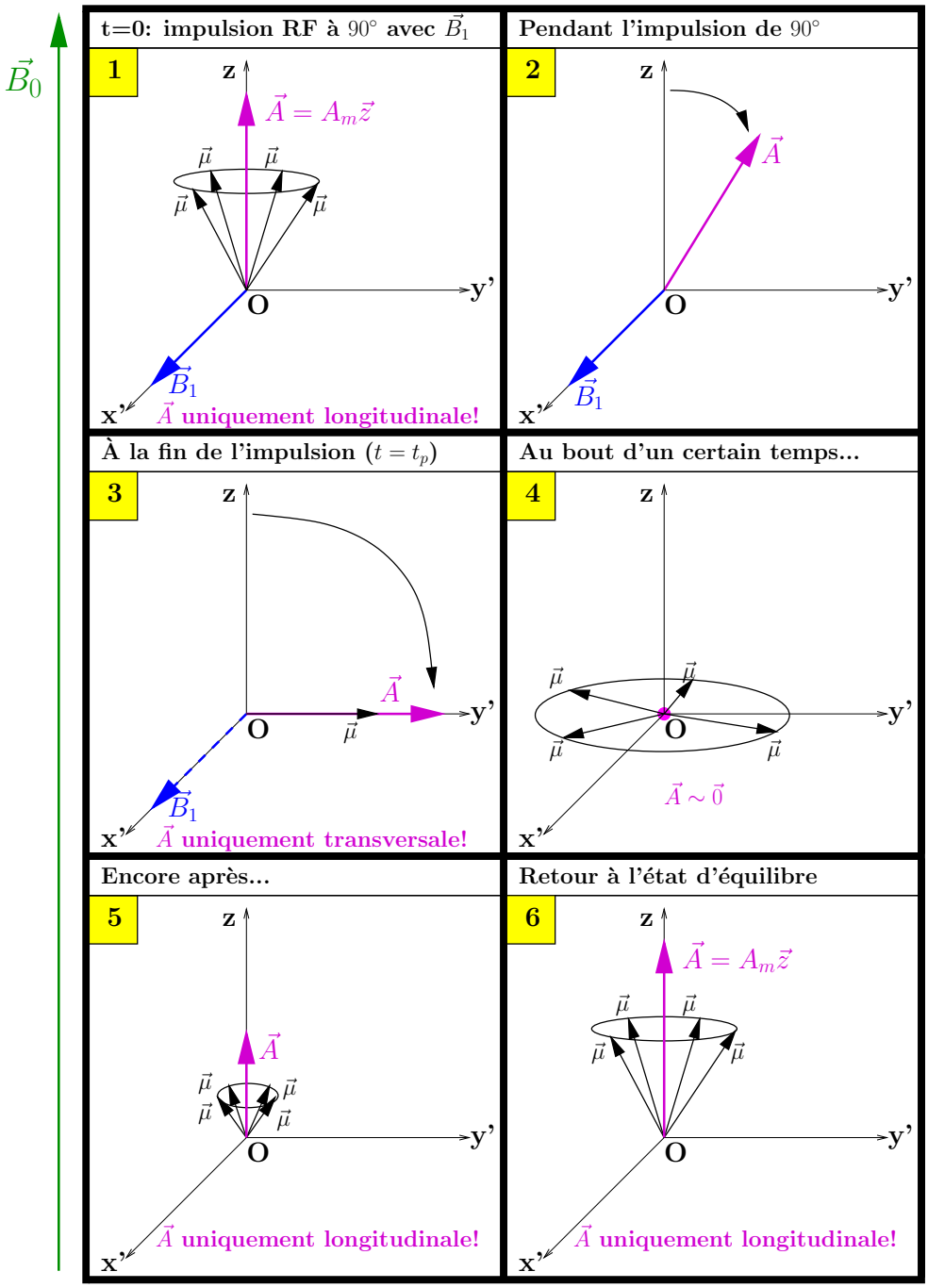


Figure 2: Effet d'une impulsion RF à 90° RF sur le vecteur d'aimantation $\vec{A}$. Cette impulsion RF crée un champ magnétique tournant $\mathbf{B_1}$ générant le phénomène de résonance magnétique (vignettes 1, 2 et 3). Puis, le phénomène de relaxation a lieu (vignettes 4,5 et 6).

Bases d'une séquence d'écho de spin.

La figure 3 indique les bases d'une séquence d'écho de spin [Hahn (1950)]. La séquence débute immédiatement après l'impulsion à 90°. À $t = 0$, les spins ont la même phase et

l'aimantation transversale est à son maximum. Puis, les spins se déphasent entre eux: c'est la relaxation T_2^*, T_2^* étant le temps de relaxation transversale effectif. À $t = T_E/2$, une impulsion RF à 180° est appliquée. Elle permet une refocalisation des spins, comme expliqué sur la figure 4. À $t = T_E$, les spins sont à nouveau en phase et l'aimantation transversale génère le signal d'écho. Le module de l'aimantation transversale $A_\perp$ est alors égal à:

$$\boxed{A_\perp(T_E) = A_m \exp[-T_E/T_2]}. \tag{2}$$

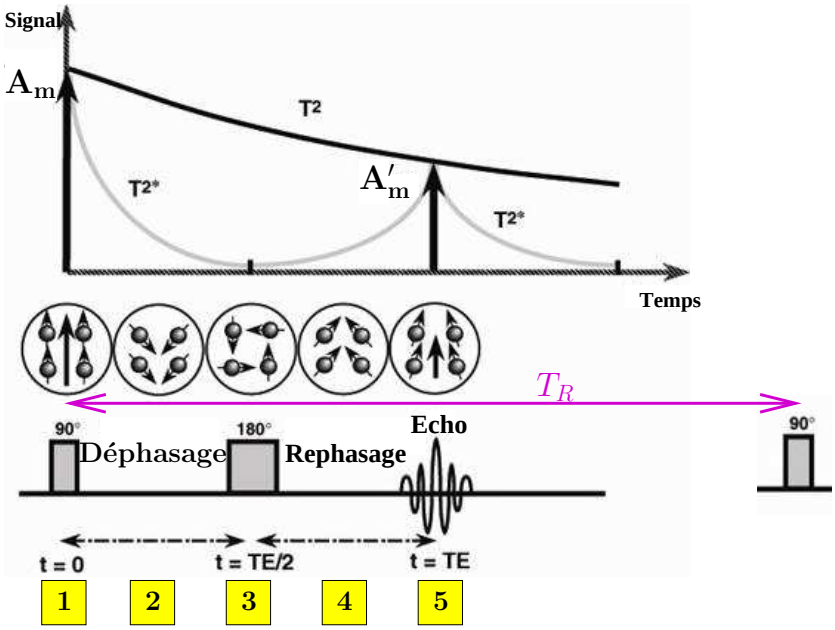


Figure 3: Bases d'une séquence d'écho de spin. Le bas de la figure indique les étapes de la séquence constituée d'une impulsion RF à 90° et d'une autre à 180°. L'origine du temps est fixée immédiatement après l'impulsion à 90°. L'écho est obtenu au temps T_E. La séquence est ensuite répétée au temps T_R, le temps de répétition. Le haut de la figure indique l'évolution de l'aimantation transversale, ainsi que le temps de relaxation T_2, et le temps effectif T_2^* plus court, dû aux inhomogénéités du champ $\mathbf{B_0}$. La configuration des spins est montrée sous le graphe. Figure adaptée de [Kastler et al. (2006)]. Les instants numérotés en jaune sont détaillés à la figure 4.

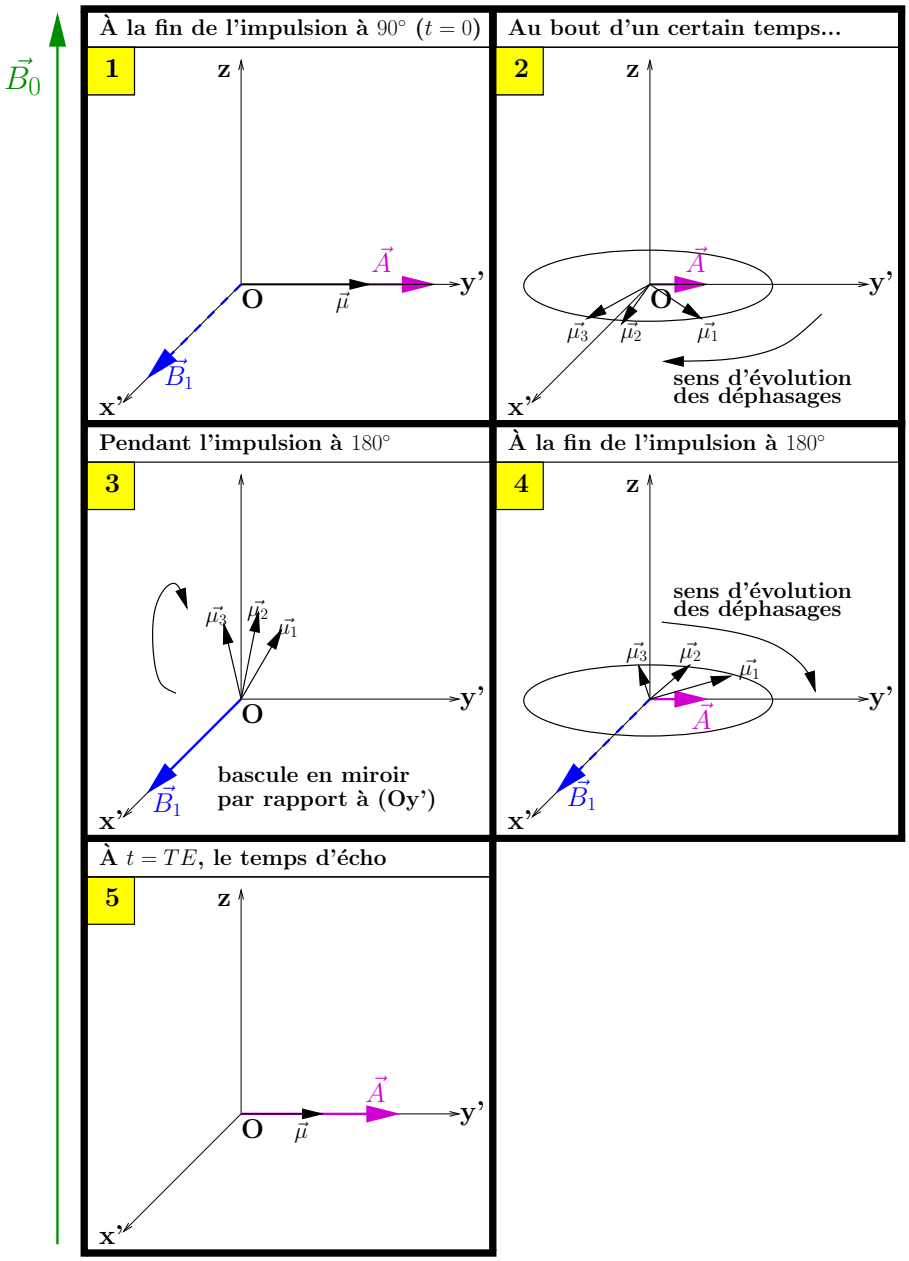
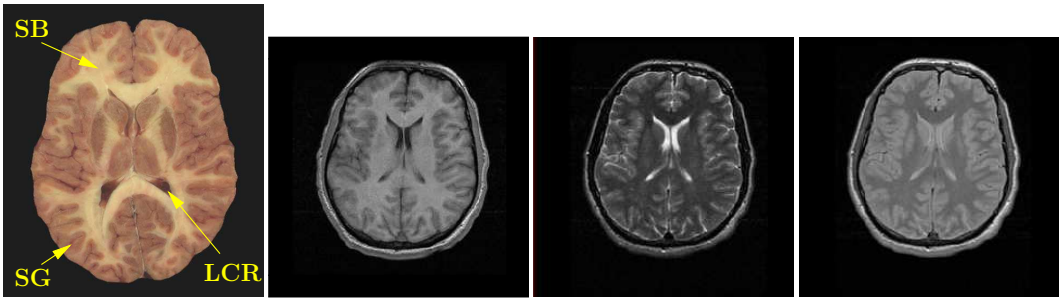


Figure 4: Figure détaillé sur l'écho de spin. Les numéros des vignettes correspondent aux numéros indiqués sur la figure 3.

Le contraste en IRM.

L'aimantation transversale, qui est le signal mesuré en IRM, n'est pas la même suivant les tissus du cerveau. Ce contraste est dû aux différences des temps de relaxation, ainsi que du taux de proton selon le tissu. Il est possible de choisir les paramètres de séquence pour révéler le contraste suivant T_1, T_2 ou bien la densité protonique, comme indiqué sur la figure 5.



Coupe histologique | **Coupe pondérée en T_1** | **Coupe pondérée en T_2** | **Coupe pondérée en densité de protons**

Figure 5: Coupe histologique d'un cerveau humain (image extraite de http://www.valley-neuroscience.com) et différents contrastes en IRM à 1.5T: coupes pondérées en T_1, T_2 et en densité de protons (extraits de http://users.fmrib.ox.ac.uk/~peterj/lectures/hbm_1/sld031.htm).

L'encodage spatial du signal d'IRM.

L'encodage spatial du signal d'IRM est réalisé par une technique reposant sur l'utilisation de gradients de champ magnétique [Carr and Purcell (1954); Lauterbur (1973)]. Tout d'abord, un gradient $\mathbf{G_{SS}}$, dit de sélection de coupe, permet d'exciter les spins d'un plan de coupe particulier. Ce gradient est appliqué en même temps que l'impulsion d'excitation RF. Une fois la coupe sélectionnée, un encodage du signal en 2D est réalisé par un gradient de codage de phase $\mathbf{G_\Phi}$ (noté aussi $\mathbf{G_y}$) et un gradient de fréquence $\mathbf{G}_\omega$ (noté aussi $\mathbf{G_x}$), également dit gradient de lecture car il est appliqué lors de la lecture du signal. $\mathbf{G_\Phi}$ modifie les phases des spins contenues dans la coupe sélectionnée par $\mathbf{G_{SS}}$ selon des lignes de voxels. $\mathbf{G}_\omega$ permet de différencier les voxels d'une ligne donnée selon leur fréquence.

La reconstruction de l'image.

Le signal d'aimantation transversale est mesuré par l'antenne de réception. Celle-ci mesure une tension $e(t)$, dont l'expression découle de la loi d'induction de Faraday et du principe de réciprocité [Haacke et al. (1999)]:

$$e(t) = -\frac{d\Phi(t)}{dt} = -\frac{d}{dt}\int_{\text{volume}} \mathbf{B}^{receive}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r},t) d^3\mathbf{r}, \tag{3}$$

avec $\Phi(t)$ le flux magnétique induit dans la bobine de réception par l'aimantation transversale émise, $\mathbf{B}^{receive}(\mathbf{r})$ le champ magnétique créé par la bobine à la position $\mathbf{r}$ pour un courant unitaire et enfin $\mathbf{A}(\mathbf{r},t)$ l'aimantation locale. Après une démodulation du signal $e(t)$ et d'un filtrage passe-bas, il résulte un signal complexe exprimé dans l'espace $\mathbf{k}$ tel que:

$$\boxed{s(k_x(t),k_y(t)) = \omega_0 \Lambda B_\perp \iint_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(\mathbf{r},0) e^{-i2\pi(k_x(t)x + k_y(t)y)} dxdy}, \tag{4}$$

avec ω_0 la fréquence de Larmor, Λ un facteur de proportionnalité, $B_\perp$ l'amplitude (module) du champ magnétique $\mathbf{B}^{receive}$ et $A_\perp(\mathbf{r}, 0)$ l'amplitude initiale de l'aimantation transversale. $k_x(t) = \frac{\gamma}{2\pi}\int_0^t G_x(t')dt'$ et $k_y(t) = \frac{\gamma}{2\pi}\int_0^t G_y(t')dt'$ sont les coordonnées dans l'espace $\mathbf{k}$ définies en fonction des gradients et du rapport gyromagnétique du proton γ. Une double transformée de Fourier inverse appliquée à $s(k_x(t), k_y(t))$ pour $t = T_E$ conduit au signal final:

$$\boxed{s_{final}(x, y) = \omega_0 \Lambda B_\perp \int e^{-T_E/T_2^*(x,y,z)} A_\perp(x, y, z, 0)dz}. \tag{5}$$

L'amplitude de $s_{final}(x, y)$ est le signal reconstruit en IRM. La figure 6 en donne un exemple.

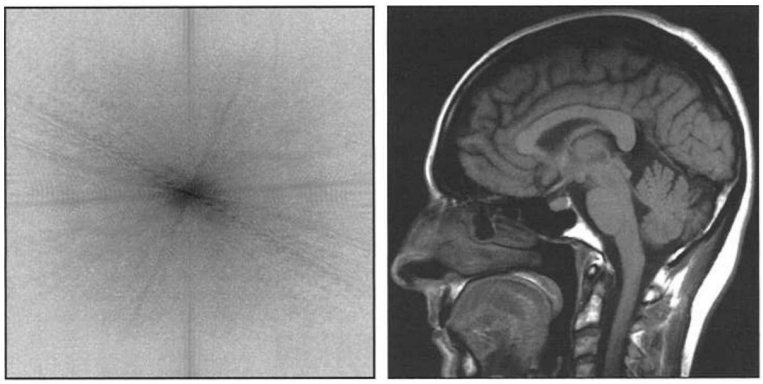


Figure 6: Amplitude du signal dans l'espace $\mathbf{k}$ (gauche) et dans le domaine spatial (droite). Pour passer de l'image de gauche à l'image de droite, une transformée de Fourier inverse 2D doit être appliquée. Extrait de [Kastler et al. (2006)].

Techniques d'IRM parallèle.

L'IRM parallèle consiste à utiliser un réseau d'antennes pour recueillir le signal d'aimantation transversale [Roemer et al. (1990)], comme montré sur la figure 7.

L'IRM parallèle permet d'exploiter l'information spatiale délivrée par chaque antenne grâce à son profil de sensibilité pour remplacer l'encodage spatial du signal effectué par les gradients de codage de phase et de fréquence en IRM classique. Cela permet de diminuer le nombre de lignes acquises de l'espace $\mathbf{k}$ et ainsi de réduire le temps d'acquisition de la séquence. Parmi les techniques de reconstruction les plus utilisées en IRM parallèle, on peut compter l'algorithme SENSE (pour sensitivity-encoding for fast MRI) qui réalise la reconstruction de l'image dans le domaine spatial. À cause du sous-échantillonnage de l'espace $\mathbf{k}$, un artéfact de repliement apparaît dans le domaine spatial et l'algorithme SENSE permet, à partir des informations de sensibiltés des antennes, de "déplier" l'image. Un autre algorithme très connu est la technique GRAPPA (generalized autocalibrating partially parallel acquisition) qui, elle, reconstruit l'image complète dans le domaine de

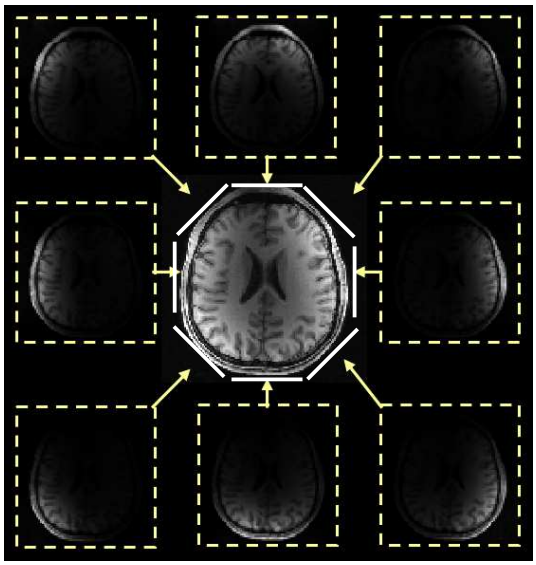


Figure 7: Principe de l'IRM parallèle avec ici huit antennes. L'image centrale correspond à la combinaison des images obtenues par chaque antenne. Figure adaptée de http://black.bme.ntu.edu.tw/courses/course_neuroimaging_fall10/slide_pmri_fhlin.pdf.

Fourier. Il faut pour cela déterminer les lignes manquantes, dues au sous-échantillonnage, grâce à une technique d'autocalibrage et de reconstruction par blocs.

Chapitre 3: Connaissances sur l'IRM de diffusion (IRMd) utilisées dans cette thèse

L'IRM de diffusion (IRMd) a été introduite en 1985 par [LeBihan and Breton (1985); Merboldt et al. (1985); Taylor and Bushell (1985)]. Cette modalité permet de diagnostiquer des accidents cérébraux ischémiques à un stade précoce [Moseley et al. (1990); LeBihan et al. (1992)]. Dans ces pathologies, un œdème apparaît, générant une diminution du processus de diffusion de l'eau visible en IRMd. D'autre part, l'IRMd est aujourd'hui le seul moyen non invasif d'explorer la connectivité anatomique cérébrale *in vivo*. Les connexions dans le cerveau influent sur la diffusion des molécules d'eau, qui est détectable par l'IRMd.

Phénomène physique de la diffusion de l'eau.

Dans un milieu sans obstacle ni restriction, des molécules sont gouvernées par un mouvement brownien [Brown (1828)]. On définit alors le coefficient de diffusion D qui mesure la capacité des molécules à diffuser. Dans un milieu à obstacles comme la substance blanche (SB), le mouvement des molécules n'est plus brownien car il se heurte à la microarchitecture du tissu (figure 8). On définit alors le coefficient de diffusion apparent (ADC), différent du coefficient de diffusion libre.

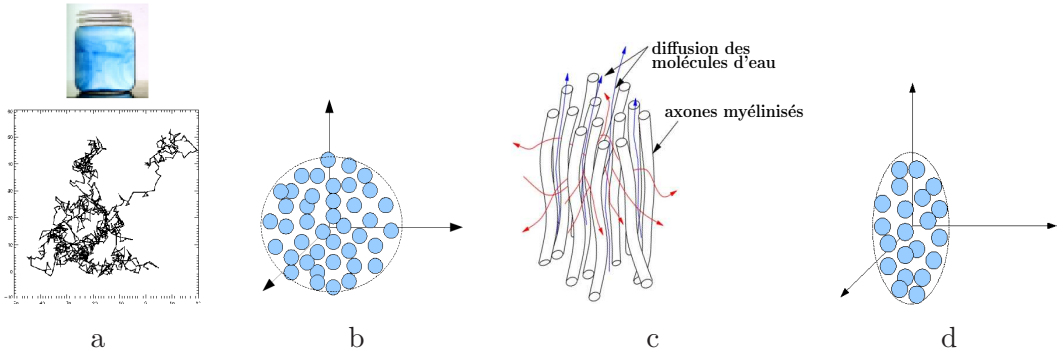


Figure 8: Représentation du phénomène de diffusion et du libre parcours moyen dans un environnement sans obstacle (a et b) et dans la substance blanche (c et d). En (a), la diffusion est libre et donc isotrope, comme indiqué en (b). En (c), la diffusion est restreinte à l'intérieur des axones et entravée de part et d'autre. Le processus de diffusion est donc anisotrope, avec des directions privilégiées le long des axones (flèches bleues) et des directions perpendiculaires aux axones (flèches rouges) pour lesquelles le mouvement est inhibé. Figure (a) extraite de Wikipedia, figures (b) et (d) extraites de [Poupon (2010)] et figure (c) extraite de [Poupon (1999a)].

Outre l'ADC, on définit également la fonction densité de probabilité de déplacement des protons qu'on appelle le propagateur et que l'on note $P(\mathbf{r}, \tau)$. Cette fonction représente la probabilité des protons à se déplacer d'une distance $\mathbf{r}$ pendant une durée τ.

Mesure de la diffusion de l'eau avec l'IRMd.

On peut mesurer la diffusion avec une séquence dite pulse gradient spin echo développée par [Stejskal and Tanner (1965)] et décrite par la figure 9. Cette séquence se distingue d'une séquence de spin écho par les deux gradients de diffusion $\mathbf{g(o)}$ orientés selon la direction donnée par le vecteur $\mathbf{o}$. Le premier gradient sert à marquer la position des spins le long de la direction $\mathbf{o}$. S'il y a diffusion des spins, alors ceux-ci se déphasent et ne sont pas refocalisés suite à l'application du deuxième gradient de diffusion après l'impulsion à 180°. Le signal d'aimantation est donc atténué, ce qui permet d'avoir une indication sur la diffusion des protons le long de la direction testée $\mathbf{o}$.

En faisant l'approximation d'un propagateur de diffusion gaussien, le signal mesuré, dit alors pondéré en diffusion, s'écrit tel que:

$$S(b, \mathbf{g}) = S_0 \exp[-b\mathrm{ADC}(\mathbf{o})] \propto \exp[-T_E/T_2]\exp[-b\mathrm{ADC}(\mathbf{o})], \tag{6}$$

où b en $s \cdot mm^{-2}$ est le paramètre de pondération en diffusion ou *b-value* [LeBihan (1991)], S_0 est le signal pondéré en T_2 qui serait obtenu sans l'application des gradients de diffusion. Dans le cas de gradients de diffusion de forme rectangulaire: $b = \gamma^2\|\mathbf{g}\|^2\delta^2(\Delta - \delta/3) = \gamma^2\|\mathbf{g}\|^2\delta^2\tau$ (γ étant le rapport gyromagnétique du proton et $\|\cdot\|$ l'opérateur de norme Euclidienne), avec $\tau = \Delta - \delta/3$ le temps de diffusion effectif comportant le terme correctif $\delta/3$ dû à la diffusion qui a lieu lors de la durée d'application des gradients. Les temps Δ et δ sont indiqués sur la figure 9.

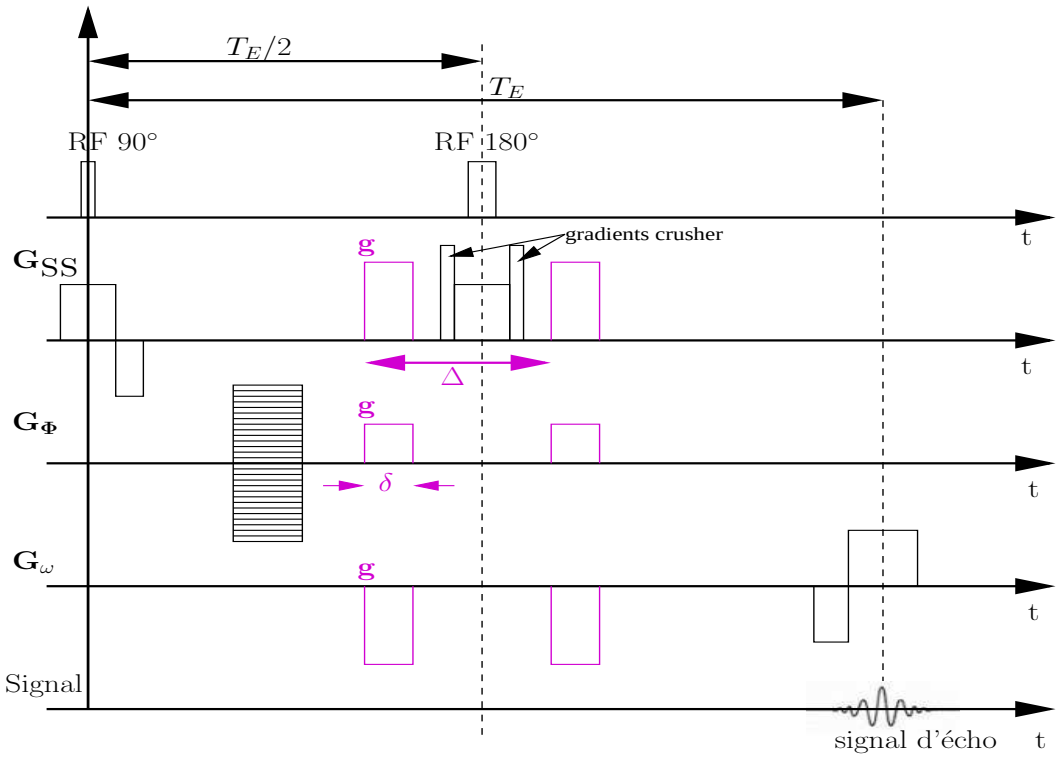


Figure 9: Séquence PGSE. Pour mesurer le processus de diffusion, deux gradients de diffusion (en magenta), notés **g**, sont ajoutés, l'un avant l'impulsion RF à 180° et l'autre après.

Modèles locaux de diffusion.

En IRMd, la séquence PGSE permet d'obtenir une grande quantité de volumes de coupes pondérées en diffusion, chaque volume correspondant à une mesure de la diffusion selon une orientation donnée. Il s'agit ensuite de rassembler les informations fournies par ces volumes pour obtenir un seul volume faisant la synthèse de ces données. Pour cela, plusieurs modèles locaux de diffusion ont été introduits. Ces modèles tentent d'inférer au plus près le propagateur de diffusion $P(\mathbf{r}, \tau)$ qui contient l'information de probabilité sur les trajectoires de diffusion des protons dans le cerveau. On peut se limiter à reconstruire la projection radiale de ce propagateur [Tuch (2004)]:

$$\Psi(\theta_i, \phi_i) = \int_0^\infty P(r, \theta_i, \phi_i, \tau) dr, \tag{7}$$

où $\Psi(\theta_i, \phi_i)$ est la fonction de distribution des orientations de diffusion (dODF). En réalité, cette définition ci-dessus est incomplète. La projection radiale du propagateur s'écrit correctement telle que [Tristán-Vega et al. (2009); Aganj et al. (2010)]:

$$\Psi_c(\theta_i, \phi_i) = \int_0^\infty P(r, \theta_i, \phi_i, \tau) r^2 dr, \tag{8}$$

avec Ψ_c que l'on appelle la dODF corrigée (cdODF). Ces deux définitions sont utilisées par certains modèles locaux et permettent de reconstruire une information intéressante concernant les orientations privilégiées par le phénomène de diffusion dans le cerveau.

De ces orientations, il est possible de déduire les trajectoires des faisceaux de fibres du cerveaux qui sont à l'origine de cette anisotropie de diffusion des protons. [Tournier et al. (2004)] a introduit la function de distribution des orientations des fibres (fODF) qui donne directement les trajectoires les plus probables de fibres cérébrales. Les trois fonctions dODF, cdODF et fODF peuvent être reconstruites en utilisant l'espace $\mathbf{q}$. Cet espace, dont le vecteur unitaire est $\mathbf{q} = \gamma\delta\mathbf{g}$, est l'espace dual de l'espace du propagateur de diffusion. Exprimé dans cet espace, le signal pondéré en diffusion $S(\mathbf{q},\tau)$ peut être relié à la transformée de Fourier du propagateur de diffusion:

$$\boxed{E(\mathbf{q},\tau) = \int_{\Re^3} P(\mathbf{r},\tau)\cdot e^{-2\pi i\mathbf{q}^{\mathrm{T}}\mathbf{r}}d\mathbf{r} = \mathcal{F}[P(\mathbf{r},\tau)]}, \tag{9}$$

avec $E(\mathbf{q},\tau) = S(\mathbf{q},\tau)/S_0$. De cette équation découlent plusieurs modèles locaux de diffusion [Callaghan (1991)].

Le modèle du tenseur de diffusion (DTI) suppose le propagateur de diffusion gaussien et en déduit l'expression suivante pour le signal pondéré en diffusion:

$$S(\mathbf{q},\tau) = S_0 e^{-\tau\mathbf{q}^{\mathrm{T}}\mathbf{D}\mathbf{q}}, \tag{10}$$

avec $\mathbf{D}$ le tenseur de diffusion [Basser et al. (1994)]. Le modèle DTI reconstruit ce tenseur de diffusion, qui permet de caractériser la diffusion des protons en 3D. De ce modèle, on peut calculer des cartes, comme celles de l'ADC et de l'anisotropie fractionnelle (FA). Une carte de FA peut être réalisée avec un codage en couleurs indiquant l'orientation des directions d'anisotropie en 3D (carte RGB) (figure 10).

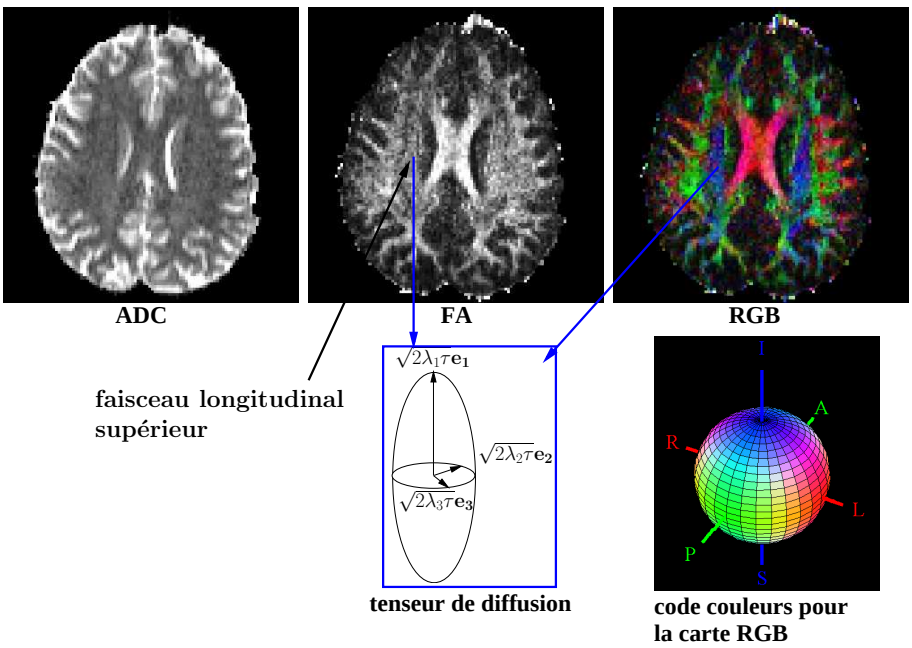


Figure 10: Cartes classiques du modèle DTI.

Cependant, ce modèle présente l'inconvénient de mal retranscrire les configurations de croisement de faisceaux de fibres à cause de l'hypothèse d'un processus gaussien de

diffusion ne pouvant représenter correctement qu'une seule population de fibres. C'est pourquoi, des modèles dits à haute résolution angulaire (HARDI) ont été développés pour mieux représenter ces configurations à plusieurs groupes de fibres.

Un de ces modèles HARDI, appelé Q-ball numérique [Tuch (2004)], repose sur un échantillonnage d'une sphère de l'espace **q**. Ce modèle travaille donc à une seule *b-value* et permet d'inférer les orientations privilégiées du processus de diffusion des protons en reconstruisant la dODF précédemment introduite. Cette méthode numérique repose sur une interpolation du signal qui peut dégrader la reconstruction de la dODF. Pour y remédier, le modèle du Q-ball analytique a été développé. Il repose sur une décomposition analytique du signal pondéré en diffusion sur une base d'harmoniques sphériques (HS) modifiées adaptée à la sphère de l'espace **q** [Frank (2002); Hess et al. (2006); Descoteaux et al. (2007)]: $\mathbf{E} = \mathbf{B}\mathbf{C}^{\mathrm{DW}}$, avec **E** le vecteur contenant le signal pondéré en diffusion, acquis pour différentes orientations de l'espace, normalisé par le signal pondéré en T_2. **B** est la matrice de la base des HS modifiées et $\mathbf{C}^{\mathrm{DW}}$ est le vecteur des coefficients du signal **E** sur la base. De ce modèle, on peut déterminer les coefficients $\mathbf{C}^{\mathrm{ODF}}$ de la décomposition de la dODF sur la base des HS modifiées. Ces coefficients permettent de reconstruire la dODF, donnant lieu à une des cartes montrées sur la figure 11. Il est également possible de reconstruire la cdODF, grâce au modèle analytique utilisant l'angle solide (modèle sa-aQBI) [Tristán-Vega et al. (2009); Aganj et al. (2010)].

Au lieu de reconstruire des indicateurs des orientations probables du processus de diffusion, on peut directement inférer les orientations des faisceaux de fibres en faisant l'hypothèse qu'un faisceau de fibres homogène présente une réponse impulsionnelle au processus de diffusion, représentée par un noyau gaussien. Il est alors possible de représenter la fODF par une déconvolution sphérique du signal pondéré en diffusion [Anderson and Ding (2002); Tournier et al. (2004); Jian and Vemuri (2007); Tournier et al. (2007)].

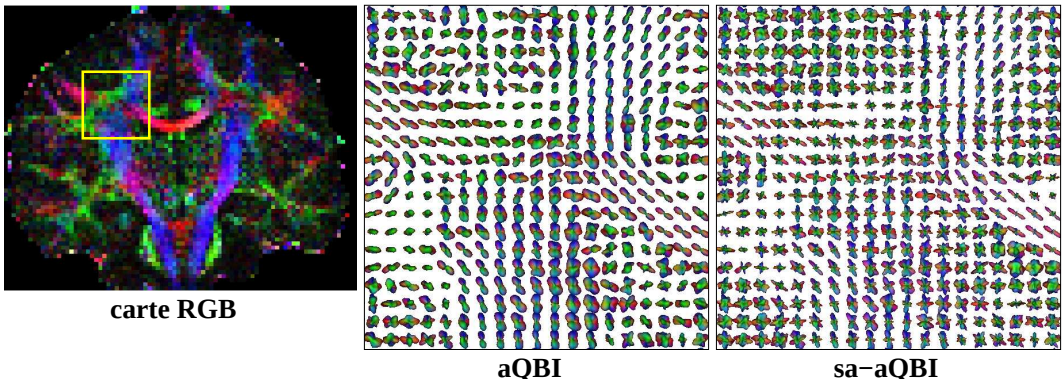


Figure 11: Carte d'ODFs à $b = 3000s \cdot mm^{-2}$ obtenues avec les modèles aQBI et sa-aQBI avec un ordre maximal d'HS L fixé à 8. Sur la gauche, la région d'intérêt utilisée est indiquée par un cadre jaune sur la carte RGB.

Techniques de tractographie.

Les techniques de tractographie construisent *in vivo* et de manière *non invasive* des chemins de fibres, appelés tractogrammes, représentant la connectivité anatomique cérébrale. Parmi les différentes classes de techniques de tractographie, celle dite déterministe par lignes de courant est la plus simple et plus rapide. Elle consiste à suivre la direction la plus probable indiquée par un modèle local de diffusion. Cette méthode n'est cependant pas capable de traduire des croisements de faisceaux. D'autres techniques, plus longues en temps de calcul, comme les classes de tractographies probabiliste et globale, proposent des tractogrammes plus proches de la réalité anatomique (figure 12).

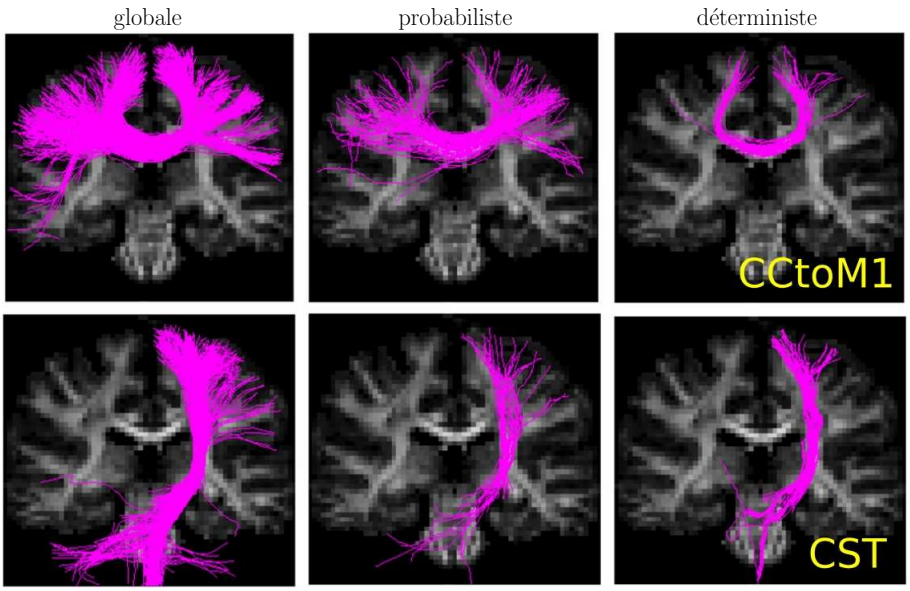


Figure 12: Comparison entres les techniques de tractographie globale, probabiliste et déterministe. La ligne du haut indique les tractogrammes reliant le corps calleux (CC) à la zone motrice primaire (CCtoM1). La ligne du bas indique les tractogrammes correspondant au faisceau cortico-spinal. Figure extraite de [Reisert et al. (2011)].

Chapitre 4: Etat de l'art de l'IRMd en temps réel (IRMdtr)

Introduction.

Le principe de l'IRMdtr est de réaliser le traitement d'un volume pondéré en diffusion immédiatement après son acquisition, avant l'acquisition du prochain volume. Cette approche, introduite par [Poupon et al. (2008b)], est innovante car le traitement des volumes acquis par les modèles locaux de diffusion se fait habituellement bien après l'acquisition (en post-traitement). La méthodologie d'IRMdtr nécessite que le traitement d'un volume ne dure pas plus longtemps que le temps entre l'acquisition de deux volumes, dit temps de répétition T_R. Un exemple de la mise en œuvre de ce procédé temps-réel (TR) avec le

modèle du DTI est montré sur la figure 13.

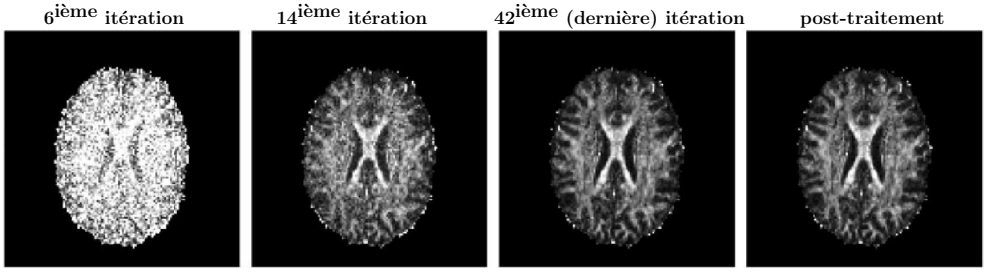


Figure 13: Résultat de l'IRMdtr sur une carte de FA (trois images les plus à gauche) et comparaison avec le résultat du post-traitement (dernière image à droite). La carte de FA est initialisée après six itérations. Puis, la carte est affinée à chaque nouvelle itération. À la $14^{\text{ième}}$ itération, la qualité visuelle de la carte est déjà bonne, alors que seulement 34% de l'acquisition a été effectué. Figure extraite de [Poupon et al. (2008b)].

Les motivations pour développer un traitement TR sont nombreuses. Pour le patient, ce procédé apporte plus de confort car l'examen peut être écourté si la qualité du résultat est suffisante avant la fin des itérations. De plus, la qualité des données étant vérifiée en TR, il y a nettement moins de gaspillage de données. À ce sujet, rappelons qu'aujourd'hui 75% des examens effectués dans le cadre d'études prénatales et néonatales sont jetés à cause des mouvements incontrôlés du sujet qui rendent les données illisibles. Un traitement TR peut permettre un contrôle fiable sur la qualité de l'examen. Du point de vue méthodologique, cette nouvelle technique d'imagerie en direct offre un cadre pour corriger des artéfacts en TR. Une alternative à une correction en TR est de redémarrer une acquisition si nécessaire. D'autre part, cette méthodologie TR rend possible la modification de paramètres d'acquisition ou de paramètres de traitement en fonction des résultats obtenus en direct. Cette possibilité d'optimisation du réglage des paramètres se fait grâce au feedback. Enfin, ce procédé d'imagerie en TR permet un diagnostic immédiat du médecin.

Dans le cadre particulier de l'IRMd, le TR permet de mieux répondre aux exigences de diagnostic rapide dans le cas d'accidents ischémiques. L'IRMdtr peut informer de l'état de santé des fibres cérébrales, ce qui peut orienter la décision médicale vers un traitement spécifique. D'un point de vue plus méthodologique, l'IRMdtr facilite aussi l'utilisation de certains modèles locaux qui demandent un long temps d'acquisition, avec beaucoup d'orientations de diffusion: en effet, le processus TR permet d'ajuster le choix d'orientations pour n'utiliser que le nombre d'itérations nécessaires lorsque le résultat apparaît suffisant.

Cadres incrémentaux.

Parmi les nombreux cadres incrémentaux possibles pour réaliser un traitement d'images d'IRMd en TR, [Poupon et al. (2008b)] se sont intéressés au filtre de Kalman (KF) [Kalman (1960)]. Ce filtre travaille à partir d'un modèle linéaire reliant des observations à un vecteur

d'état à estimer. Le KF permet de mettre à jour et ainsi affiner le vecteur d'état au fur et à mesure que des observations sont faites. Le modèle linéaire sur lequel repose le KF dans notre cadre d'IRMdtr peut se résumer par:

$$\boxed{\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}}, \text{ avec } \begin{cases} \mathbf{y}\text{: le vecteur de taille K}\times 1 \text{ des observations,} \\ \mathbf{A}\text{: la matrice de dessin de taille K}\times\text{N,} \\ \mathbf{x}\text{: le vecteur d'état de taille N}\times 1, \\ \boldsymbol{\epsilon}\text{: le vecteur de bruit de taille K}\times 1\text{, dont les composantes sont} \\ \text{distribuées selon } \mathcal{N}(0, \mathbf{R}). \end{cases} \tag{11}$$

Dans ce modèle, le vecteur $\mathbf{y}$ est rempli par une nouvelle observation à chaque itération i, i allant de 1 à K. $\mathbf{x}$ est le vecteur d'état que l'on cherche à estimer. Les composantes du vecteur $\boldsymbol{\epsilon}$ doivent être des variables aléatoires, indépendantes et distribuées selon une gaussienne à moyenne nulle et de matrice de covariance diagonale $\mathbf{R}$ (avec ses éléments diagonaux notés R_i). Le filtre de Kalman propose l'estimation de $\mathbf{x}$ grâce aux équations suivantes:

$$\begin{cases} \text{innovation: } \nu_i = y_i - \mathbf{a}_i \hat{\mathbf{x}}_{i-1}, \\ \text{covariance de l'innovation: } s_i = R_i + \mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathbf{T}}, \\ \text{gain: } \mathbf{k}_i = s_i^{-1} \cdot \mathbf{P}_{i-1} \mathbf{a}_i^{\mathbf{T}}, \\ \text{vecteur d'état à l'itération } i\text{: } \hat{\mathbf{x}}_i = \hat{\mathbf{x}}_{i-1} + \nu_i \mathbf{k}_i, \\ \text{covariance de l'erreur d'estimation à l'itération } i\text{: } \mathbf{P}_i = \mathbf{P}_{i-1} - \mathbf{k}_i \mathbf{a}_i \mathbf{P}_{i-1}, \end{cases} \tag{12}$$

avec $\mathbf{a}_i = [\mathbf{A}_{i1}, ..., \mathbf{A}_{iN}]$ la $i^{\text{ième}}$ ligne de la matrice de dessin $\mathbf{A}$. La figure 14 récapitule le fonctionnement du KF (qui est ici utilisé comme un filtre de Wiener).

Pour appliquer le KF, il reste à initialiser $\hat{\mathbf{x}}$ et $\mathbf{P}$, ce qui, en pratique, se fait souvent de la façon suivante: $\hat{\mathbf{x}}_0 = \mathbf{0}$ et $\mathbf{P}_0 = V\mathbf{I}$, avec V assez grand de manière à ce que le KF donne peu de poids à l'estimation initiale.

Adaptation des modèles DTI, aQBI et sa-aQBI au TR avec le filtre de Kalman.

Avant de décrire l'adaptation des modèles DTI, aQBI et sa-aQBI à une utilisation en TR grâce au filtre de Kalman, il est nécessaire de faire une première analyse du bruit en IRM. Pour un voxel et une orientation de diffusion donnés, on peut écrire la mesure du signal pondéré en diffusion telle que: $M = S + \epsilon$, avec S le signal sans bruit et ϵ le bruit d'acquisition qui est non-gaussien et fait que le signal bruité M suit une distribution ricienne [Henkelman (1985); Bernstein et al. (1989)] dans le cas d'une acquisition simple-canal, et une distribution χ non centrée dans le cas d'une acquisition multi-canaux [Constantinides et al. (1997)]. Le bruit ϵ ne peut alors pas être considéré comme un bruit

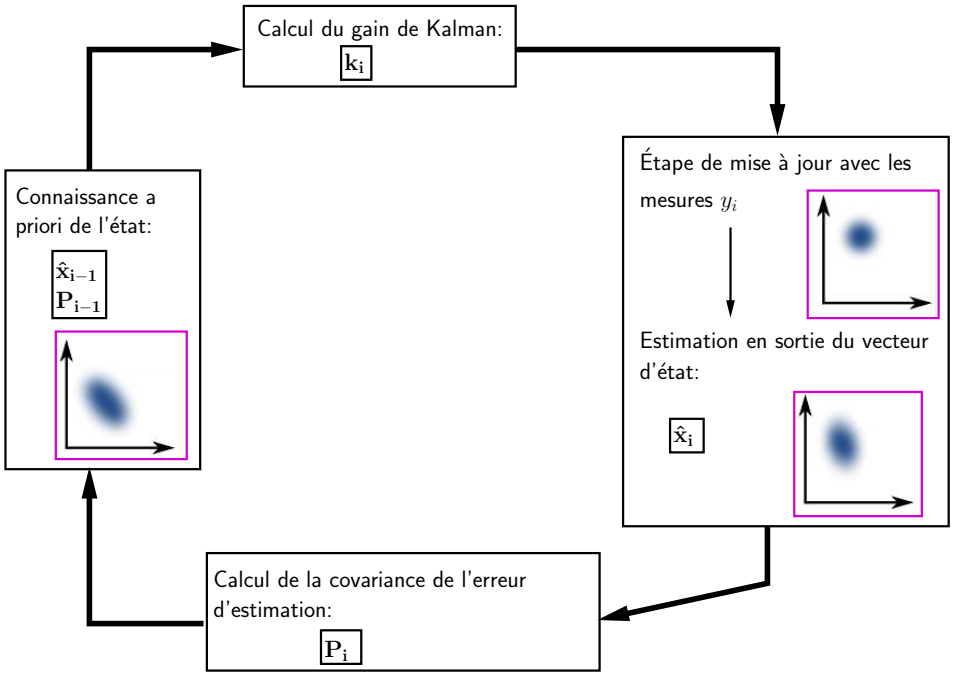


Figure 14: Schéma d'un KF utilisé comme un filtre de Wiener. Figure adaptée de en.wikipedia.org et de ccar.colorado.edu. L'estimation de l'état est affinée à chaque nouvelle itération i.

gaussien à moyenne nulle, en particulier plus le rapport signal à bruit est faible (SNR $\leq$ 4) et plus il y a de canaux utilisés.

Cependant, comme première approximation, les modèles DTI, aQBI et sa-aQBI ont été adaptés pour le TR en supposant un bruit gaussien à moyenne nulle pour respecter le cadre d'utilisation du KF. Pour le DTI, [Poupon et al. (2008b)] ont proposé le modèle linéaire suivant, de la forme $\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}$:

$$\boxed{ln\left(\frac{S_0}{M(\mathbf{o}_i)}\right) = b_i \mathbf{o}^T \mathbf{D} \mathbf{o}_i - ln\left(1 + \frac{\mu_i}{S_0} e^{b_i \mathbf{o}_i^T \mathbf{D} \mathbf{o}_i}\right)}, \tag{13}$$

dans lequel $\mathbf{o}_i$ est l'orientation de diffusion testée à l'itération i et μ_i est le bruit d'acquisition à l'itération i. En utilisant l'équation 13, le vecteur d'état à estimer par le KF correspond aux six coefficients du tenseur de diffusion: $\mathbf{x} = [D_{xx}, D_{xy}, D_{xz}, D_{yy}, D_{yz}, D_{zz}]^T$.

De manière similaire, on peut écrire deux formes linéaires pour utiliser le KF sur le modèle aQBI. La première est:

$$\boxed{\mathbf{M_E} = \mathbf{B} \cdot \mathbf{C}^{\mathrm{DW}} + \boldsymbol{\epsilon}}, \tag{14}$$

où $\mathbf{M_E} = [M(0)/S_0, ..., M(K)/S_0]^T$ correspond au vecteur des signaux pondérés en diffusion normalisés par le signal pondéré en T_2. $\mathbf{C}^{\mathrm{DW}}$ est le vecteur des coefficients du signal sans bruit $\mathbf{E}$ et correspond dans le cadre de Kalman au vecteur d'état à estimer en TR. Enfin, $\boldsymbol{\epsilon}$ représente le vecteur du bruit normalisé par le signal pondéré en T_2. La deuxième forme linéaire que l'on peut utiliser est la suivante [Poupon et al. (2008b); Deriche et al.

(2009)]:

$$\boxed{\mathbf{M_E} = \left(\mathbf{BP}^{-1}\right)\mathbf{C}^{\mathrm{ODF}} + \boldsymbol{\epsilon},} \tag{15}$$

dans laquelle le vecteur d'état est le vecteur des coefficients de la dODF permettant de calculer la dODF.

Enfin, on peut écrire une forme linéaire également pour le modèle sa-aQBI [Caruyer et al. (2010)]:

$$\boxed{ln\left(-ln\mathbf{M_E}\right) = \mathbf{B}\cdot\mathbf{C}^{\mathrm{SA}} + \boldsymbol{\epsilon}}, \tag{16}$$

où $\mathbf{C}^{\mathrm{SA}}$ est le vecteur des coefficients de $ln\left(-ln\mathbf{E}\right)$ sur la base d'HS modifiée. $\boldsymbol{\epsilon}$ est un vecteur contenant $ln\left(-ln\mu_i\right)/S_0\ \forall i \in [\![1;K]\!]$, avec μ_i le bruit d'acquisition à l'itération i. Dans ce modèle, le vecteur d'état est donc $\mathbf{C}^{\mathrm{SA}}$. À partir de ce vecteur, on peut en déduire le vecteur des coefficients de la cdODF que vise à reconstruire le modèle sa-aQBI:

$$\begin{cases} \hat{\mathbf{C}}^{\mathrm{ODF}_c}(j) = \frac{1}{2\sqrt{\pi}} & \text{si } j = 1, \\ \hat{\mathbf{C}}^{\mathrm{ODF}_c}(j) = \frac{1}{16\pi^2}\mathbf{P}\cdot\mathbf{N}\cdot\hat{\mathbf{C}}^{\mathrm{SA}}(j) & \text{si } j > 1. \end{cases} \tag{17}$$

Pour les modèles aQBI et sa-aQBI, il est possible d'intégrer une régularisation dans le processus. Cette régularisation doit seulement apparaître lors de l'initialisation de la matrice de covariance $\mathbf{P}$, telle que: $\mathbf{P}_0 = ((1/V)\mathbf{I} + \lambda\mathbf{L})^{-1}$, avec $V = 10^6$, $\lambda = 0.006$ le facteur de régularisation et $\mathbf{L}$ la matrice de Laplace-Beltrami.

Dans chacune de ces formes linéaires dédiées au KF, le bruit est considéré comme un bruit gaussien à moyenne nulle, ce qui peut conduire à des estimations erronnées des vecteurs d'état, en particulier à de faibles SNR et/ou pour un nombre de canaux utilisés élevé.

Contraintes du TR.

Lorsqu'on veut réaliser le traitement et l'analyse des signaux pondérés en diffusion en TR, le procédé sera optimal si chaque itération apporte une information nouvelle et complémentaire à celles obtenues lors des précédentes itérations. Ce que l'on souhaite, c'est avoir dès les premières itérations une information de base cohérente, puis l'affiner au fur et à mesure des itérations. Pour appliquer ce principe à l'IRMdtr, il est nécessaire d'avoir un jeu d'orientations de diffusion à tester qui soit approximativement uniforme, non seulement à la fin de toutes les itérations, mais également aux itérations intermédiaires, dans le cas où l'examen serait arrêté avant la fin de l'acquisition. Parmi les algorithmes proposant de générer ces jeux d'orientations, nous avons choisi de travaillé avec le modèle de [Dubois et al. (2006)], dont une application est montrée à la figure 15.

Pour réaliser un traitement de données d'IRMd en TR, il est absolument nécessaire d'utiliser des outils logiciels efficaces et rapides. La figure 16 indique en rouge l'architecture à haute performance de calcul déployée pour le TR et qui permet de traiter les données acquises par l'imageur et reconstruites par l'unité de reconstruction, puis de les renvoyer

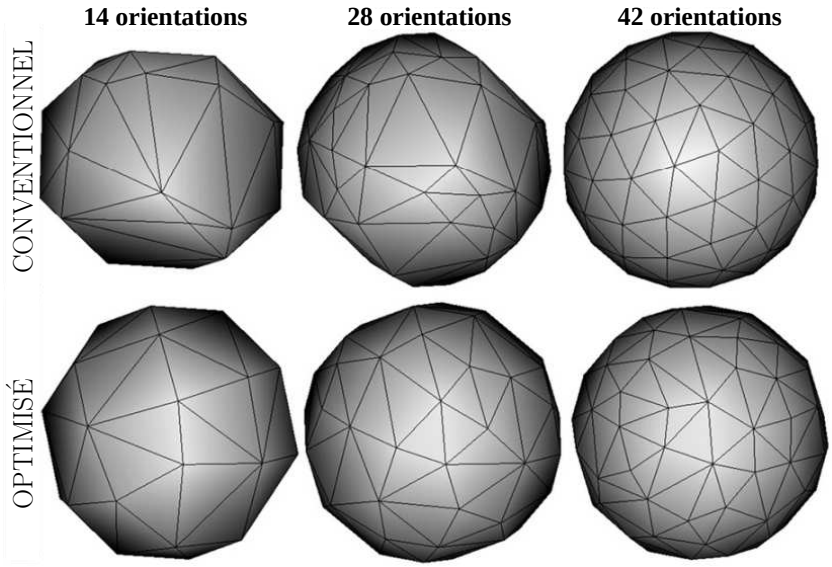


Figure 15: Comparaison entre des jeux d'orientations conventionnels (en haut) et optimisés (en bas) pour une séquence avec 42 orientations au total. Les jeux optimisés intermédiaires (à la $14^{\text{ième}}$ et $28^{\text{ième}}$ itérations) sont plus uniformes que les jeux conventionnels correspondants. Figure extraite de [Poupon et al. (2008b)].

vers la console. Le système de calcul utilise un cluster de 80 processeurs (ou CPUs) pour arriver à une réduction du temps de traitement d'un volume qui doit être obtenu en-dessous du T_R. De plus, cette architecture dédiée au TR permet de réaliser un feedback sur des paramètres de séquence par exemple.

Preuve de concept des modèles adaptés au TR.

Le modèle DTI adapté au TR a été validé sur des données réelles à $b = 3000s \cdot mm^{-2}$ comme le montre la figure 17. De même, les modèles aQBI et sa-aQBI ont été validés sur les mêmes données. La figure 18 montre les résulats obtenus avec un ordre maximal des HS fixé à 6 et avec l'utilisation de la régularisation dans l'initialisation de la matrice de covariance $\mathbf{P}$. Les temps de traitements ont été mesurés avec 1 CPU et en utilisant le cluster de 80 CPUs. Si les temps dépassent parfois le T_R en utilisant un seul CPU, ils tombent bien en-dessous du T_R avec le cluster.

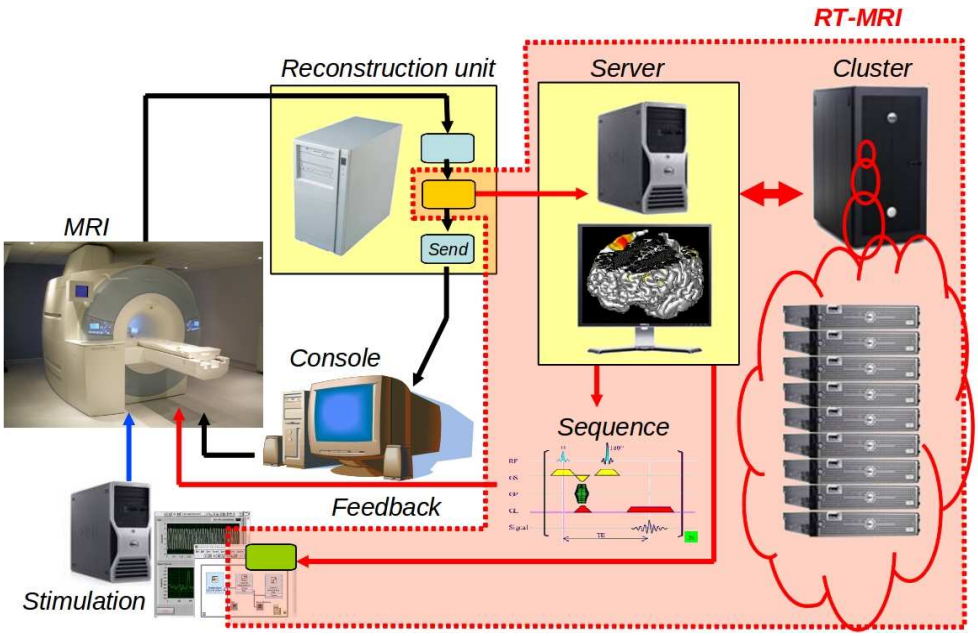


Figure 16: L'architecture hardware adaptée au projet TR. Le bloc rouge "RT-MRI" constitue la partie hardware prenant en charge le traitement TR. Avant que l'unité de reconstruction envoie les données à la console, ces données sont envoyées vers un serveur qui commande l'exécution des algorithmes de traitement TR sur les données. Pour réduire le temps de traitement, ces algorithmes sont parallélisés et distribués sur un cluster de 80 CPUs. De plus, un feedback peut être réalisé sur les paramètres de séquence ou sur les paramètres de contrôle des stimuli dans le cas de l'IRMf TR. Figure extraite de [Poupon and Riff (2009)].

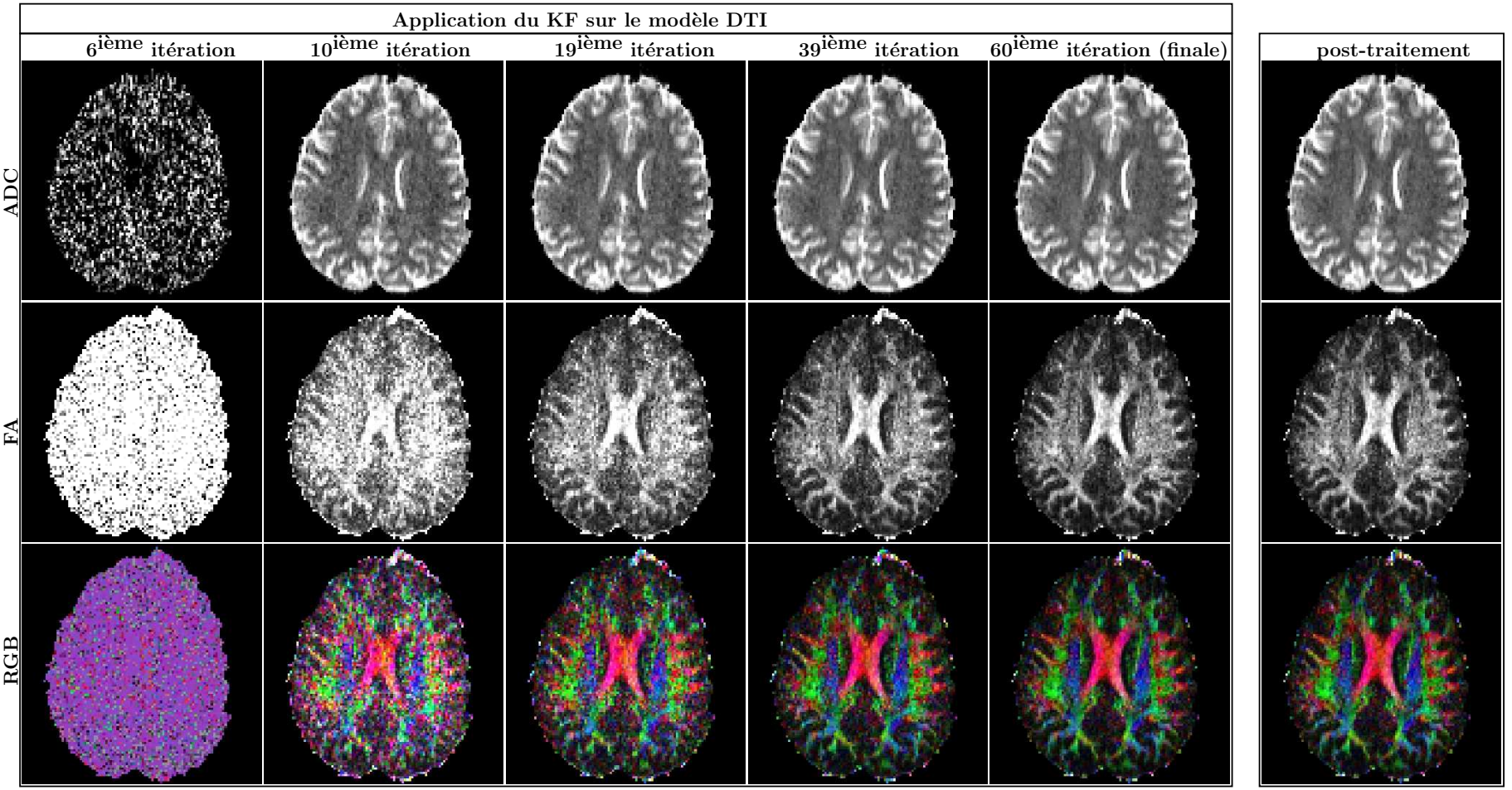


Figure 17: Application du KF sur le modèle DTI.

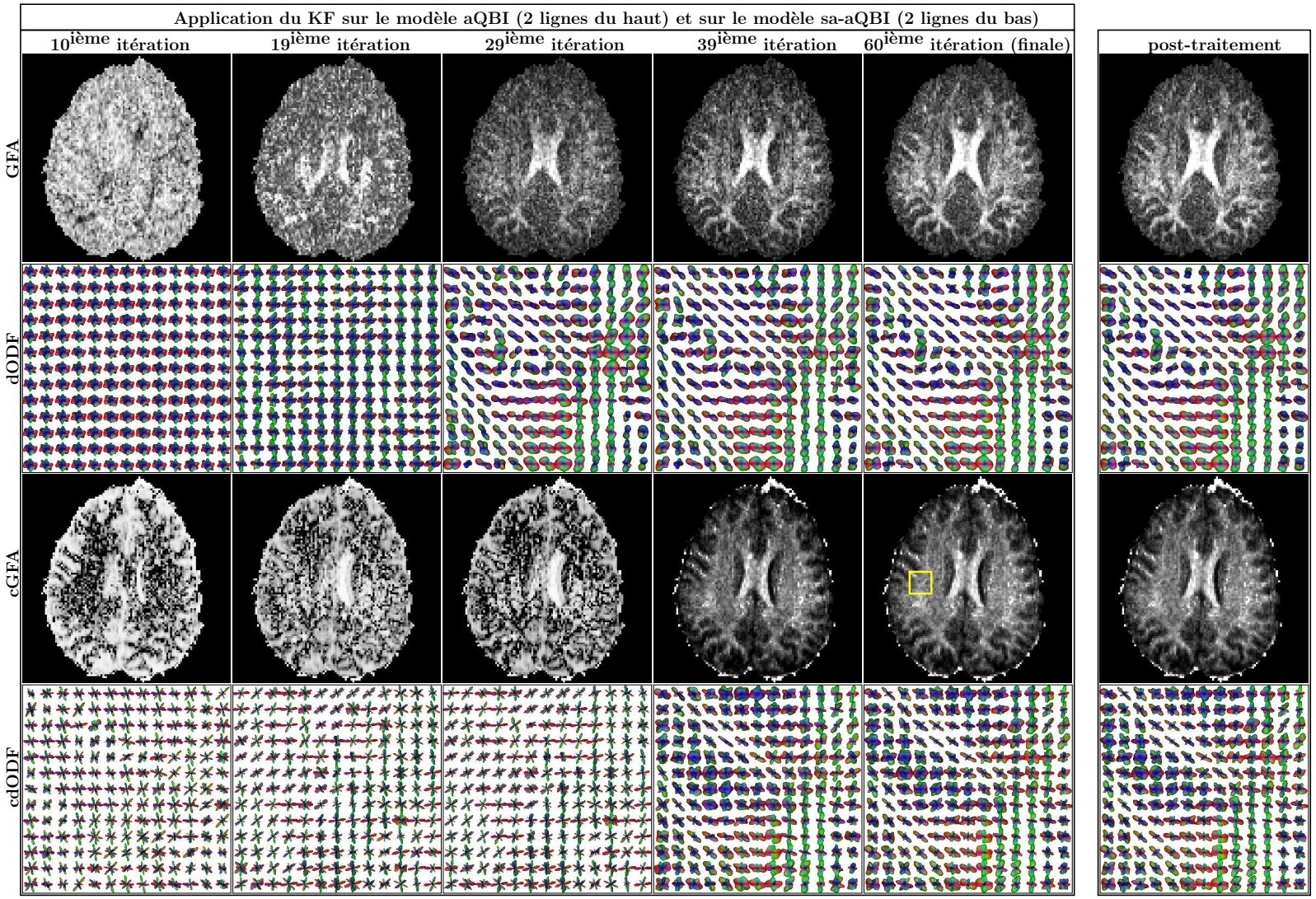


Figure 18: Application du KF sur les modèles aQBI et sa-aQBI. L'ordre maximal des HS était fixé à 6 et la régularisation a été utilisée. La région dans laquelle sont montrées les ODFs est indiquée par un cadre jaune sur la carte RGB de la quatrième colonne.

Chapitre 5: Analyse du bruit en IRM

Origines du bruit en IRM.

Le bruit en IRM provient principalement de l'agitation thermique des porteurs de charge dans le circuit de réception du signal d'IRM. On distingue le bruit provenant des tissus conducteurs du sujet se trouvant dans l'imageur et le bruit provenant de l'antenne de réception du signal. En IRM parallèle, des corrélations entre les signaux de bruit provenant des différents canaux existent [Roemer et al. (1990); Hayes and Roemer (1990); Harpen (1992); Redpath (1992); Brown et al. (2007)]. Elles sont impossibles à éliminer et doivent être prises en compte dans les modèles du bruit en IRM.

Le bruit qui corrompt le signal complexe mesuré en IRM par un canal de réception c possède une composante réelle ϵ_{r_c} et une composante imaginaire ϵ_{i_c}. Pour chaque canal de réception, ces deux composantes sont supposées suivre chacune une distribution gaussienne à moyenne nulle, et ces deux distributions ne sont pas corrélées l'une avec l'autre [Henkelman (1985)]. En IRM, il est courant de travailler sur l'amplitude M du signal mesuré qui est reconstruite par la technique de la somme des carrés (SoS):

$$\boxed{M = \sqrt{\sum_{c=1}^{n} \left[(S_{r_c} + \epsilon_{r_c})^2 + (S_{i_c} + \epsilon_{i_c})^2 \right]}}, \tag{18}$$

avec S_{r_c} et S_{i_c} les parties réelle et imaginaire, respectivement, du signal sans bruit complexe mesuré par le canal c. n est le nombre de canaux utilisés lors de l'acquisition. Les parties réelles et imaginaires des signaux mesurés peuvent éventuellement être accessibles à la fin de l'examen, mais la chaîne de reconstruction du fabricant ne permet pas ou difficilement d'y accéder pendant l'examen. D'autre part, il est plus habituel de travailler avec l'amplitude du signal qui est exempte d'artéfacts de phase [Henkelman (1985); Constantinides et al. (1997); Nowak (1999)]. À cause du calcul mathématique pour obtenir l'amplitude du signal, le bruit final ϵ qui entâche l'amplitude sans bruit $S = \sqrt{\sum_{c=1}^{n} [S_{r_c} + S_{i_c}]}$ n'est plus un bruit gaussien à moyenne nulle. C'est un bruit qui dépend du signal sans bruit S.

Dans le cas d'une acquisition à un canal, le bruit est dit ricien: l'amplitude M suit une fonction densité de probabilité (PDF) ricienne, définie pour $M \geq 0$ et $S \geq 0$, par [Bernstein et al. (1989); Rice (1952)]:

$$\underline{\text{PDF ricienne:}} \quad \boxed{p(M; S, \sigma) = \frac{M}{\sigma^2} . \exp\left(-\frac{M^2 + S^2}{2\sigma^2} \right) \cdot I_0 \left(\frac{S \cdot M}{\sigma^2} \right)}, \tag{19}$$

avec σ l'écart-type du bruit gaussien présent sur les parties réelle et imaginaire du canal de réception et I_0 la fonction de Bessel modifiée de première espèce et d'ordre 0. Cette distribution est tracée sur la figure 5.2. La PDF de ϵ se distingue d'une gaussienne à moyenne nulle en particulier à de faibles SNR. Un biais ricien apparaît alors et l'image bruitée contient une composante non nulle due au bruit dans ce cas.

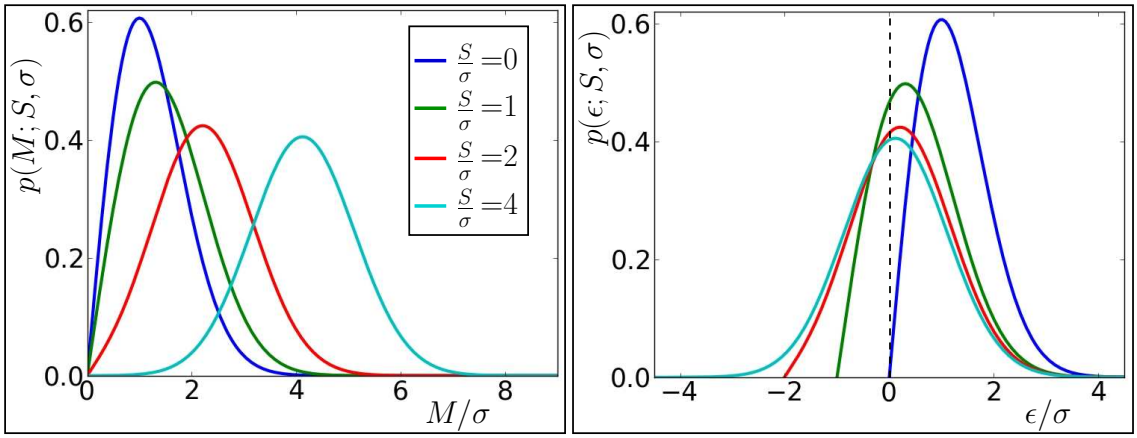


Figure 19: PDF de M (à gauche) et de ϵ (à droite) dans le cas d'une acquisition simple-canal pour plusieurs valeurs de S/σ. Plus le rapport S/σ est élevé, plus la PDF de ϵ ressemble à une gaussienne à moyenne nulle. La légende des couleurs indiquée à gauche s'applique également pour les courbes de droite.

Dans le cas d'une acquisition à plusieurs canaux, s'il n'y a pas de sous-échantillonnage de l'espace **k**, ni de corrélation entre les canaux et que la variance est la même pour chaque canal, alors l'amplitude M suit une distribution χ non centrée (χ-nc), définie pour $M \geq 0$ et $S \geq 0$, par [Constantinides et al. (1997)]:

$$\underline{\text{PDF } \chi\text{-nc:}} \quad \boxed{p(M;S,\sigma,n) = \frac{S}{\sigma^2}\left(\frac{M}{S}\right)^n \exp\left(-\frac{M^2+S^2}{2\sigma^2}\right) \cdot I_{n-1}\left(\frac{S\cdot M}{\sigma^2}\right)}, \tag{20}$$

avec σ l'écart-type du bruit gaussien présent sur les parties réelle et imaginaire de chaque canal de réception et n le nombre de canaux. I_{n-1} est la fonction de Bessel modifiée de première espèce et d'ordre $n-1$. Cette PDF est une généralisation à $n > 1$ de la distribution ricienne. Elle est indiquée sur la figure 20. Comme précédemment, le même phénomène de biais apparaît à faible SNR, ainsi que pour de grandes valeurs de n.

Il y a, en pratique, des corrélations entre les canaux, ce qui change la distribution théorique du bruit. [Aja-Fernández and Tristán-Vega (2012)] ont proposé de prendre en compte ces corrélations en calculant un nombre effectif de canaux n_{eff} et une variance effective σ^2_{eff} dépendant tous deux de la position **v** du voxel considéré. En remplaçant n et σ par ces deux paramètres, la distribution χ-nc (équation 20) est alors une bonne approximation de la vraie distribution de M dans l'image.

Si, en plus, une technique de reconstruction GRAPPA est utilisée, alors la non-stationnarité du bruit est plus conséquente. [Aja-Fernández et al. (2011)] ont proposé de calculer n_{eff} et σ^2_{eff} en tenant compte des paramètres de reconstruction GRAPPA. Ces paramètres injectés dans la distribution χ-nc conduisent à une bonne approximation de la réalité expérimentale. Dans le cas d'une reconstruction SENSE, le bruit n'est également pas stationnaire. La PDF suivie par M peut être approximée par une ricienne contenant une variance voxel-dépendante [Dietrich et al. (2008a); Rajan et al. (2012b)].

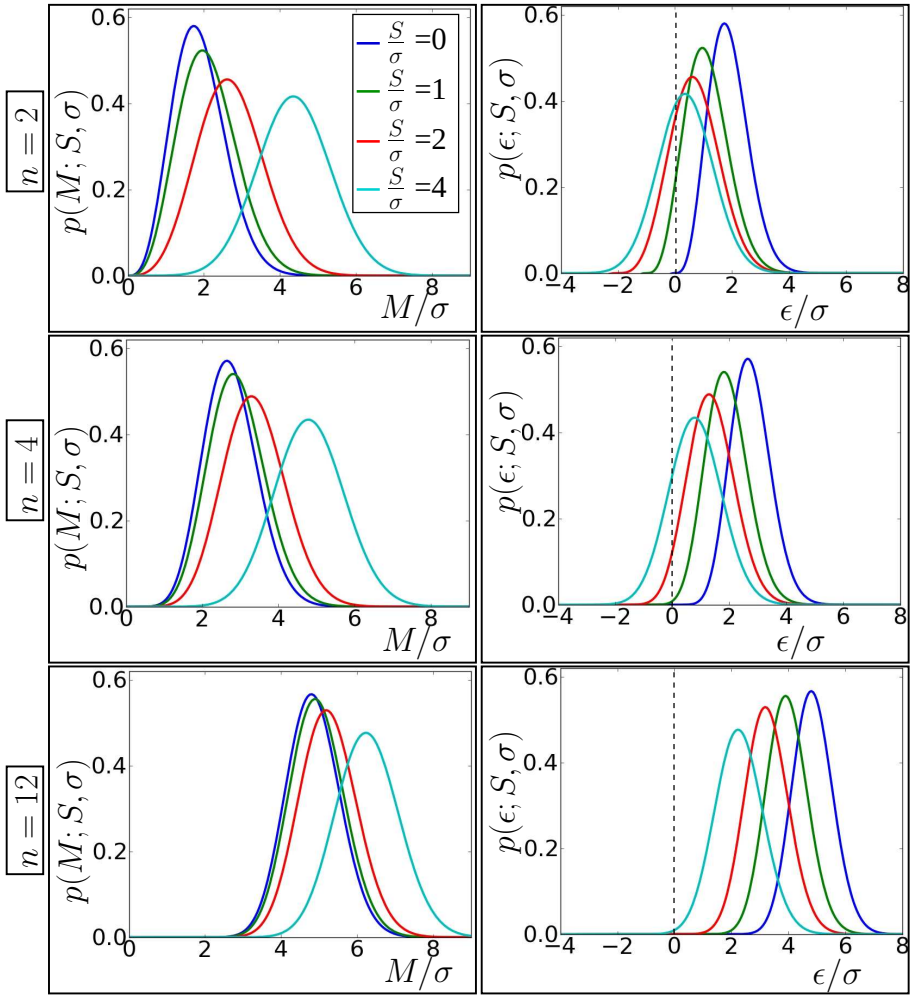


Figure 20: PDF de M (à gauche) et de ϵ (à droite) dans le cas d'une acquisition multi-canaux respectant les trois conditions suivantes: **pas de sous-échantillonnage, pas de corrélation entre les canaux, et la même variance pour chaque canal**. Les distributions sont représentées pour $n = 2/4/12$ et pour plusieurs valeurs de S/σ.

En IRMd, le signal peut facilement être noyé dans le bruit car le SNR est souvent faible, en particulier à des valeurs de b élevées ($b > 3000 s/mm^2$). Nous avons simulé un signal pondéré en diffusion en 3D correspondant à un croisement de deux faisceaux de fibres, et nous avons reconstruit la dODF correspondante. Nous avons comparé ces deux résultats sans et avec addition de bruit χ-nc à différents SNR pour $n = 4$. La figure 21 montre que le bruit dégrade la résolution angulaire. Le biais du bruit fait se rétrécir les lobes des dODF. Enfin, le bruit crée de faux pics pouvant faire croire à des directions du phénomène de diffusion à prendre en compte.

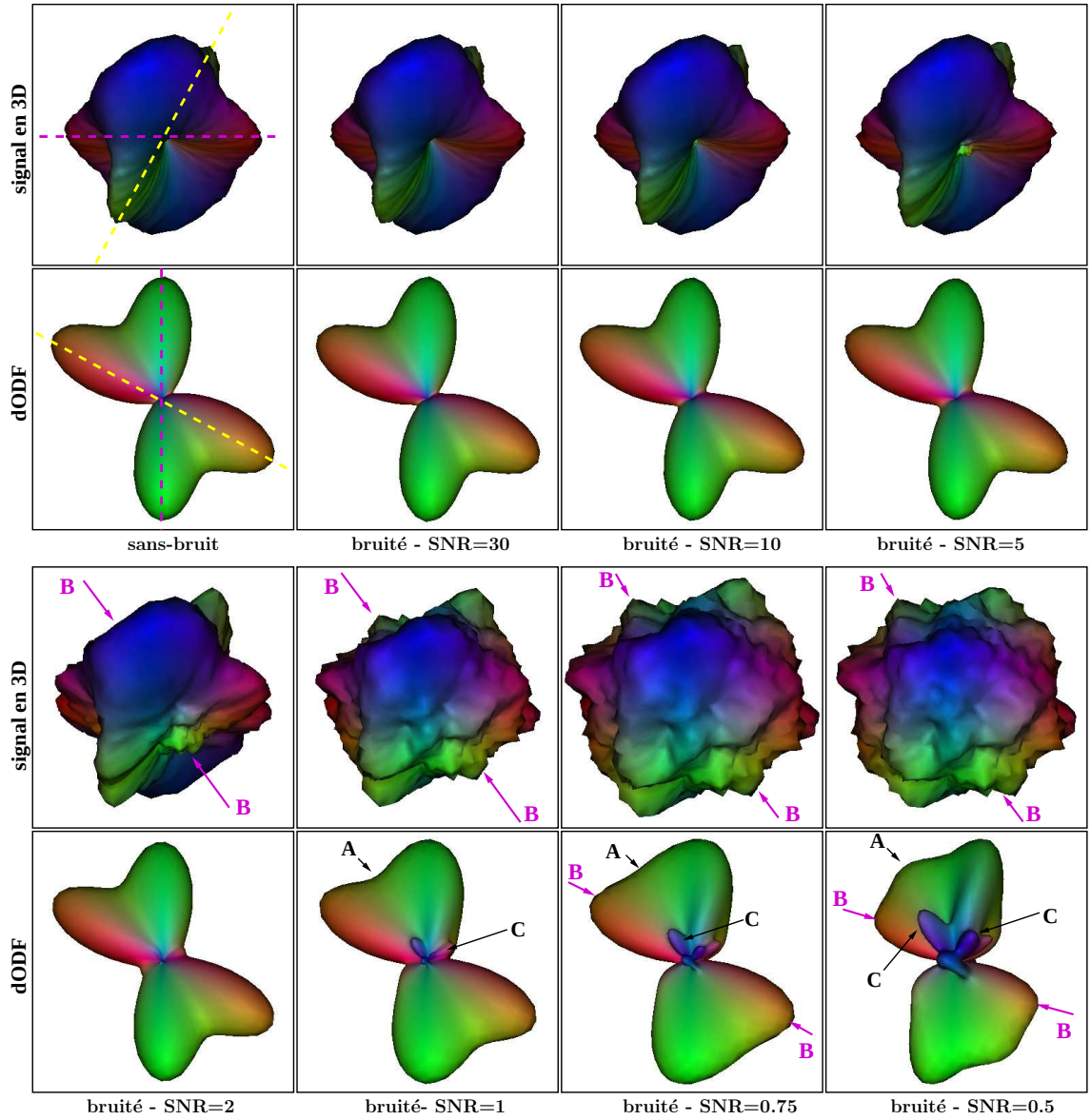


Figure 21: Simulations de signaux pondérés en diffusion en 3D sans bruit et bruités à différents SNR avec leurs dODFs correspondantes. Le SNR a été calculé à partir du signal sans bruit pondéré en diffusion ($S = 7$), et non pas à partir du signal sans bruit pondéré en T_2 ($S_0 = 200$). La lettre A indique une perte de résolution angulaire. La lettre B montre l'effet du biais qui fait se rétrécir les lobes des dODFs. La lettre C indique les pics erronés créés par le bruit sur les dODFs. Les simulations ont été réalisées à $b = 6000s \cdot mm^{-2}$, $n = 4$ et avec σ allant de 0.23 à 14.

Méthodes d'estimation du bruit en IRM.

Parmi les méthodes d'estimation du bruit en IRM, la plupart reposent sur une estimation de σ qui s'effectue dans le fond de l'image (zone sans signal utile) [Aja-Fernández et al. (2009)]. La variance obtenue est alors considérée comme valable pour tous les voxels du volume. Une méthode, reposant également sur l'analyse du fond, a été proposée par [Aja-Fernández et al. (2013)] pour tenir compte des corrélations et estimer σ_{eff} et n_{eff}

en chaque voxel du volume pour une reconstruction SoS sans sous-échantillonnage, en supposant que chaque canal a la même variance et que les signaux sans bruit reçus par les différents canaux sont les mêmes. Enfin, d'autres techniques s'intéressent au cas où le volume ne possède pas de fond [Coupé et al. (2010); Rajan et al. (2010)]. Nous avons étendu l'une de ces méthodes [Rajan et al. (2010)], dédiée à l'origine à un bruit ricien, au cas d'un bruit χ-nc [Brion et al. (2011b)]. Cette méthode estime une variance globale pour tout le volume. Elle repose sur l'analyse de la skewness dans l'image pour choisir dans une LUT (lookup table) un facteur de correction à appliquer dans l'estimation de σ faite à partir d'une estimation supposant un bruit gaussien. Notre technique donne des résultats similaires à ceux obtenus par la technique de [Rajan et al. (2010)] dans le cas d'un bruit ricien. Elle peut servir de première approximation lorsqu'il n'y a pas de fond.

Indices de qualité d'images.

Outre les indices classiques d'erreur quadratique moyenne (MSE), de SNR et de contraste signal à bruit (CNR), nous nous sommes intéressés à développer deux indices spécifiques à l'IRMd:

1. Une MSE spécifique aux modèles DTI, aQBI et sa-aQBI qui repose sur le vecteur d'état $\mathbf{x}$ qui peut être le tenseur de diffusion $\mathbf{D}$ en DTI, le vecteur des coefficients $\mathbf{C}^{\mathrm{DW}}$ en aQBI et le vecteur $\mathbf{C}^{\mathrm{SA}}$ en sa-aQBI. Cet indice se calcule ainsi:

$$\mathrm{MSE} = \frac{1}{N_v} \sum_{\mathbf{v} \in \mathcal{V}} \sum_{j=1}^{N} \left(\tilde{x}(j, \mathbf{v}) - x(j, \mathbf{v}) \right)^2 , \tag{21}$$

   avec $\tilde{x}(j, \mathbf{v})$ le $j^{\text{ième}}$ coefficient du vecteur bruité $\tilde{\mathbf{x}}$ calculé sur les données bruitées. $x(j, \mathbf{v})$ est le $j^{\text{ième}}$ coefficient du vecteur sans bruit $\mathbf{x}$. Dans l'équation 21, $\mathcal{V}$ est le volume considéré, N_v correspond au nombre de voxels de ce volume, et N est le nombre de coefficients contenus dans le vecteur d'état. Cet indice MSE requiert l'utilisation d'une référence sans bruit et ne peut donc être utilisé que sur des données simulées.

2. Un rapport de FA, GFA ou cGFA se calculant ainsi (pour la GFA ici):

$$\text{GFA ratio} = \frac{\mathrm{GFA_A}}{\mathrm{GFA_B}}, \tag{22}$$

   avec $\mathrm{GFA_A}$ et $\mathrm{GFA_B}$ des moyennes de GFA calculées sur la carte de GFA en deux régions A et B, respectivement. Cet indice a l'avantage de pouvoir être appliqué sans nécessité d'image de référence sans bruit. Il peut aussi être appliqué sur une image pour laquelle on ne connaît pas le modèle de bruit (typiquement une image filtrée). Ce calcul de rapport correspond à une mesure du contraste entre deux régions d'intérêt. La région A est choisie avec un niveau uniforme d'anisotropie moyenne et la région B avec une niveau uniforme d'anistropie faible. Ainsi, plus le bruit est élevé, plus l'indice est faible.

Chapitre 6: La correction du bruit en TR

Contraintes provenant du TR.

Parmi la multitude de méthodes de correction du bruit, on peut distinguer plusieurs groupes répertoriés à la figure 22.

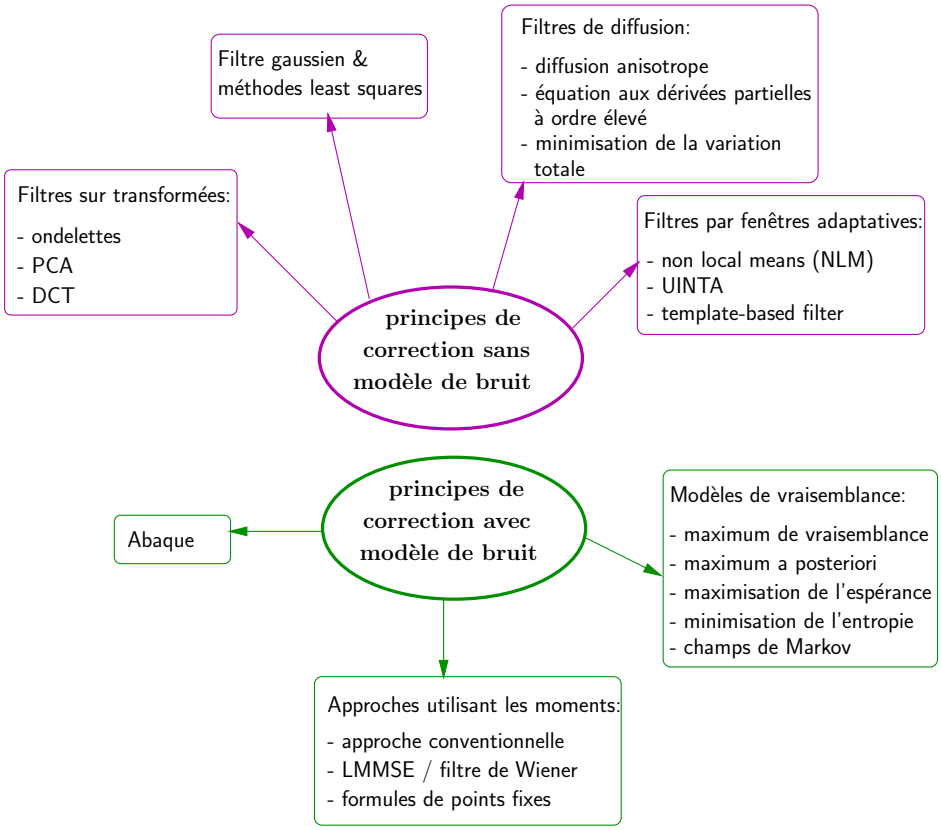


Figure 22: Principaux groupes de méthodes de correction du bruit avec une application aux images d'IRM. Les méthodes peuvent être distinguées suivant qu'elles tiennent compte ou non du caractère ricien ou χ-nc du bruit en IRM.

Notre objectif de correction en TR pour l'IRMdtr impose de respecter certaines conditions obligatoires:

1. la méthode ne doit pas nécessiter d'avoir la connaissance de tous les volumes pondérés en diffusion à la fois,
2. la méthode doit pouvoir s'appliquer à autant de modèles locaux de diffusion que possible,
3. la méthode ne doit pas nécessiter plus de mesures que celles fournies par un protocole clinique habituel,
4. la méthode doit s'exécuter en moins de temps que le T_R (généralement égal à une dizaine de secondes).

Ces conditions nous ont conduits à éliminer de notre choix les méthodes reposant sur une correction utilisant l'ensemble des volumes pondérés en diffusion acquis lors de l'examen [Fillard et al. (2007); Clarke et al. (2008); Brion et al. (2009); Martin-Fernández et al. (2009); Tristán-Vega and Aja-Fernández (2010)]. Nous avons également écarté les techniques spécifiques à un certain modèle local de diffusion [Chefd'Hotel et al. (2002); Wang et al. (2004b); Chang et al. (2005); Salvador et al. (2005); Koay et al. (2006); Assemlal et al. (2007); Caan et al. (2010); Tristán-Vega et al. (2012)]. Nous avons exclu les méthodes nécessitant des mesures supplémentaires à celles effectuées lors d'un protocole usuel d'IRMd [Koay and Basser (2006); Koay et al. (2009a)]. Enfin, la quatrième condition nous a invités à favoriser une technique rapide, compatible avec une distribution sur un cluster de CPUs. Notre choix s'est porté sur l'estimateur linéaire qui minimise l'erreur quadratique moyenne (LMMSE) de [Aja-Fernández et al. (2008a,b)], qui a été à l'origine développé pour le bruit ricien. Nous avons étendu ce LMMSE à la correction d'un bruit χ-nc pour traiter des images acquises en IRM parallèle avec la reconstruction GRAPPA.

Correction avec un estimateur linéaire qui minimise l'erreur quadratique moyenne (LMMSE).

Le LMMSE de [Aja-Fernández et al. (2008a)] adapté à un bruit ricien s'écrit sous la forme:

$$\boxed{\hat{S}^2 = \underbrace{\langle M^2 \rangle - 2\sigma^2}_{J_R} + \underbrace{\left(1 - \frac{4\sigma^2[\langle M^2 \rangle - \sigma^2]}{\langle M^4 \rangle - \langle M^2 \rangle^2}\right)}_{K_R} \times \left(M^2 - \langle M^2 \rangle\right)}, \tag{23}$$

avec $\langle \cdot \rangle$ l'opérateur de moyenne spatiale effectué localement sur un voisinage de voxels. Pour utiliser l'équation 23, il faut estimer l'écart-type du bruit σ. Le terme noté J_R revient à appliquer une moyenne et à en soustraire le biais dû au bruit ricien. Le terme K_R est un terme d'attache aux données qui module l'effet de J_R dans les régions hétérogènes pour conserver les détails des contours.

Le LMMSE adapté au bruit χ-nc s'écrit sous la forme [Brion et al. (2011b,c,a)]:

$$\boxed{\hat{S}^2 = \langle M^2 \rangle - 2n\sigma^2 + \left(1 - \frac{4\sigma^2\left[\langle M^2 \rangle - n\sigma^2\right]}{\langle M^4 \rangle - \langle M^2 \rangle^2}\right) \times \left(M^2 - \langle M^2 \rangle\right)}, \tag{24}$$

avec n le nombre de canaux. Lorsque $n = 1$, on retrouve l'équation 23.

Pour tenir compte des corrélations entre canaux en IRM parallèle, nous proposons une méthode estimant les paramètres effectifs n_{eff} et σ_{eff}. Notre technique repose sur une estimation valable pour tout le volume, et non voxel à voxel, ce qui permet un traitement rapide. Notre détermination de ces deux paramètres est empirique: il s'agit de tester différentes valeurs pour n_{eff}, puis de calculer les σ_{eff} correspondants tels que:

$$\underline{\text{ModeM1-}\chi\text{:}} \quad \boxed{\hat{\sigma}_{eff} = \left(\sqrt{2}(n_{eff})^{(1/2)}\right)^{-1} \text{mode}\left(\langle M_{bg}(\mathbf{v}) \rangle\right)}, \tag{25}$$

où $M_{bg}(\mathbf{v})$ est l'amplitude mesurée au voxel $\mathbf{v}$ dans le fond du volume et mode $(\langle M_{bg}(\mathbf{v})\rangle)$ est le mode de la distribution de la moyenne locale de $M_{bg}(\mathbf{v})$; $(n_{eff})^{(1/2)}$ est calculé en appliquant la définition du symbole de Pochhammer. Une fois σ_{eff} calculé pour différentes valeurs de n_{eff}, nous testons le LMMSE avec ce panel de couples de paramètres et choisissons le couple $(n_{eff};\ \sigma_{eff})$ qui produit le rapport de GFA —comme défini dans la sous-section — le plus élevé.

Résultats & discussion.

Nous avons appliqué le LMMSE original ("Rice LMMSE") et notre extension au bruit χ-nc ("nc-χ LMMSE"), avec un voisinage spatial de 5×5×5 voxels, sur nos données simulées à $b = 4500s \cdot mm^{-2}$ bruitées avec un bruit χ-nc stationnaire avec $n = 4$ et $\sigma = 20$. Nous avons reconstruit les cartes de dODFs du modèle aQBI avec un ordre maximal d'HS de 8 et le facteur de régularisation fixé à 0.006. La figure 23 montre que le nc-χ LMMSE permet de mieux retrouver le croisement de faisceaux au voxel zoomé.

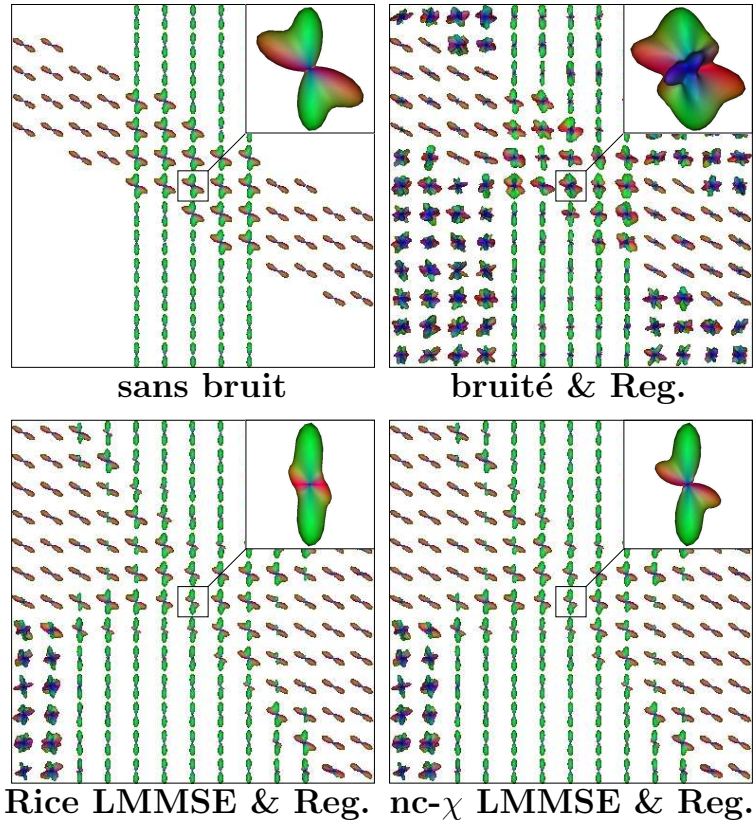


Figure 23: Comparaison entre le "Rice LMMSE" et le "nc-χ LMMSE" sur données simulées pondérées en diffusion et corrompues par un bruit χ-nc stationnaire avec $n = 4$ et $\sigma = 20$. Les cartes dODF ont été obtenues avec régularisation.

Nous avons ensuite comparé les deux LMMSEs sur nos données réelles acquises avec l'algorithme GRAPPA à $b = 1500/3000/4500/6000s \cdot mm^{-2}$. Nous avons déterminé les

paramètres n_{eff} et σ_{eff} de façon globale —comme expliqué précédemment— et nous avons comparé cette technique à la méthode d'estimation de ces paramètres voxel à voxel de [Aja-Fernández et al. (2013)]. Nos résultats (figure 24) donnent des images plus nettes avec la méthode "nc-χ LMMSE" globale, mais sujettes à des hyperintensités à $b = 6000s \cdot mm^{-2}$.

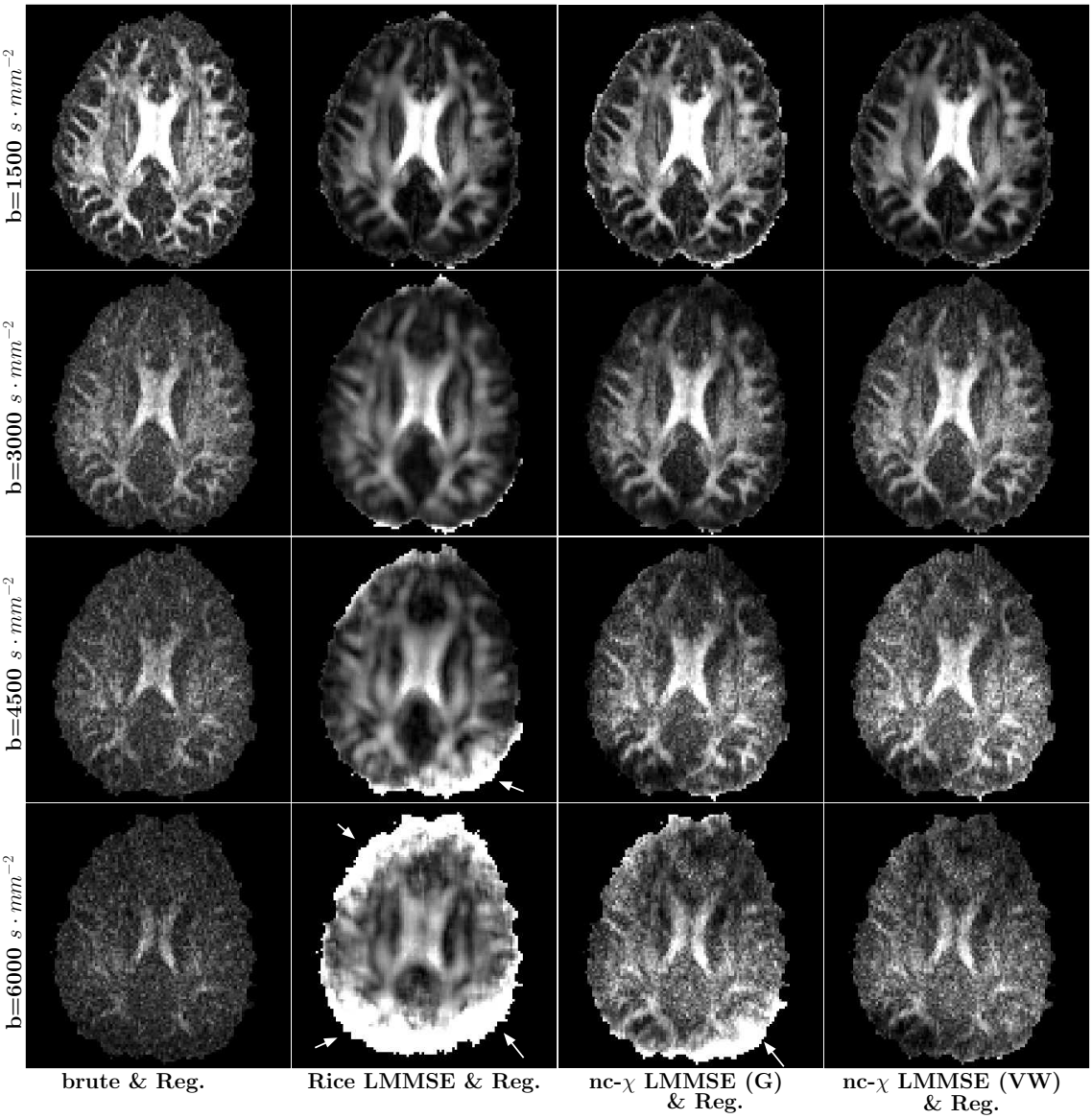


Figure 24: Comparaison entre le "Rice LMMSE" et le "nc-χ LMMSE" sur cartes de GFA réelles avec la méthode d'estimation globale (G) des paramètres effectifs ou la méthode voxel à voxel (VW) à $b = 1500/3000/4500/6000s \cdot mm^{-2}$. Les hyperintensités sont indiquées par des flèches blanches.

Nous avons choisi d'exécuter le LMMSE avec un voisinage de 5×5×5 voxels, ce qui induit un effet de lissage (dû au calcul de moyennes nécessaire au LMMSE) modéré, ainsi qu'un temps d'exécution faible qui est inférieur au T_R pour la méthode globale.

Chapitre 7: Méthodes de correction du bruit dédiées à l'IRMd en temps-réel

Intégration du LMMSE dans un cadre de KF dédié au TR.

Cette nouvelle méthode est schématisée à la figure 25. Elle consiste à appliquer le LMMSE sur la mesure du signal pondéré en diffusion M, puis à injecter le signal corrigé $\hat{S}$ dans le KF qui, pour le modèle aQBI par exemple, en déduit les coefficients $\hat{\mathbf{C}}^{\mathrm{DW}}$ corrigés du bruit. En outre, ces résultats sont utilisés via une boucle de feedback pour calculer une pondération à intégrer dans le calcul des moyennes pour limiter l'effet de lissage du LMMSE. Cette pondération est calculée ainsi pour un voxel $\mathbf{v}'$ voisin du voxel central $\mathbf{v}$:

$$w(\mathbf{v},\mathbf{v}') = \underbrace{\exp\left(\frac{-(\mathbf{v}'-\mathbf{v})^2}{2\alpha^2}\right)}_{w_{SPATIAL}} \underbrace{\exp\left(\frac{-\sum_{j=1}^{N}\left(\hat{\mathbf{C}}^{\mathrm{DW}}_{\mathbf{v}'}(j)-\hat{\mathbf{C}}^{\mathrm{DW}}_{\mathbf{v}}(j)\right)^2}{2\beta^2}\right)}_{w_{STRUCTUREL}}, \tag{26}$$

avec $\hat{\mathbf{C}}^{\mathrm{DW}}_{\mathbf{v}'}$ et $\hat{\mathbf{C}}^{\mathrm{DW}}_{\mathbf{v}}$ les vecteurs des coefficients sur la base des HS du voisin et du voxel central, respectivement. $w_{SPATIAL}$ pondère l'influence des voisins selon leur distance au voxel central à corriger et $w_{STRUCTUREL}$ favorise les voisins structurellement équivalents au voxel central. L'algorithme "LMMSE & KF" ainsi constitué est facilement adaptable aux modèles DTI et sa-aQBI.

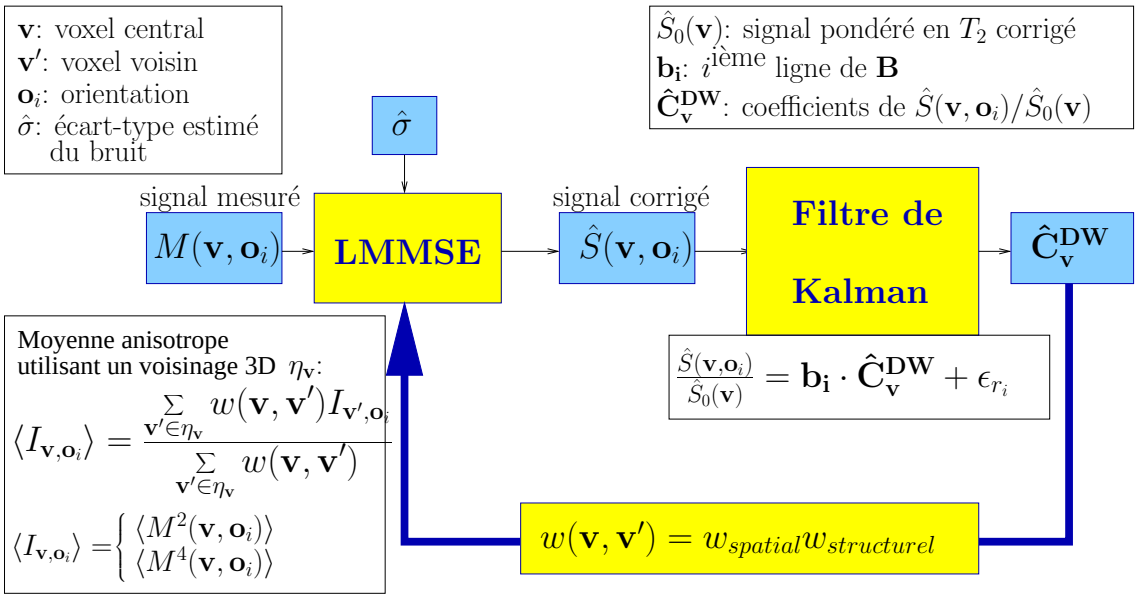


Figure 25: Diagramme de l'algorithme de correction du bruit en TR reposant sur un LMMSE et un KF associés à une boucle de feedback. Ici, la méthode est adaptée pour le modèle aQBI.

La méthode nécessite le réglage de α et de β intervenant dans l'équation 26. $\alpha = 2$ est

un bon compromis entre surlissage et manque d'information voisine. β est choisi de telle sorte que le "LMMSE & KF" appliqué sur des données simulées bruitées par un bruit χ-nc à $n = 4$ et à différents niveaux de bruit produise la plus faible MSE. Ces deux paramètres fixés, on peut comparer cette nouvelle technique au "nc-χ LMMSE" précédent (figure 26) et constater la forte diminution du lissage apportée par le feedback. Les MSE calculés confirment ce gain visuel de qualité.

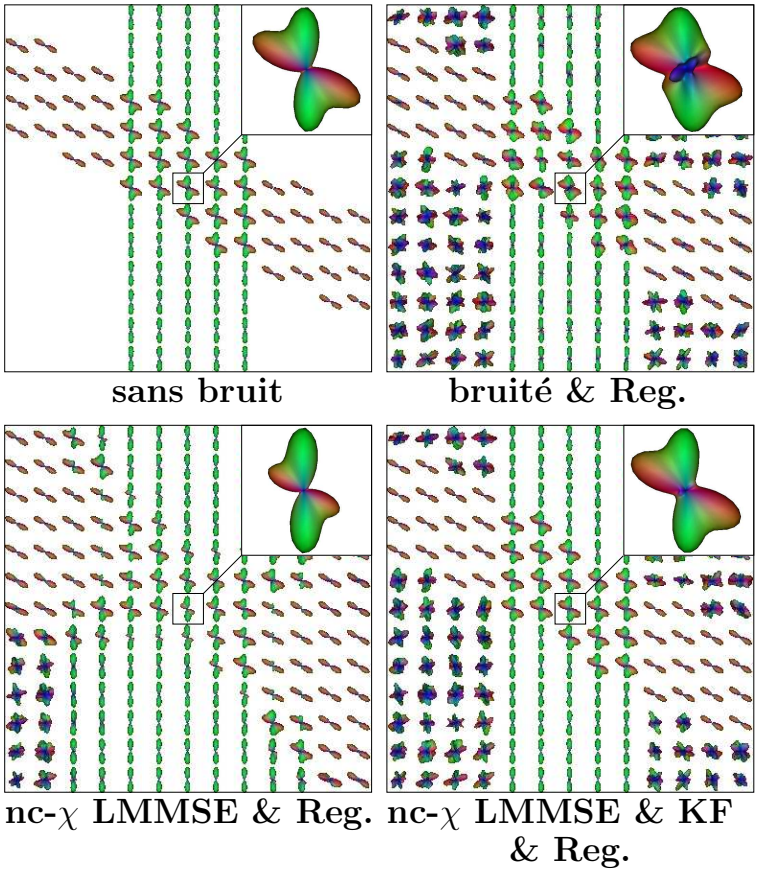


Figure 26: Comparaison entre le "nc-χ LMMSE" et le "LMMSE & KF" sur données simulées pondérées en diffusion et corrompues par un bruit χ-nc avec $n = 4$ et $\sigma = 16$ à $b = 1500s \cdot mm^{-2}$. Les cartes dODF ont été obtenues avec régularisation.

Nous avons effectué une comparaison similaire sur les données réelles, en reportant les valeurs $(\alpha; \beta)$ obtenues pour les données simulées pour le "LMMSE & KF" (figure 27). L'indice de qualité fourni par le rapport de GFA a été mesuré en utilisant les deux mêmes régions dans chaque volume et ces mesures confirment le gain apporté par le "LMMSE & KF".

Nous avons ensuite observé l'impact de la correction à des itérations intermédiaires (figure 28), qui a validé l'intérêt de la méthode en TR. Enfin, la méthode "LMMSE & KF" (version globale) parallélisée et distribuée sur un cluster de 80 CPUs a un temps d'exécution en-dessous du T_R, rendant la méthode totalement compatible pour une utilisation en TR.

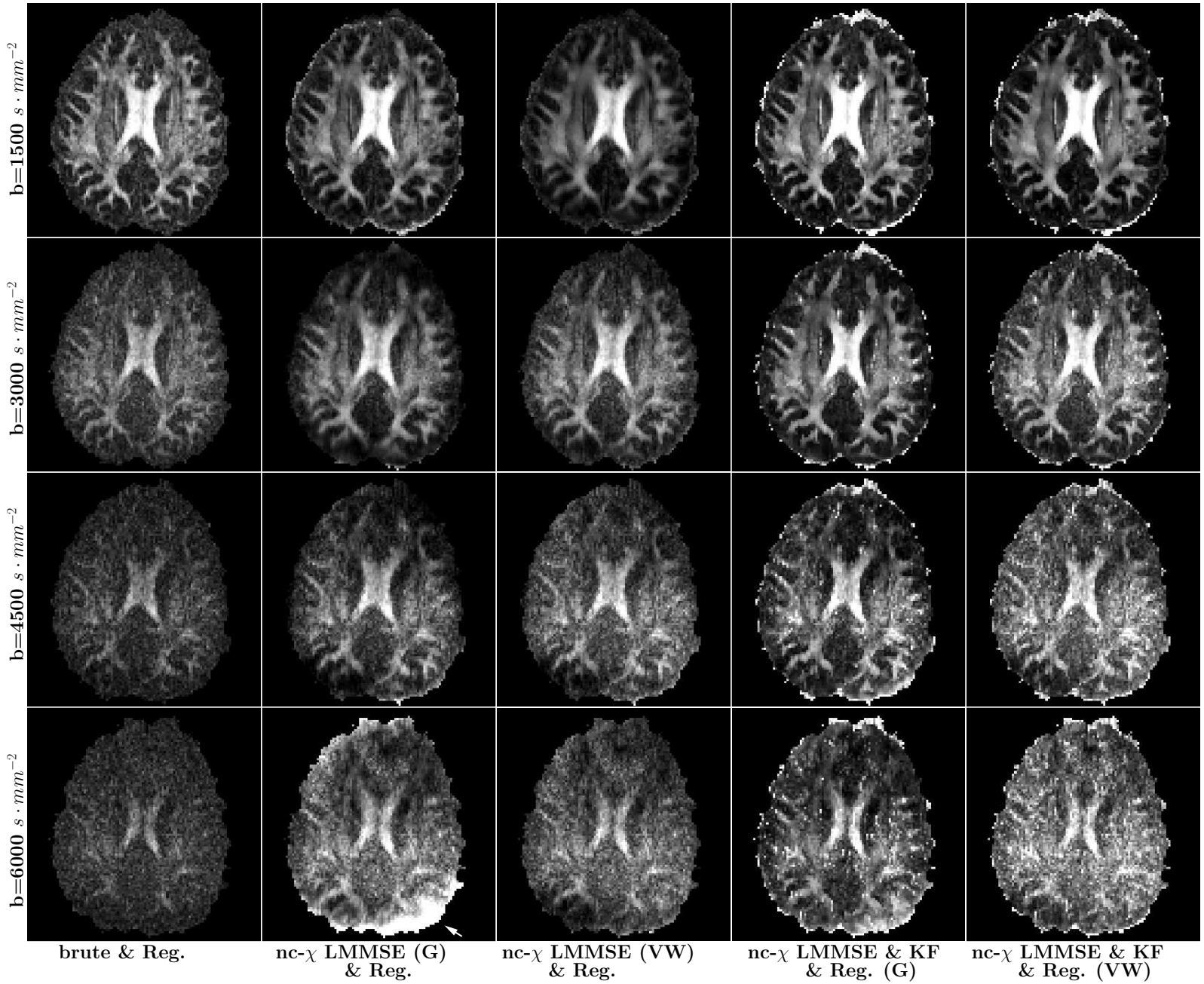


Figure 27: Comparison entre le "nc-χ LMMSE" et le "LMMSE & KF" (versions globale (G) et voxel à voxel (VW)) avec régularisation à $b = 1500/3000/4500/6000 s \cdot mm^{-2}$. Les hyperintensités sont indiquées par des flèches blanches.

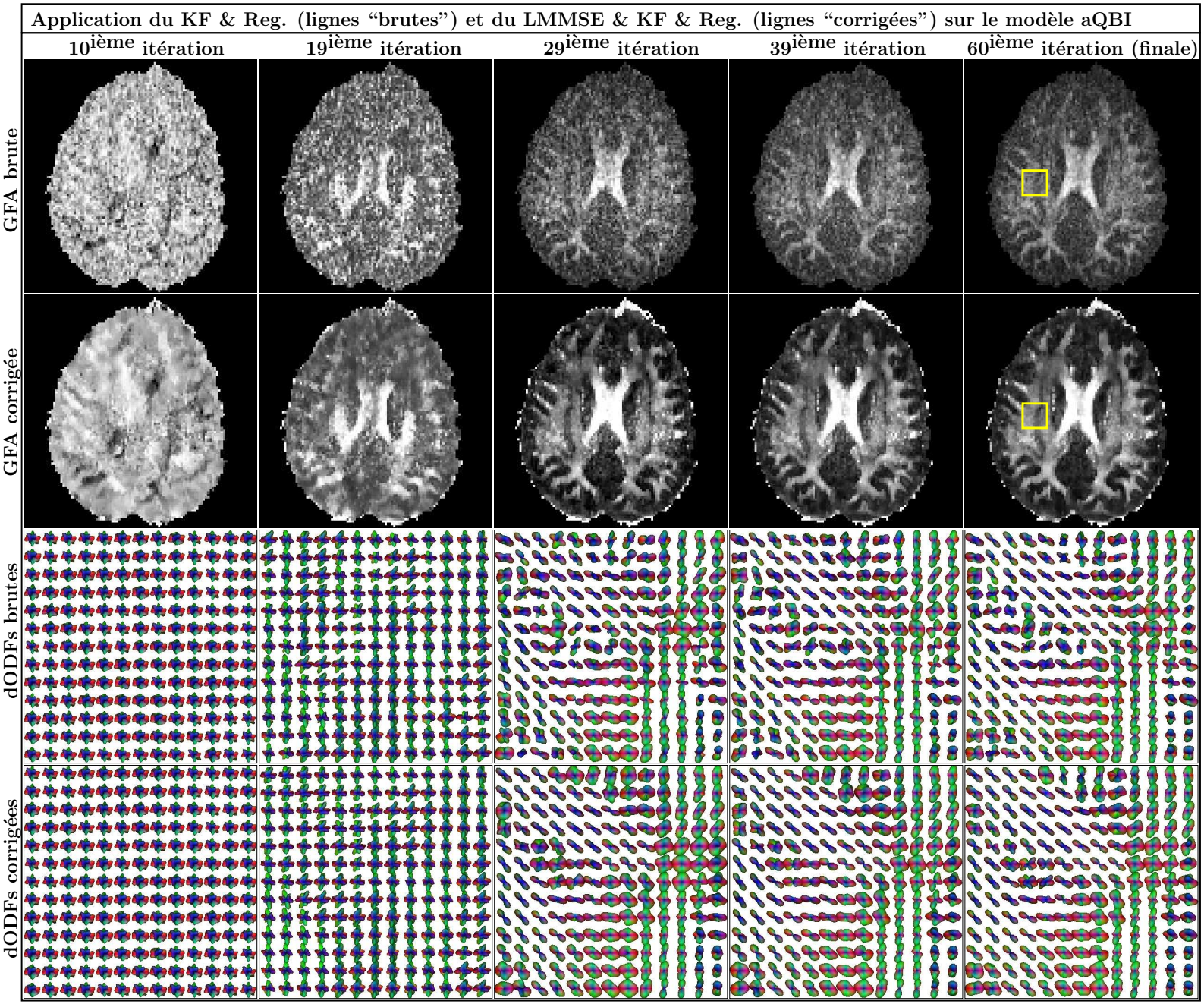


Figure 28: Résultats du "LMMSE & KF" en TR à $b = 3000s \cdot mm^{-2}$ avec le modèle aQBI avec un ordre maximal d'HS de 6 et le facteur de régularisation $\lambda = 0.006$.

Filtrage de bruit gaussien dans l'espace de Fourier.

Nous proposons une seconde méthode de correction du bruit en TR devant être insérée avant la reconstruction SoS, dans la chaîne de reconstruction du fabricant. Cette technique présente l'avantage de corriger les données dans l'espace **k** d'un bruit gaussien non corrélé et à moyenne nulle sur chaque canal de réception. Si le modèle de bruit est plus simple à appréhender par cette technique, la difficulté de cette méthode réside dans l'accès à la chaîne de reconstruction du fabricant. Dans notre cas, nous avons inséré une étape de traitement codée en ICE dans la chaîne du système Siemens 3T Tim Trio. Notre traitement a consisté en un filtre gaussien passe-bas appliqué sur chaque position (k_x, k_y) de l'espace **k**, sur la composante réelle et sur la composante imaginaire du signal. Notons par exemple m_{k_x,k_y} l'une de ses composantes, le signal sans bruit estimé s'écrit alors:

$$\boxed{\hat{s}_{k_x,k_y} = m_{k_x,k_y} \times G(k_x, k_y, \sigma_{Gauss})}, \tag{27}$$

avec $G(k_x, k_y, \sigma_{Gauss}) = e^{-\frac{(k_x-o_x)^2+(k_y-o_y)^2}{2}\sigma^2_{Gauss}}$ une gaussienne en 2D d'écart-type σ_{Gauss} dans l'espace **k** centré en (o_x, o_y). Nous avons testé ce filtre passe-bas sur des données réelles pondérées en diffusion à $b = 1400s \cdot mm^{-2}$ pour différentes valeurs de σ_{Gauss} (figure 29). Cette méthode est très rapide (la correction dure moins de 10ms) et donc valide pour une exécution en TR. Le risque de cette technique est le surlissage lorsque σ_{Gauss} est trop élevé.

Correction reposant sur un filtre de Kalman parallèle (PKF).

Nous avons proposé une troisième méthode de correction du bruit en TR, qui, comme le "LMMSE & KF", s'applique sur l'amplitude du signal et appréhende donc un bruit non gaussien. Cette technique repose sur un filtre de Kalman parallèle (PKF) [Platanio-tis et al. (1997)] qui utilise l'approximation d'une distribution de bruit non gaussienne par une somme de gaussiennes pondérées. De cette approximation, le PKF en déduit une mixture de gaussiennes approximant la PDF de l'observation $p(M_E)$ (M_E étant la mesure du signal pondéré en diffusion normalisé par le signal pondéré en T_2). Puis, le PKF détermine une gaussienne finale approximant $p(M_E)$ pour pouvoir utiliser un KF en respectant l'hypothèse de gaussianité nécessaire au KF. Même si $p(M_E)$ est approximée par une gaussienne, celle-ci est plus proche de la vraie distribution de $p(M_E)$ que ne le serait la gaussienne approximant $p(M_E)$ par un KF classique.

L'ensemble de la méthode reposant sur un PKF est schématisé sur la figure 30. Une première étape consiste à obtenir la somme de gaussiennes représentant la PDF du bruit grâce à une approximation nécessitant d'estimer le paramètre σ de la courbe χ-nc et le signal sans bruit S. Puis, à partir de cette somme de gaussiennes, le PKF détermine les coefficients $\hat{\mathbf{C}}^{\mathrm{DW}}$ corrigés du bruit et utilise une boucle de feedback, comme dans la technique "LMMSE & KF", pour améliorer l'estimation du LMMSE. Les résultats par cette méthode avec une approximation utilisant trois gaussiennes pour représenter la PDF du bruit sont convaincants sur les données simulées. Pour améliorer la robustesse de la

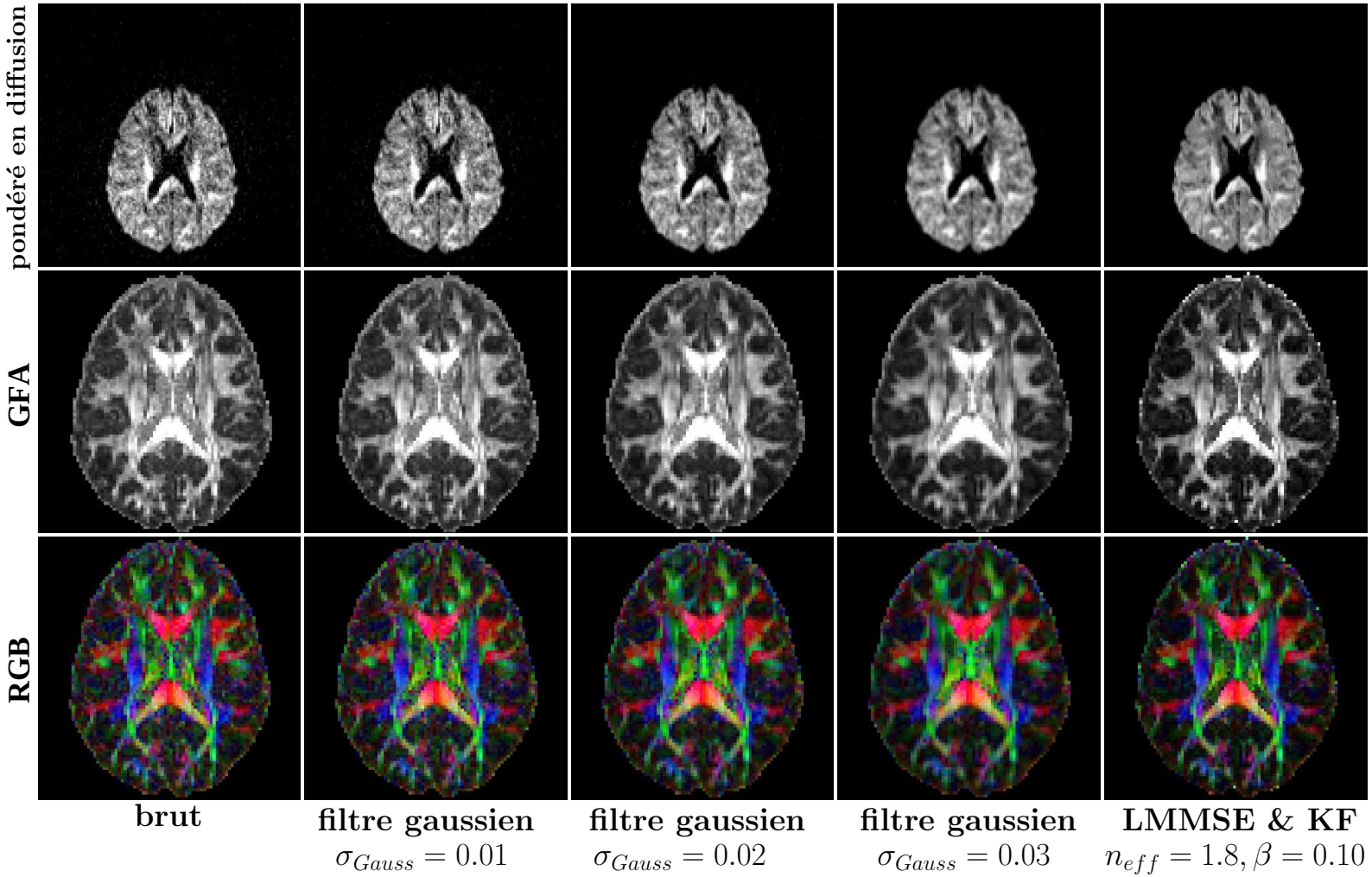


Figure 29: Comparaison entre le filtre gaussien passe-bas, appliqué avec $\sigma_{Gauss} = 0.01/0.02/0.03$ et la méthode LMMSE & KF appliquée avec $n_{eff} = 1.8$, $\alpha = 2$ et $\beta = 0.10$ sur des données pondérées en diffusion acquises à $b = 1400s \cdot mm^{-2}$.

méthode sur les données réelles, plutôt que d'injecter la mesure bruitée M à l'entrée du PKF, nous injectons une moyenne anisotrope de M, ce qui nous donne les résultats des figures 31 et 32.

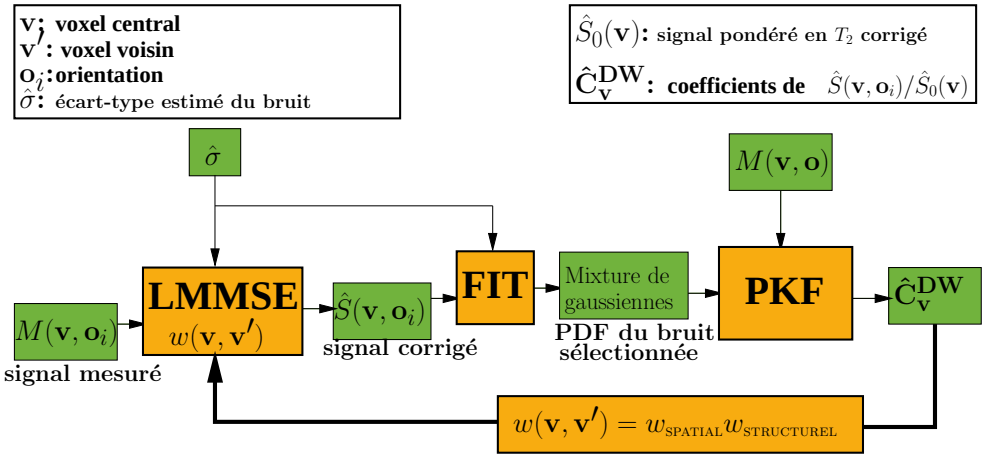


Figure 30: La méthode de correction du bruit en TR reposant sur un PKF.

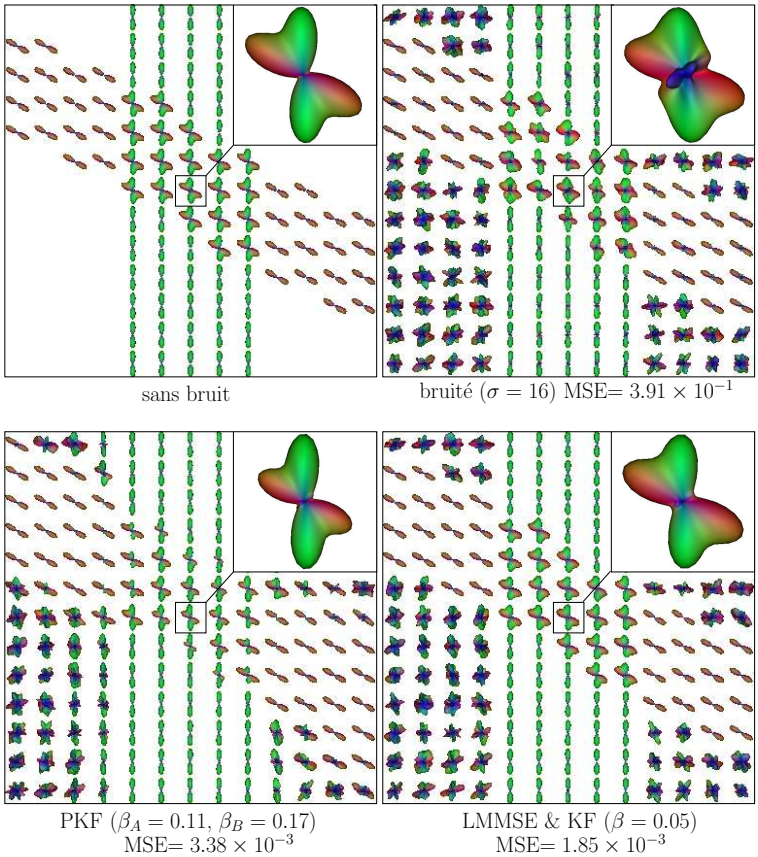


Figure 31: Comparaison entre les méthodes PKF et "LMMSE & KF" sur les données simulées à $b = 4500s \cdot mm^{-2}$, avec un bruit χ-nc de paramètres $\sigma = 16$ et $n = 4$.

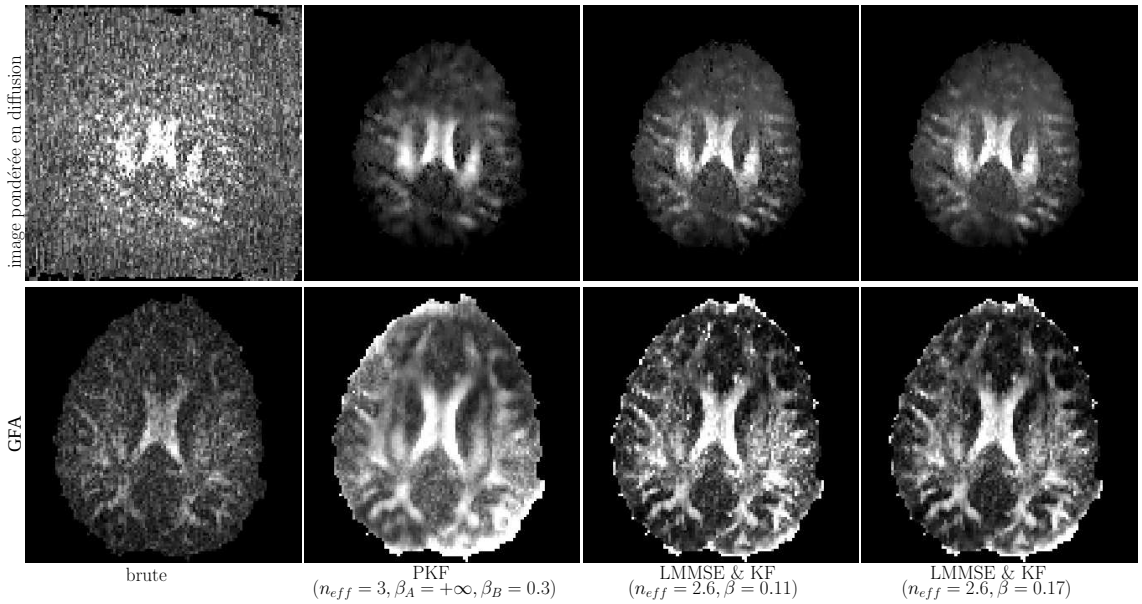


Figure 32: Application de la méthode PKF sur les données réelles à $b = 4500s \cdot mm^{-2}$ et comparaison avec deux configurations de la technique "LMMSE & KF".

Si les résultats avec le PKF sont quasiment équivalents à ceux obtenus avec le "LMMSE & KF" sur les données simulées, ils sont légèrement moins précis sur les données réelles avec un effet de lissage plus important que dans le cas du "LMMSE & KF".

Chapitre 8: Vers la détermination de la connectivité en TR

Dans ce dernier chapitre, nous nous sommes intéressés à la faisabilité de l'exécution d'une technique de tractographie en TR. Comme première étape de notre démarche, nous avons étudié la qualité des résultats intermédiaires de tractographie déterministe de suivi de fibres. De plus, nous avons analysé l'impact de notre correction de bruit en TR "LMMSE & KF" sur ces résultats. La comparaison sur des données simulées est montrée à la figure 33 et celle sur des données réelles est indiquée à la figure 34. Nous pouvons voir que les résultats de tracts en TR sont informatifs dès la $19^{\text{ième}}$ itération. Leur qualité s'affine avec les itérations. D'autre part, les résultats confirment le gain visuel apporté par le "LMMSE & KF".

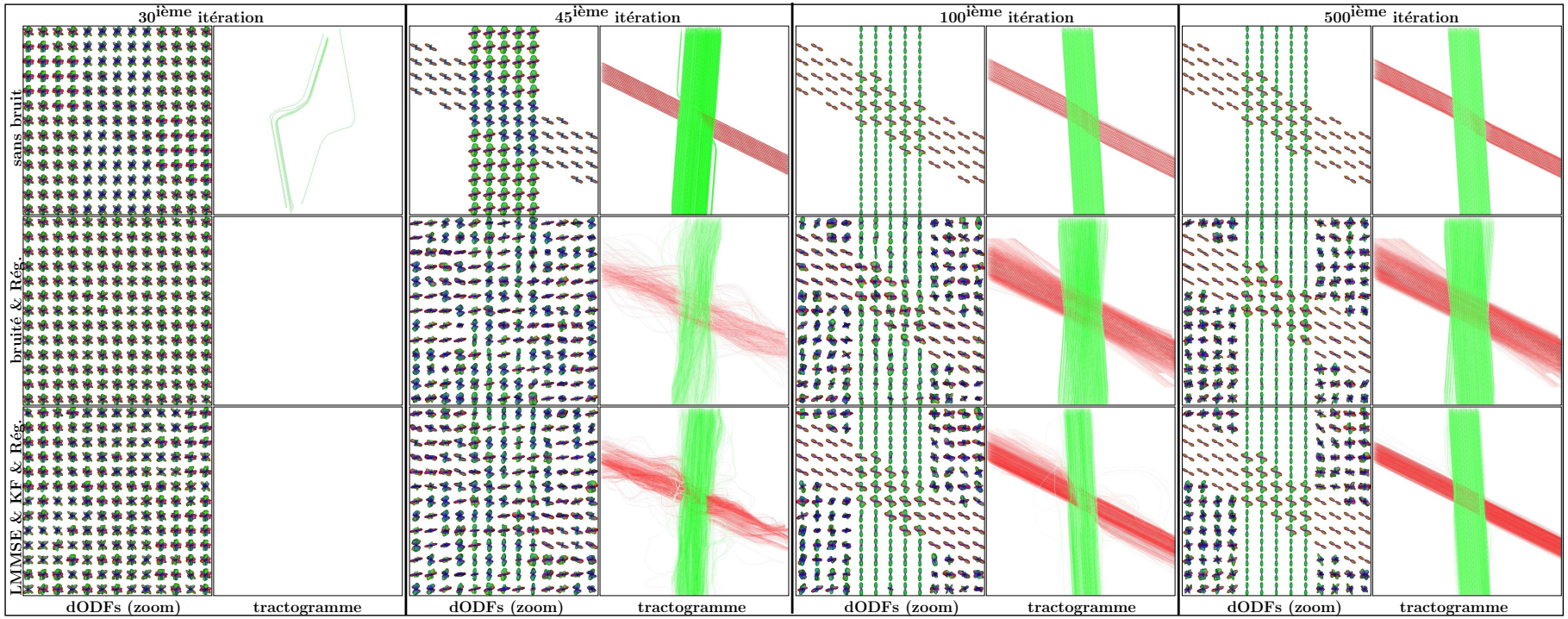


Figure 33: Résultats de dODFs et de tractographie en TR sur les données simulées à $b = 4500s \cdot mm^{-2}$. L'algorithme de tractographie déterministe de suivi de fibres a été exécuté sur les cartes de dODF obtenues par le modèle aQBI avec un ordre maximal d'HS fixé à 4. Quand "Rég." est indiqué, cela signifie que le facteur de régularisation Laplace-Beltrami était fixé à 0.006 (sinon il était fixé à 0).

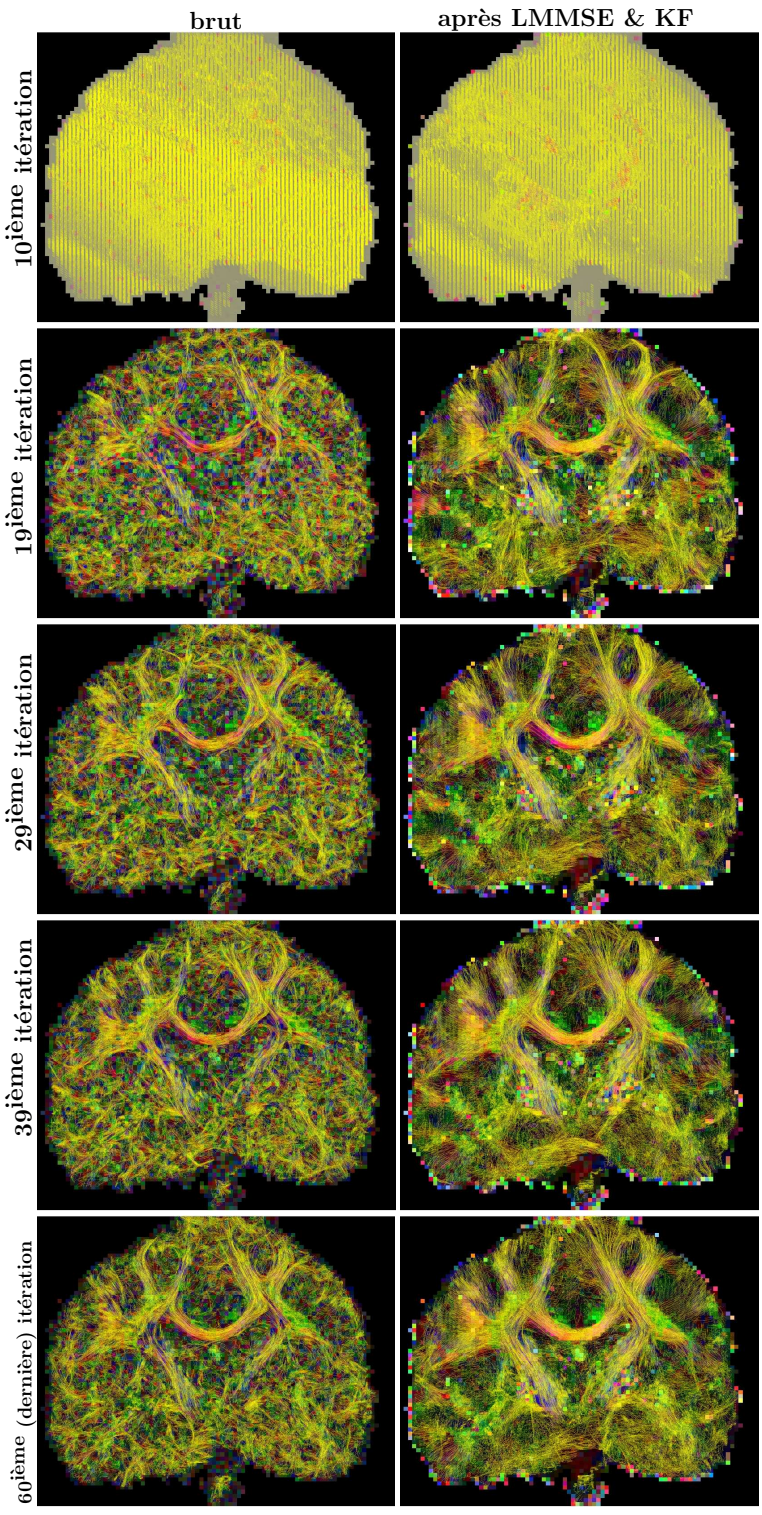


Figure 34: Résultats de tractographie en TR sur les données réelles à $b = 4500s \cdot mm^{-2}$. L'algorithme de tractographie déterministe de suivi de fibres a été exécuté sur les cartes de dODF obtenues par le modèle aQBI avec un ordre maximal d'HS fixé à 4. Le facteur de régularisation Laplace-Beltrami était fixé à 0.006.

Chapitre 9: Conclusion

Dans cette thèse, nous avons proposé des méthodes d'estimation et de correction du bruit pour l'IRMdtr, avec une ouverture finale sur la possibilité de faire de la tractographie en TR. Ce manuscrit a tout d'abord présenté les motivations pour développer l'IRMdtr et le cadre mathématique sur lequel repose le traitement TR en IRMd. Puis, nous avons expliqué la modélisation du bruit en IRM et analysé les méthodes de débruitage de la littérature en regard de nos contraintes TR. Nous nous sommes focalisés sur une solution fondée sur un LMMSE que nous avons étendu à la correction d'un bruit χ non centré (χ-nc). Puis, nous avons développé trois méthodes de correction du bruit en TR, prenant en compte la modélisation théorique du bruit en IRM. La première méthode consiste en la combinaison du LMMSE et d'un filtre de Kalman avec une boucle de feedback. Cette première méthode prend en compte le caractère χ-nc du bruit. Notre deuxième méthode repose sur un débruitage effectué avant la reconstruction SoS, directement sur les données de l'espace **k** qui sont corrompues par un bruit gaussien plus facile à appréhender. Cependant, la technique nécessite de modifier la chaîne de reconstruction du fabricant. Enfin, nous avons proposé une troisième méthode appliquée, comme la première, sur les données d'amplitude du signal. Cette dernière technique repose sur un filtre de Kalman parallèle incorporant la nature χ-nc du bruit. Nous avons comparé nos trois méthodes TR sur des données de diffusion simulées et réelles. Enfin, le dernier chapitre de cette thèse a proposé une étape supplémentaire dans l'inférence de la connectivité du cerveau humain en TR, consistant à proposer un cadre permettant de réaliser une tractographie en direct, pendant l'examen IRM du patient. Ce dernier travail a permis de mettre en évidence le gain de qualité obtenu sur les cartes de tractographie TR après application d'une de nos méthodes de correction du bruit en TR.

Contributions mineures.

Nous avons développé une méthode d'estimation de la variance du bruit dédiée à des données corrompues par un bruit χ-nc et qui ne contiennent pas de fond [Brion et al. (2011b)]. Nos résultats ont montré que notre méthode est équivalente à la méthode de [Rajan et al. (2010)] dédiée à un bruit ricien sans considérer le fond. Cette technique suppose un bruit stationnaire, ce qui est légèrement simpliste dans le cas d'une acquisition à plusieurs canaux. Elle peut néanmoins servir de première approximation pour estimer le bruit sur des données ne présentant pas de fond.

Nous avons proposé un nouvel indice de qualité d'image dédié aux images d'IRMd pour lequel il n'y a besoin ni d'image de référence sans bruit, ni d'une estimation de la variance du bruit [Brion et al.]. L'avantage de cet indice est qu'il peut être utilisé sur des images brutes ou filtrées, en se fondant uniquement sur l'intensité des voxels de deux régions choisies sur une carte d'IRMd, comme une carte de GFA.

Contributions majeures.

Nous avons proposé une extension du LMMSE pour corriger d'un bruit χ-nc [Brion et al. (2011a,b,c)]. Cette méthode, dite LMMSE χ-nc, est plus adaptée au bruit présent dans des images acquises en IRM parallèle par rapport à la méthode LMMSE originale développée pour un bruit ricien.

De plus, nous avons suggéré de prendre en compte d'éventuelles corrélations entre les canaux par une solution empirique consistant à tester le LMMSE avec différentes valeurs pour le nombre de canaux effectif n_{eff} [Brion et al.].

Nous avons conçu une première méthode de correction du bruit en TR dédiée à l'IRMdtr, reposant sur l'association du LMMSE χ-nc et d'un filtre Kalman (KF) permettant de corriger chaque nouveau volume pondéré en diffusion immédiatement après son acquisition et avant celle du volume suivant [Brion et al. (2010, 2011d,e,f); Brion et al.]. La méthode exploite les résultats obtenus en TR en les injectant comme donnée supplémentaire au LMMSE pour améliorer l'estimation des volumes suivants.

Nous avons également développé un filtre passe-bas gaussien appliqué dans l'espace **k** sur les parties réelle et imaginaire de chaque canal. Cette méthode nécessite d'être insérée dans la chaîne de reconstruction du fabricant et est donc spécifique à une constructeur donné.

Enfin, nous avons proposé une troisième méthode de correction du bruit en TR qui repose sur un filtre de Kalman parallèle et qui est appliquée, comme la première méthode sur les données d'amplitude entachées d'un bruit χ-nc [Brion et al. (2012a,b,c)]. Cette dernière technique repose sur l'approximation de la distribution de bruit par une somme de gaussiennes injectées dans un filtre de Kalman parallèle.

Notre dernière investigation concerne la tractographie en TR que nous avons abordée en exécutant un algorithme de suivi de fibres déterministe à chaque nouvelle acquisition d'un volume pondéré en diffusion.

Perspectives.

Nous pensons que les outils développés dans cette thèse contribuent à l'essor de l'IRMdtr. Nos outils de débruitage vont permettre de gagner en qualité d'images obtenues en TR. La méthode LMMSE & KF sera d'ailleurs prochainement intégrée dans le logiciel Connectomist, pour être utilisée par la communauté. Par la suite, nous aimerions élaborer de nouvelles techniques en vue d'applications innovantes de l'IRMdtr. Un brevet [Poupon et al. (2008a)] a été déposé en 2008, axé sur l'optimisation en TR du jeu de gradients de diffusion pour affiner la mesure du signal de diffusion et explorer au mieux la structure anatomique sous-jacente. Ceci est particulièrement valable pour des structures comme la moelle épinière où les faisceaux sont orientés selon une direction principale. L'avantage apporté par le TR est la possibilité de régler le gradient de diffusion à l'itération i en fonction des résultats obtenus aux précédentes itérations. En outre, nous aimerions exploiter l'immense potentiel de l'IRMdtr pour des applications cliniques pédiatriques. Le projet

PEDIART, mené avec le Professeur Chiron, le Docteur Hertz-Pannier et le Docteur Sévin, a d'ailleurs été financé, avec, comme axe d'application, l'épilepsie chez des enfants pour qui la maladie est pharmaco-résistante. Un traitement en TR des données permettrait notamment de mieux contrôler l'impact des mouvements incontrôlés des patients sur les résultats des examens. Cela permettrait d'accélérer l'étude de ces cas pathologiques difficiles à traiter. Enfin, l'IRMdtr offre des possibilités très intéressantes comme l'adaptation d'outils de segmentation automatique à une mise en œuvre TR [Poupon (1999b); Marrakchi-Kacem (2011)] qui permettrait d'accélérer l'étude d'une structure anatomique spécifique du cerveau. Il serait également passionnant d'étendre des outils de classification de groupes de faisceaux de la substance blanche ([Guevara Alvez (2011)]) en vue d'une exploitation en TR. Enfin, une calibration axonale de fibres [Assaf et al. (2008)] se faisant en direct lors de l'examen d'IRMd permettrait d'accélérer l'étude de la distribution de la tailles des axones dans le cerveau, ainsi que d'améliorer des mesures spécifiques de distributions en utilisant les résultats obtenus en TR injectés, par feedback, comme donnée d'entrée au réglage des paramètres de séquence.

Part I

Introduction

Chapter 1

Introduction

Context

Diffusion magnetic resonance imaging (dMRI), introduced in [LeBihan and Breton (1985); Merboldt et al. (1985); Taylor and Bushell (1985)], has become a commonly used imaging modality going from the early diagnosis of ischemia to the inference of the brain connectivity. This modality relies on the measurement of the water diffusion in the brain and allows, from this measurement, to infer the directions of the white matter fiber bundles in the brain. In the last twenty years, several methods were explored to represent at best the map of the pathways connecting neuronal regions together. With the increase of the accuracy of these 3D maps of the routes in the brain, the enthusiasm grew a lot in the community and several other fields were investigated to always image more from the brain microarchitecture.

Recently, a further advance for dMRI, as well as for functional MRI (fMRI) was proposed with the outcome of real-time (RT) MRI workflows. RT MRI is a new concept enabling to perform the usual MRI post-processing online, during the acquisition with the patient in the MRI scanner. Concerning the RT fMRI technique introduced by [Cox et al. (1995)], it was shown to bring new insights in many fields. It was used for example to communicate with completely paralyzed patients, who were thought to be in a vegetative state [Birbaumer et al. (1999)]. RT fMRI was shown to be a new way towards brain computer interfaces (BCI) [Weiskopf et al. (2004); deCharms (2007)], enabling to reduce the chronic pain of patients through a neurofeedback training [deCharms (2008)]. In parallel to these studies opening the possibility to read out functional brain processes in RT, the outcome of RT dMRI (rtdMRI), introduced by [Poupon et al. (2008b)], opened the possibility to read out the structural brain architecture in RT during the exam. In the frame of this thesis, we will focus on the RT workflow dedicated to the dMRI modality.

rtdMRI allows to see in RT the dMRI maps, which are refined each time a new diffusion measurement is performed. Common dMRI maps like the fractional anisotropy, the apparent diffusion coefficient, the diffusion or fiber orientation distribution functions'

maps can be evaluated nearly instantaneously as the dMRI sequence is running. The possibilities brought by such a workflow are powerful. The online results fully permit to take decision about the quality of the scan and putatively allow to correct for any unexpected problem such as motion of the subject, hardware failure... And, it also gives the opportunity to insert any further incremental processing of the acquired MR data to improve the workflow whenever it is possible, thus enabling to possibly stop the acquisition when the actual estimates provide enough information with respect to the diagnostic.

Diffusion imaging is characterized by an exponential decay and is consequently inherently corrupted by noise. While this noise remains Gaussian on the two complex acquisition channels, it unfortunately follows a Rician or a noncentral χ distribution when considering the magnitude of the reconstructed MR data. In addition, the actual tendency with the most novel high angular resolution diffusion imaging (HARDI) and hybrid diffusion imaging (HYDI) models is to increase the diffusion sensitization, thus enforcing the noise level. Furthermore, it can be also interesting to decrease the voxel size to increase the spatial resolution of the dMRI images, making the signal-to-noise ratio decrease. Consequently, an RT denoising method becomes essential to use the rtdMRI workflow at its full extent.

In this thesis, we addressed the feasibility of noise correction in RT which can be challenging when considering the laps of dozen of seconds available between the acquisition of two volumes. First, we focused on one of the state-of-the-art offline denoising techniques which copes with our constraints, the linear minimum mean square error (LMMSE) estimator, originally dedicated to Rician noise, and we extended it to noncentral χ noise correction. Then, we proposed a first RT correction method based on an association of this extended LMMSE estimator with a Kalman filter embedded together with a feedback loop to tackle the oversmoothing inherently introduced by the LMMSE. In addition, we proposed another RT denoising technique relying on a unique tool: a parallel Kalman filter accounting for a non-Gaussian noise. Each technique was applied on the sum of squares (SoS) recombined signals as an RT process. We also developed in this thesis an RT noise correction which directly applies on the signals before the SoS reconstruction. This method has the advantage of accounting for a simple zero-mean Gaussian noise. All this RT environment was tested on a clinical Tim Trio Siemens 3.0T MRI system but can easily be adapted to any clinical MRI system. We also addressed the feasibility of performing tractography in RT to see the tracts being refined after the acquisition of each new measurement. We studied the impact of the RT noise correction for this tractography application.

Organization and contributions of this thesis

This thesis manuscript is divided in three main parts (without including the introduction and conclusion parts). The first part is a large state-of-the-art report on rtdMRI. It contains all the literature background studied for this thesis. The second part focuses on the main axis of this thesis: the RT noise correction dedicated to rtdMRI. This part contains

a review of the scientific literature on the MRI noise. It then details our contributions mostly concerning RT noise issues. Finally, the third part proposes to infer the connectivity in RT with a RT tractography application. Fig. 1.1 shows the organization as a diagram, showing the connections between the different chapters.

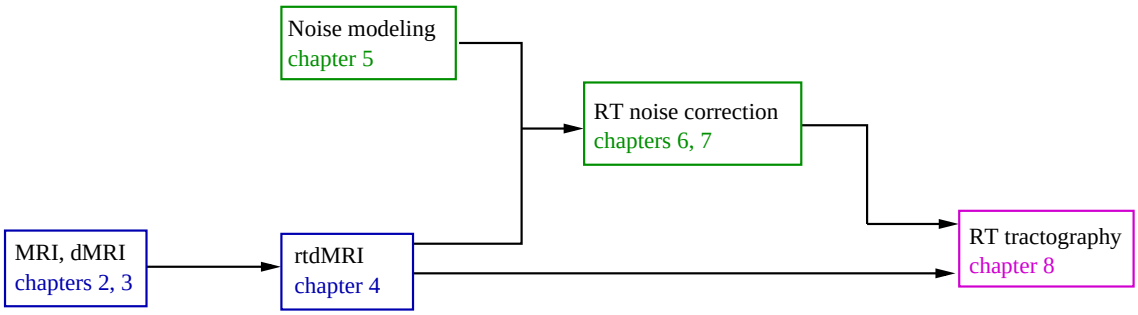


Figure 1.1: Organization of the thesis chapters. The chapters in blue concern the first main part of the thesis with a large state-of-the-art report converging to the rtdMRI method. The chapters in green are included in the second main part of the thesis. They contain the contributions of this thesis dedicated to the RT noise correction. Finally, the chapter in magenta is the RT tractography application. The latter is performed after the RT noise correction.

Part II: State-of-the-art report on real-time diffusion magnetic resonance imaging (rtdMRI)

This part is a large state-of-the-art report converging to rtdMRI. It contains three chapters, which go from basics in MRI to detailed specifications for rtdMRI.

Chapter 2 - Basic principles of nuclear magnetic resonance imaging (MRI): This chapter presents the basic principles of MRI. It goes from the physical principles used in MRI to the image formation process. This chapter is not fundamental to understand this thesis. We wrote it in a spirit to cover at best the MRI modality from which dMRI is derived. However, certain elements like the image reconstruction techniques followed by the parallel MRI techniques will serve as a reference for the next chapters, especially chapter 5 concerning the MRI noise analysis, which is dependent on the MRI reconstruction.

Chapter 3 - Knowledge about diffusion MRI (dMRI) used in this thesis: This chapter moves from MRI to dMRI. It covers all the dMRI knowledge used in this thesis: from the physical phenomenon, on which the dMRI modality relies to essential notions, like the propagator, the orientation distribution function and the $\mathbf{q}$ space, used in different local diffusion models. In this chapter, we focus on the diffusion tensor, the Q-ball and spherical deconvolution models, which will be used for the rest of the thesis. The chapter ends with a presentation of the various classes of tractography techniques stemming from

the dMRI literature. This last overview will be useful for the chapter 8, which proposes to infer the connectivity in RT.

Chapter 4 - State-of-the-art report on real-time dMRI (rtdMRI): Chapter 4 is a central chapter in this thesis. All the next chapters will rely on this one and refer to it. At this point of the thesis, we explain how rtdMRI was thought and elaborated. We first explain the numerous motivations behind rtdMRI and then explain the mathematical Kalman frameworks proposed to achieve the RT goal. These frameworks are adapted to the diffusion tensor and the Q-ball models presented previously in chapter 3. Finally, we address in this chapter the hardware architecture necessary to the RT perspective of making *online* post-processing. This chapter also gives some insights about the noise issue in MRI that will be further detailed in chapter 5.

Part III: Real-time noise correction for real-time dMRI (rtdMRI)

This part about the RT noise correction for rtdMRI presents the main tools developed for this thesis. It goes from the description of the MRI noise to the proposition of solutions to answer the RT dMRI noise correction issue.

Chapter 5 - Modeling of noise in MRI: This chapter contains an important state-of-the-art review about the MRI noise modeling and quantification analysis. It describes the different statistical distributions of the measured noisy MRI signal and their impact on the visual dMRI results. This will serve as a basis for the next chapters 6 and 7 about denoising solutions. Different MRI noise estimation techniques are presented, with a new one emerging from this thesis work. Finally, various image quality indices are detailed, with a new one proposed specifically for dMRI images which will be used to quantify the noise removal obtained with our solutions proposed in te next two chapters.

Chapter 6 - Correcting noise in RT Chapter 6 tackles the RT noise problematic, which is the heart of this thesis work. As this problematic had never been considered before, the considerations here are completely new and specific to this thesis. This chapter makes a review of the various offline MRI noise correction methods and highlights the constraints brougt by an RT denoising objective. The offline methods are therefore analyzed regarding their potential for a RT application. From this review, the linear minimum mean square error (LMMSE) estimator is selected as a good solution regarding our RT constraints.

Chapter 7 - Noise correction methods dedicated to real-time dMRI: This chapter presents three different RT noise corection methods. The first RT denoising solution relies on an LMMSE estimator embedded in a Kalman framework with a feedback loop. We will show that this technique generates an important improvement in the dMRI results obtained with the rtdMRI workflow. A second RT denoising possibility is addressed,

relying on a correction applied directly in the **k**-space, before the sum of squares (SoS) recombination to obtain the MRI magnitude signal. This other strategy, for which the method has to be implemented in the manufacturer reconstruction pipeline, is very simple regarding the noise analysis, as it has the advantage of accounting for a simple zero-mean Gaussian noise. However, we highlight the difficulty to implement it in practice in the manufacturer reconstruction system. Finally, we propose a third RT denoising solution, which is applied on the magnitude signal, as the first technique presented in this chapter. This last technique tackles the RT noise removal from a more rigorous point of view than the first technique and is based on a parallel Kalman filter. We explain in details the technique and the results obtained. To conclude the chapter, all three methods are compared on their different performances on the theoretical and practical aspects, including the image quality improvement, as well as the time required to run these techniques in RT.

Part IV: Real-time tractography application

This part proposes a chapter containg a prototype for a RT tractography application.

Chapter 8 - Inference of the connectivity in RT: Chapter 8 presents the possibility to perform a tractography algorithm in RT. The technique proposed is at its first development stage and represents a prototype to show the feasibility of the concept. The idea here is to run a streamline deterministic tractography algorithm at each iteration of the rtdMRI workflow. Therefore, no algorithmic change in the method is proposed yet, but the main objective of this chapter is to highlight the effect on the first RT denoising method presented in chapter 7 on the tracts.

Part V: Appendices

Appendix A - Acquisition of real human brain data: This appendix details the generation of real human brain data, which were used to test our algorithms. These data were generated at NeuroSpin in the "Archi database" project on the Magnetom Tim Trio 3T MRI system (Siemens Healthcare, Erlangen, Germany).

Appendix B - Some mathematical functions used in this thesis: This appendix gives the mathematical expressions of some mathematical functions used in this thesis.

Appendix C - Simulation of DW data: This appendix details the generation of simulated DW data without and with noise. The noise-free data serve as reference when evaluating the performances of the algorithms.

Software contributions

All the methods developed in this thesis were integrated into the Ptk library directed by Cyril Poupon and Fabrice Poupon. Some of the methods have been integrated to the RT environment developed under the supervision of Fabrice Poupon with Olivier Riff. In the future, they will be integrated into the Connectomist software.

Part II

State-of-the-art report on real-time diffusion magnetic resonance imaging (rtdMRI)

Chapter 2

Basic principles of nuclear magnetic resonance imaging (MRI)

How does nuclear magnetic resonance imaging (MRI) operate and what are its applications? Which physical principles does it rely on? How are the images generated? This chapter covers the background to answer these questions. First, we present the physical phenomena which are at the origin of this imaging modality. Then, we describe the elements of the well-known spin-echo sequence created by [Hahn (1950)] and which generates a signal revealing the different brain tissues. Finally, we explain the image reconstruction, as well as the parallel MRI techniques that enable to accelerate the acquisition. This recap is inspired in majority by the book of [Kastler et al. (2006)], from which we took a lot of figures, but also by my courses at Phelma given by Dr. Françoise Hippert and Dr. Emmanuel Barbier, by thesis chapters of [Poupon (1999a)], [Marrakchi-Kacem (2011)], and finally by the websites http://users.fmrib.ox.ac.uk, http://www.valley-neuroscience.com and http://www.imaios.com/fr/e-Cours/e-MRI.

2.1 Introduction

MRI was introduced in the 1970's by Paul Lauterbur [Lauterbur (1973)] and Peter Mansfield [Mansfield (1977)], who received, for their discovery, the Nobel Prize in physiology and medicine in 2003. Using the nuclear magnetic resonance (NMR) principles, Paul Lauterbur generated the first two dimensional (2D) MRI images and Peter Mansfield developed, among other works, a method to acquire an image rapidly (in a few seconds), known under the name of echo planar imaging, which is a key scheme in functional and diffusion MRI. MRI allows to make images of the whole body, in a non invasive way, contrary to X-rays for example. MRI contributed to significantly improve the comprehension of the brain and its pathologies. Today, MRI is widely used in clinical medicine, for the diagnosis of pathologies (for example: tumors, neurodegenerative diseases, epilepsies, stroke), and also

in cognitive neuroscience research.

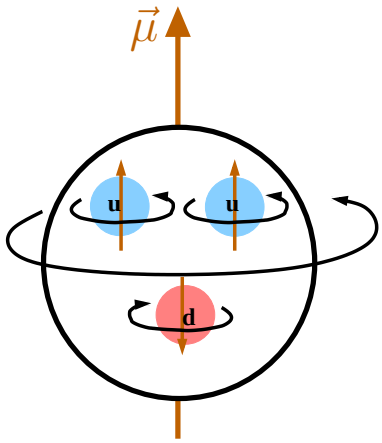


Figure 2.1: Representation of a proton (black circle) and its magnetic moment $\vec{\mu}$. The two letters u designate the two quarks *up* and the letter d designates the quark *down* which together constitute the proton.

This imaging technique is based on the response of the hydrogen protons to imposed magnetic disturbances. Nuclei other than those of hydrogen can sometimes be considered, but in this thesis, we will only consider hydrogen protons. Those are constituted by three subparticules: two quarks *up* and one quark *down* represented by the symbols u and d, respectively (fig. 2.1). The *up* and *down* quarks rotate about the same axis but in opposite directions, causing the proton rotation. A rotating particle has an intrinsic kinetic moment aligned along the particle rotation axis and called "spin", and the rotating proton is called **spin** by extension. A quark is a both charged and rotating particle and consequently induces a magnetic moment, aligned along its rotation axis. The addition of the three little magnetic moments created by the three quarks causes a magnetic moment for the proton. The latter can then be considered as a small magnet. For a set of magnetic moments, we define the vector of **magnetization A** such that: $\mathbf{A} = \sum_{i=1}^{N} \boldsymbol{\mu_i}/V$, where N is the number of spins in the observed region and V is the observed volume. In absence of external magnetic field, the direction of the spins are random, like in the fig. 2.2. The net magnetization is then nul. What does happen when an external magnetic field or a strong magnetic disturbance is applied?

First let's describe the hardware configuration of an MRI scanner (fig. 2.3). Such a machine is composed of several hardware elements that we will describe. First, a **superconductive coil** generates an homogeneous magnetic field $\mathbf{B_0}$ within a given field of view around its isocenter. This coil has the advantage of creating a high field, without consuming any electric current thanks to its superconductivity. It defines the magnetic equilibrium status of the subject, who is located inside the MRI scanner. The MRI machine also contains an **antenna**, which consists of a coil and permits to disturb the magnetic equilibrium status. It is a key element of the **magnetic resonance** that we detail in sec-

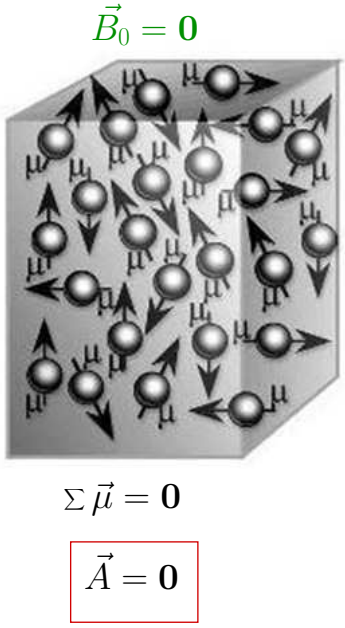


Figure 2.2: Representation of the spins in absence of magnetic field. The magnetization $\vec{A}$ is then zero. Extracted from [Kastler et al. (2006)].

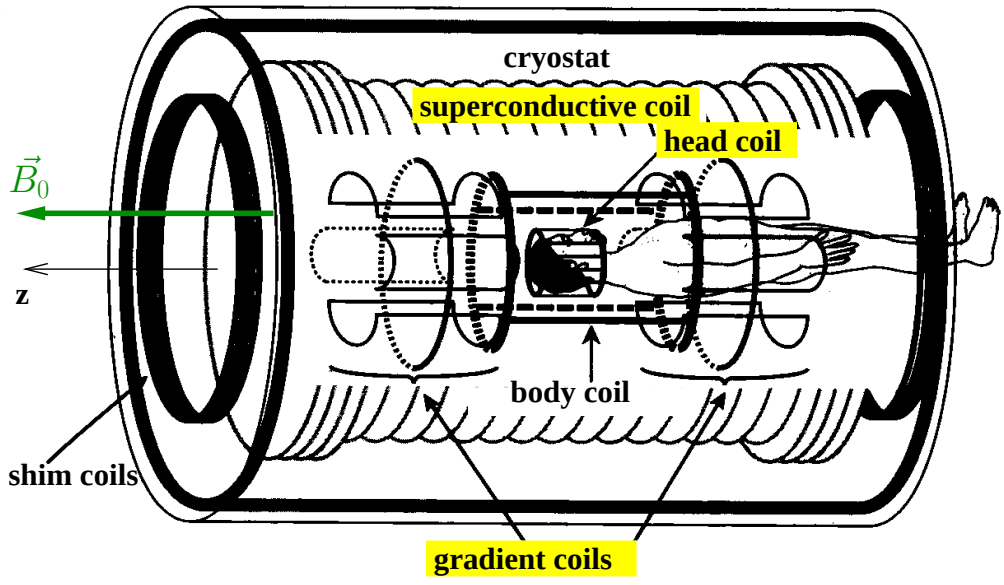


Figure 2.3: Structure of an MRI scanner. Adapted fig. from [Kastler et al. (2006)].

tion 2.2. This coil can also sometimes play the role of the MRI signal reception antenna, *e.g.* receive the protons' response to the magnetic disturbance. In general, the reception antenna is adapted to the geometry of the object of interest, the head in our case. And

finally, the **gradient coils** permit to spatially encode the MRI signal and thus to create an image. Other elements exist in the MRI scanner like the cryostat, which contains liquid helium at a temperature of 4° Kelvin (*i.e.* -269°C) and which consequently enables to keep the coil in its superconductive state. The shim coils help to maintain a static field $\mathbf{B_0}$ homogeneous in a region centered around the isocenter of the magnet. Finally, the room, where the MRI machine is located, is protected from any external radio-frequency (RF) waves by a Faraday cage.

2.2 Magnetic resonance and relaxation phenomena

When the machine is on, the superconductive coil creates a static magnetic field $\mathbf{B_0}$ directed along the z-axis indicated in fig. 2.3. This field starts the **precession** of the spins around the z-axis at an angular frequency called the Larmor frequency: $\omega_0 = \gamma B_0$ with γ the proton gyromagnetic ratio ($\gamma/(2\pi) \sim 43 MHz.T^{-1}$). Actually, the spins precess at $\gamma\,(B_0 + b)$ with b the time fluctuating field created partly by electrons around the hydrogen nuclei. As $b << B_0$, the spins' precession Larmor frequency can be approximated to ω_0. The spins precess in a direction such that their magnetic moment is pointed in the same direction as $\mathbf{B_0}$ or in the opposite direction. As the spins are dephased, those that precess with a precession axis in the same direction as $\mathbf{B_0}$ generate a global magnetic moment along $\mathbf{B_0}$ (fig. 2.4). The other generate a global magnetic moment in the opposite direction. Thanks to a small majority of spins oriented along $\mathbf{B_0}$ and because living tissues contain water in abundance, the resulting magnetic moment is not zero. Consequently, a magnetization vector $\mathbf{A}$ of the observed tissue is induced. It is static and oriented along $\mathbf{B_0}$. The system of spins is then in a magnetic equilibrium state. How can we disturb this state and what does happen then?

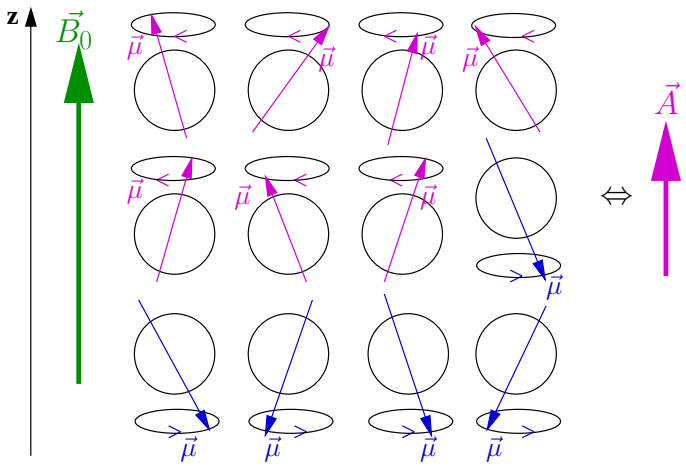


Figure 2.4: Spin precession with magnetic moments designated as $\vec{\mu}$ in an environment occupied by a magnetic field $\mathbf{B_0}$. This precession produces a magnetization vector $\vec{A}$.

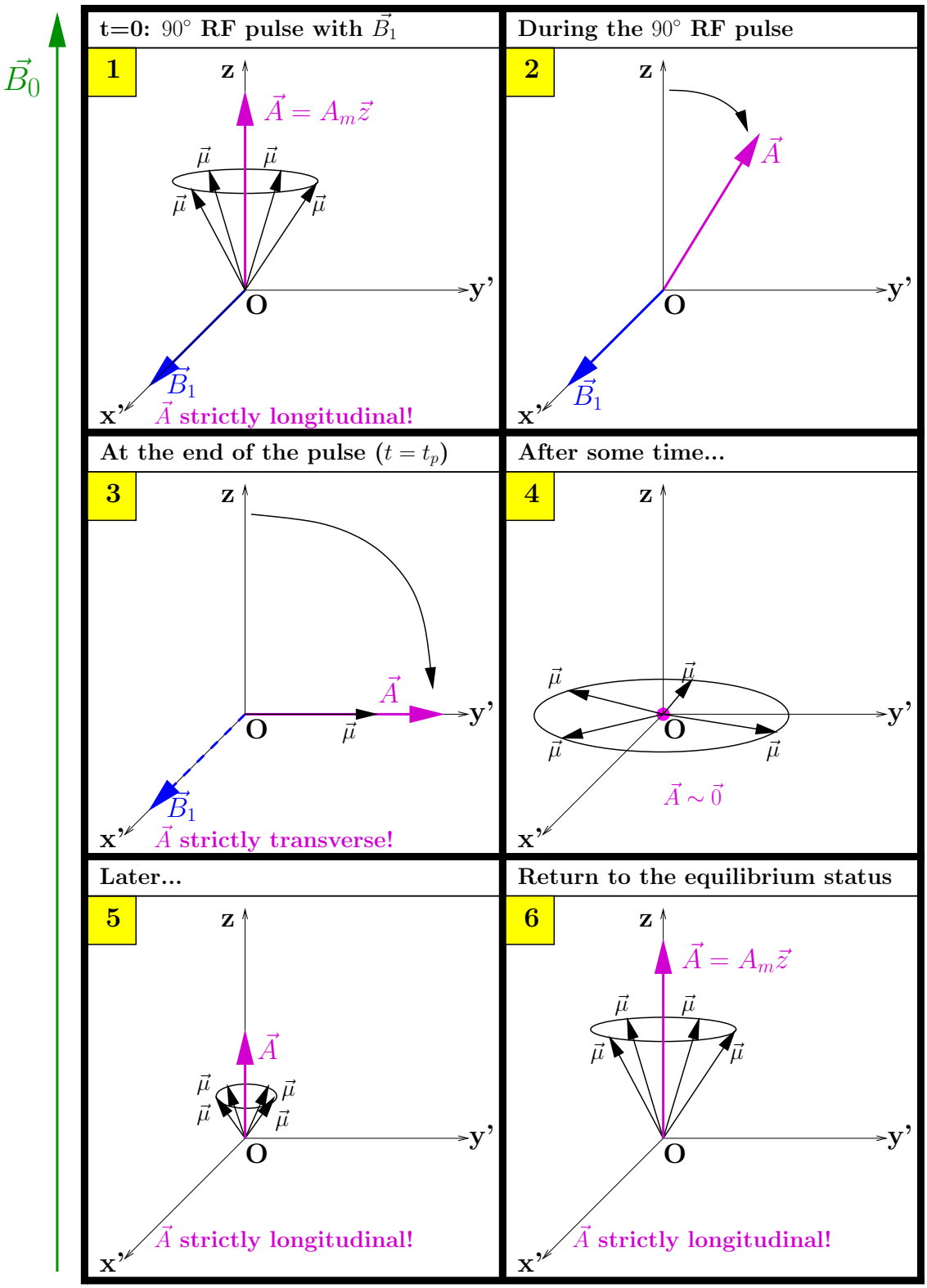


Figure 2.5: Effect of a 90° RF pulse on the magnetization vector $\vec{A}$, introduced in fig. 2.4. This 90° RF pulse creates the rotating magnetic field $\mathbf{B_1}$. The magnetic resonance phenomenon (vignettes 1, 2 and 3) and the relaxation phenomenon (vignettes 4,5 and 6) occur.

The equilibrium state of the magnetization $\mathbf{A}$ of the observed tissue can be disturbed by an RF antenna (the head coil visible in the fig. 2.3 for example), which creates a rotating magnetic field $\mathbf{B_1}$ which orientation is perpendicular to the z-axis and which angular velocity is $\omega_{exc} = \gamma B_1$. The components of the vector $\mathbf{B_1}$ are:

$$\begin{cases} B_{1x}(t) = B_1 \cos(\omega_{exc} t), \\ B_{1y}(t) = B_1 \sin(\omega_{exc} t), \\ B_{1z}(t) = 0. \end{cases} \tag{2.1}$$

The energy transfer required to excite the spin system only occurs when the angular velocity ω_{exc} satisfies the **resonance condition**: $\omega_{exc} = \omega_{res}$. In practice, resonance happens when ω_{exc} is sufficiently close to ω_{res} so that $|\omega_{exc} - \omega_{res}| << \gamma B_1$. Here, the resonance angular velocity is ω_0. When the resonance condition is satisfied, the spins will precess around $\mathbf{B_1}$, at the angular velocity given by $\omega_{exc} = \omega_0$, in addition to their precession around $\mathbf{B_0}$. They will also all get the same phase. It is the **magnetic resonance** phenomenon brought to light by Felix Bloch [Bloch (1946)] and Edward Purcell [Purcell et al. (1946)] in 1946 who received the Nobel Prize for physics in 1952 for their work. Fig. 2.5 explains the steps, in a vignette form, of the spins' excitation by an RF wave, called 90° RF, as it permits to flip the magnetization by 90°. To simplify the phenomenon description, we consider the rotating frame of reference (Ox'y'z), where the vector $\mathbf{B_1}$ is fixed and parallel to the x'-axis. At $t = 0$ (vignette 1), the magnetization $\mathbf{A}$ of the spin system is still longitudinal, *i.e.* aligned along $\mathbf{B_0}$. Little by little, because of the 90° RF pulse, the magnetization tips into the transverse plane and the spins begin to get all the same phase (vignette 2). At time t_p (vignette 3), *i.e.* immediately at the end of the RF pulse, all the spins have the same phase. The magnetization is then strictly transverse. After some time, there is the **relaxation of the transverse component** of $\mathbf{A}$ which is caused by the dephasing of the spins in the transverse plane (vignette 4). Later, there is the **relaxation of the longitudinal component** of $\mathbf{A}$ which begins to recover, because of a return to the spin equilibrium state (vignettes 5 and 6). The magnetization ends up to be strictly longitudinal, as it was before the RF pulse.

The magnetization time evolution can be mathematically described by the **Bloch equations** in the laboratory (stationary) frame of reference (Oxyz) [Bloch (1946)]:

$$\boxed{\frac{d\mathbf{A}}{dt} = \gamma \left(\mathbf{A} \times \mathbf{B}\right) - \frac{A_x \mathbf{x} + A_y \mathbf{y}}{T_2} - \frac{A_z - A_m}{T_1}\mathbf{z}}, \tag{2.2}$$

where $\times$ is the vector product operator. The solution of this equation is, in the rotating frame of reference (Ox'y'z), such that [Haacke et al. (1999)]:

$$\begin{cases} A_{x'}(t) = \exp[-t/T_2] \left(A_{x'}(0) \cos\left[(\omega_0 - \omega_{exc})t\right] + A_{y'}(0) \sin\left[(\omega_0 - \omega_{exc})t\right]\right), \\ A_{y'}(t) = \exp[-t/T_2] \left(A_{y'}(0) \cos\left[(\omega_0 - \omega_{exc})t\right] - A_{x'}(0) \sin\left[(\omega_0 - \omega_{exc})t\right]\right), \\ A_z(t) = A_z(0) \exp[-t/T_1] + A_m \left(1 - \exp[-t/T_1]\right), \end{cases} \tag{2.3}$$

with T_1 and T_2 the longitudinal and transverse relaxation times respectively ($T_1 > T_2$) and $A_m = A_z(\infty)$. The transverse relaxation is also called "spin-spin" relaxation because it is due to interactions between spins. The longitudinal relaxation can be called "spin-lattice" relaxation because it is due to a return to the spins' equilibrium inside of their lattice. The **transverse magnetization** is, in the frames of reference (Ox'y'z) and (Oxyz), defined respectively by:

$$\begin{cases} A_{x'y'}(t) = A_{x'}(t) + iA_{y'}(t), \\ A_{xy}(t) = A_x(t) + iA_y(t), \end{cases} \tag{2.4}$$

with the link between both expressions given by $A_{xy}(t) = A_{x'y'}(t)e^{-i\omega_{exc}t}$. In the particular case of a 90° RF pulse, we have the following relations at $t = t_p$, *i.e.* immediately at the end of the RF pulse (see vignette 3 in fig. 2.5):

$$\begin{cases} A_{x'}(t_p) = 0, \\ A_{y'}(t_p) = A_m, \\ A_z(t_p) = 0. \end{cases}$$

With the time origin changed and now fixed immediately after the 90° pulse, the transverse magnetization is found to be: $A_{xy}(t) = iA_m e^{-t/T_2} e^{i(-\omega_0 t + \phi(0))}$, with the constant phase $\phi(0) = -\omega_{exc} t_p$. We write the **magnitude** of the transverse magnetization as the modulus of $A_{xy}(t)$: $\boxed{A_\perp(t) = |A_{xy}(t)|}$. The curves representing the x-component $A_x(t)$ of the transverse magnetization and the longitudinal magnetization $A_z(t)$ are shown in fig. 2.6.

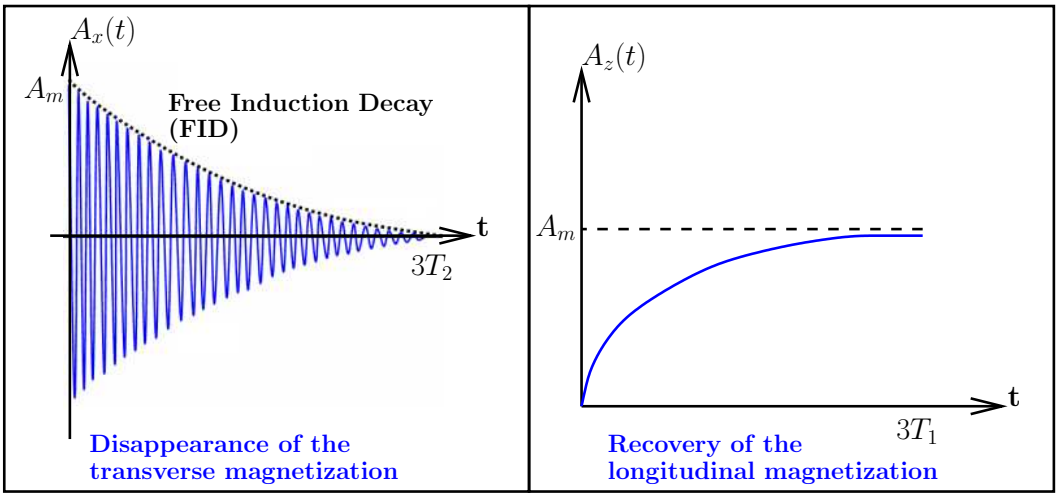


Figure 2.6: The transverse relaxation curve, also designated as the Free Induction Decay (FID), shows the disappearance of the magnetization $A_x(t)$ (left). A similar curve shape is obtained for $A_y(t)$ (not shown). The longitudinal relaxation curve shows the recovery of the longitudinal magnetization $A_z(t)$ (right).

In MRI, the signal measured by the reception antenna only corresponds to the transverse magnetization. It is maximum right after the 90° pulse, and then periodically evolves

in an exponentially decaying envelope. This is called the **free induction decay (FID)**. This signal decrease is rapid and this is partially due to **inhomogeneities in the $\mathbf{B_0}$ magnetic field**, which induce spontaneous precessions of the spins with angular velocities slightly different from ω_L. Then, the **effective transverse relaxation time** T_2^* can be defined such that: $1/T_2^* = 1/T_2 + 1/T_{2i}$, with T_{2i} the relaxation time due to the inhomogeneities of $\mathbf{B_0}$ and T_2 the "spin-spin" relaxation time due to the molecular interactions. Thus, the T_2 parameter in eq. 2.2 and in the system 2.3 has to be replaced by T_2^*. Let us rewrite the expression of the transverse magnetization in the stationary frame of reference (Oxyz) with the time origin fixed immediately after the 90° RF pulse:

$$A_{xy}(t) = iA_m e^{-t/T_2^*} e^{i(-\omega_0 t + \phi(0))}. \tag{2.5}$$

Knowing the T_1, T_2 and T_2^* parameters, as well as the transverse magnetization time evolution after a 90° pulse, we can now detail the basis of the spin-echo sequence.

2.3 Basis of the spin-echo sequence

Fig. 2.7 shows the principle of the **spin-echo sequence**, the concept of which was discovered by Erwin Hahn [Hahn (1950)] in 1950. The sequence immediately begins after the 90° RF pulse. At $t = 0$, the spins have all the same phase and the transverse magnetization is maximum, as illustrated by the vignette 1 in the fig. 2.8. Little by little, the spins are dephased and the relaxation T_2^* happens (vignette 2). At $t = T_E/2$, a 180° RF pulse is applied and enables a flip of the magnetization by an angle of 180° (vignette 3 in fig. 2.8). The flip caused by the 180° RF pulse is a mirror flip with respect to the (y'z) plane. It changes the dephasing sign, without changing the dephasing evolution direction. As the dephasing of the spins is caused by the inhomogeneities of $\mathbf{B_0}$, it only depends on the environment. Consequently, the 180° RF pulse generates a refocusing of the spins: because of the 180° flip, the spin phases begin to decay to zero and are equal to zero at the echo time T_E (vignette 5). At T_E, the spins are again in coherence and the magnetization, which is again maximum, generates an echo signal: this phenomenon is called the **spin echo**. The magnetization at $t = T_E$ is however smaller than at $t = 0$, because of the "spin-spin" relaxation due to molecular interactions. It is possible to create several successive echoes when repeating this cycle every T_R, T_R being the repetition time indicated in fig. 2.7. At the first echo, *i.e.* at $t = T_E$, the modulus of transverse magnetization is:

$$\boxed{A_\perp(T_E) = A_m \exp[-T_E/T_2]}. \tag{2.6}$$

At the next echoes, *i.e.* for $t > T_R$, if the recovery of the longitudinal magnetization is not complete —because of a too short T_R—, this phenomenon has also to be considered. Thus, the modulus of the transverse magnetization at the n^{th} echo with $n > 1$ is expressed such that:

$$\boxed{A_\perp(nT_E) = A_m \left(1 - \exp[-T_R/T_1]\right) \exp[-nT_E/T_2]}. \tag{2.7}$$

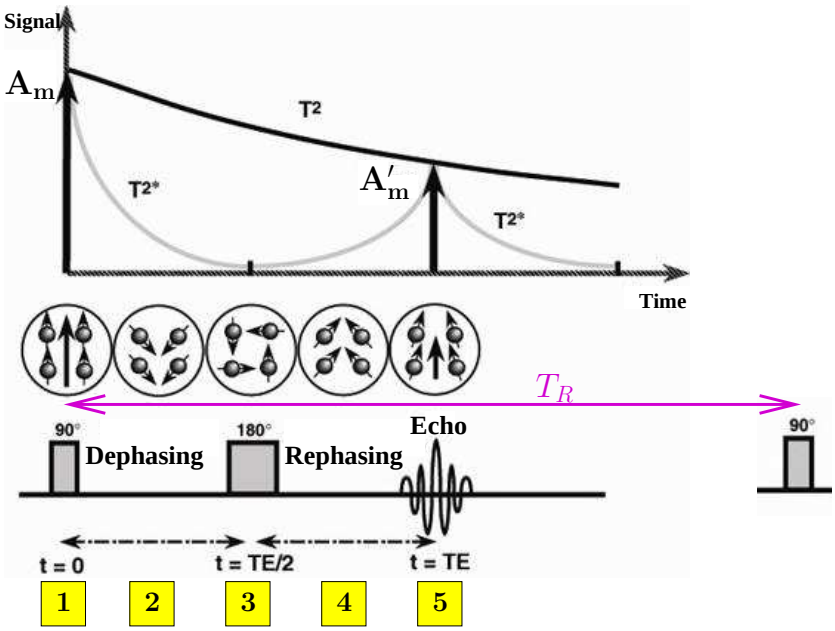


Figure 2.7: Basis of the spin-echo sequence. The bottom of the figure indicates the steps of the sequence, which is constituted by 90° and 180° RF pulses. The time origin is fixed immediately after the 90° RF pulse. The echo, *i.e.* the maximum transverse magnetization signal after a 180° pulse, is obtained at the echo time (T_E). These steps are then repeated at $t = T_R$, T_R being the repetition time after the 90° pulse. The top of the figure indicates the time evolution of the transverse magnetization enabling to reveal the relaxation time T_2 and T_2^*. The circles under the graph show the spin configuration during the sequence. At $t = T_E$, the spins are refocused, producing the echo signal of the transverse magnetization. Figure adapted from [Kastler et al. (2006)]. The numbered instants in a yellow frame are detailed in the fig. 2.8.

This cycle repeated every T_R enables to fill the lines of the Fourier plane from which we then get access to the image itself. This part of the MRI reconstruction process is described in section 2.6. But, before creating the image, how can the MRI signal reveal something interesting of the subjects' brain inside the MRI machine?

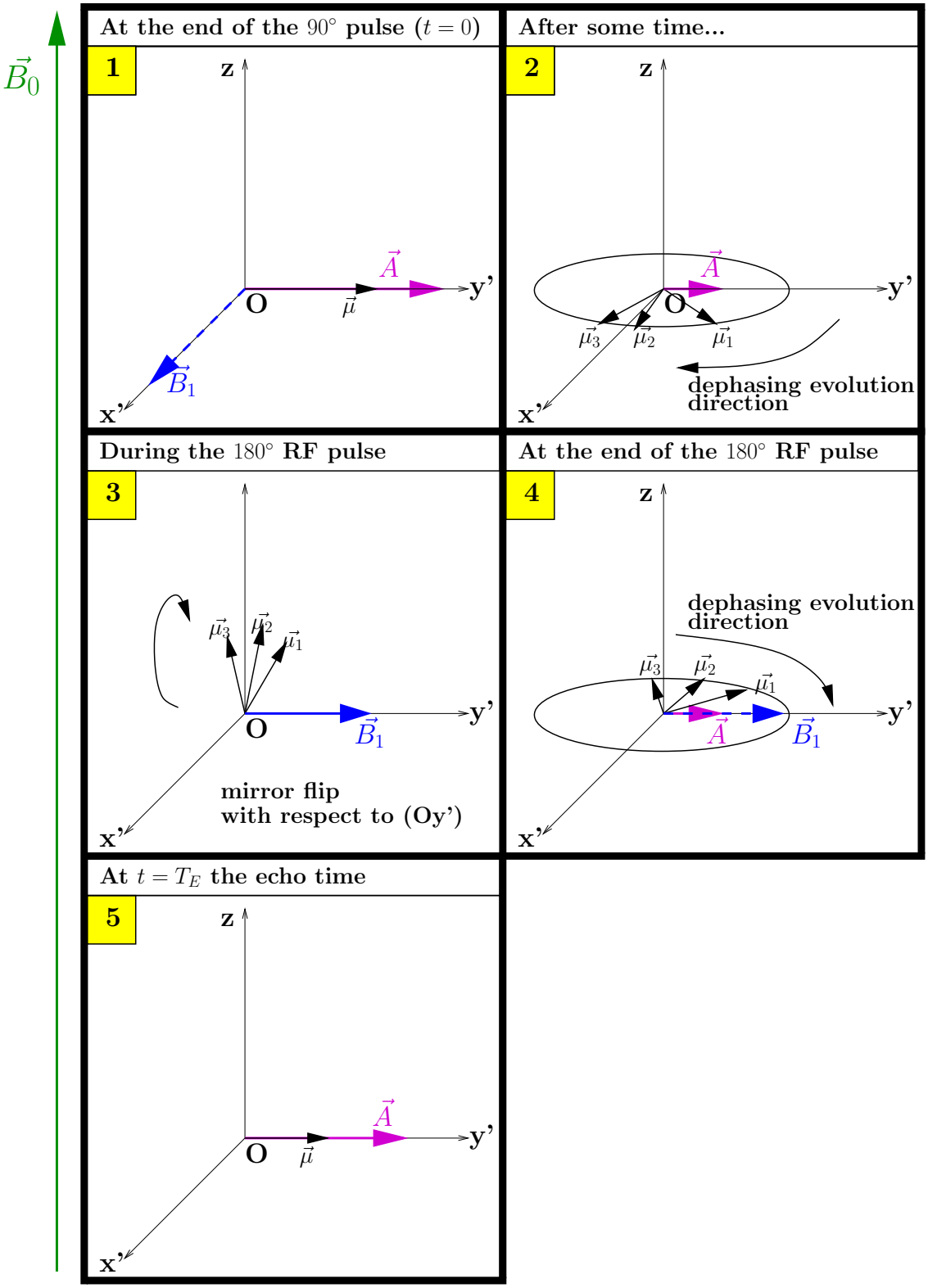


Figure 2.8: Detailed figure of the spin-echo process. The vignettes' numbers correspond to those indicated in fig. 2.7.

2.4 Contrast in MRI

In MRI, the measured signal is the transverse magnetization. To obtain contrast in the image, it is essential that this signal reveals the brain region that we want to study. And yet, depending on the brain tissue, the T_1 and T_2 relaxation times are not the same. These two times, as well as the hydrogen proton density, are given in table 2.1 for the cerebrospinal fluid (CSF), the grey matter (GM) and the white matter (WM).

	$T1$(s)	$T2$(ms)	Proton density
CSF	4	500	1
GM	1	80-90	0.92
WM	0.7	70-80	0.79

Table 2.1: Values of the T_1 and T_2 relaxation times, and of the hydrogen proton density at 1.5T in the cerebrospinal fluid (CSF), the grey matter (GM) and the white matter (WM). Data extracted from http://users.fmrib.ox.ac.uk/~peterj/lectures/hbm_1/sld031.htm

These relaxation times induce different longitudinal and transverse magnetization relaxation curves according to the tissue (fig. 2.9). Depending on the T_R and T_E parameters set up for the sequence, the contrast between the tissues is not the same. Indeed, a long T_R cancels the contrast in T_1, whereas a short T_E cancels the contrast in T_2. The table 2.2 sums up the parameters T_R and T_E to choose to get either a contrast due to the T_1 relaxation (then we do a T_1 weighting), or a contrast due to the T_2 relaxation (then we do a T_2 weighting). When the contrast in T_1 and T_2 are both canceled, the contrast in the image is only due to the proton density in the tissues. Fig. 2.10 allows to see the image results of these different weightings and to compare them with a histological section.

	Short T_R(∼14ms)	Long T_R(∼4s)
Short T_E	T_1 weighting	proton density weighting
Long T_E(∼100ms)	no contrast	T_2 weighting

Table 2.2: The T_R and T_E to choose depending on the desired weighting. For a T_1 weighting, the T_E to take is about 5ms and for a proton density weighting, the T_E to take is about 15ms. Data taken from http://users.fmrib.ox.ac.uk/~peterj/lectures/hbm_1/sld031.htm

We have seen in this section that the measured MRI signal can reveal a brain tissue. But, how do we know where this signal exactly comes from in the brain?

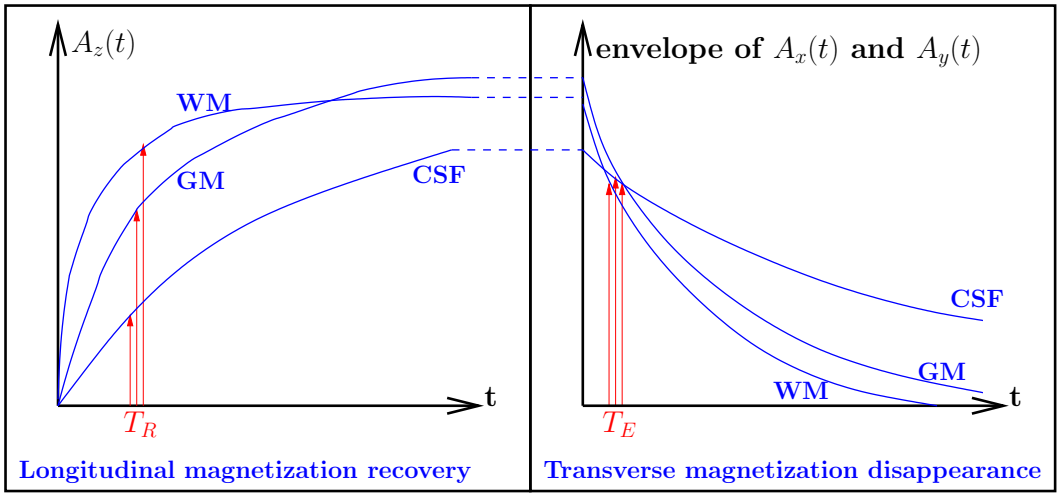


Figure 2.9: Longitudinal relaxation curve (left) and transverse relaxation curve (right) for three main brain regions: the CSF, the GM and the WM. According to the choice of T_R and T_E, priority is given to a contrast coming from the longitudinal relaxation (T_1 weighting) or from the transverse relaxation (T_2 weighting). For example, to get a T_1 weighting, we need an intermediate T_R (like in this figure) allowing to distinguish the longitudinal magnetizations of the different brain areas. Moreover, the contrast due to the transverse relaxation has to be canceled. For that, it is necessary to choose a very short T_E (like in this figure).

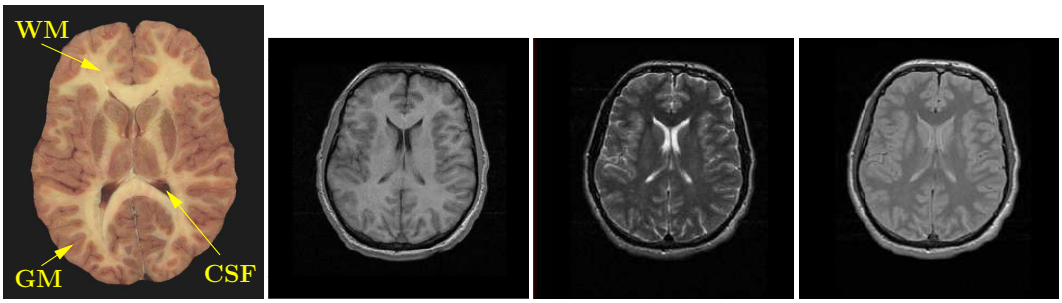


Figure 2.10: Histological section of a humain brain (image extracted from http://www.valley-neuroscience.com) and images showing three different contrasts in MRI at 1.5T: the T_1 weighting, the T_2 weighting and the proton density weighting (extracted from http://users.fmrib.ox.ac.uk/~peterj/lectures/hbm_1/sld031.htm).

2.5 The MRI signal spatial encoding

To create a brain image from the measured transverse magnetization signal, the spatial position of the protons which emit this MRI signal has to be determined. It is what is called the MRI signal spatial encoding, which is done using three pairs of gradient coils visible in the fig. 2.3. This technique that uses magnetic field gradients to spatially

encode the MRI signal was first invented by Herman Carr and Edward Purcell with only one gradient enabling to generate a 1D image [Carr and Purcell (1954)]. Twenty years later, Paul Lauterbur introduced the concept of 2D MR imaging combining three gradients [Lauterbur (1973)].

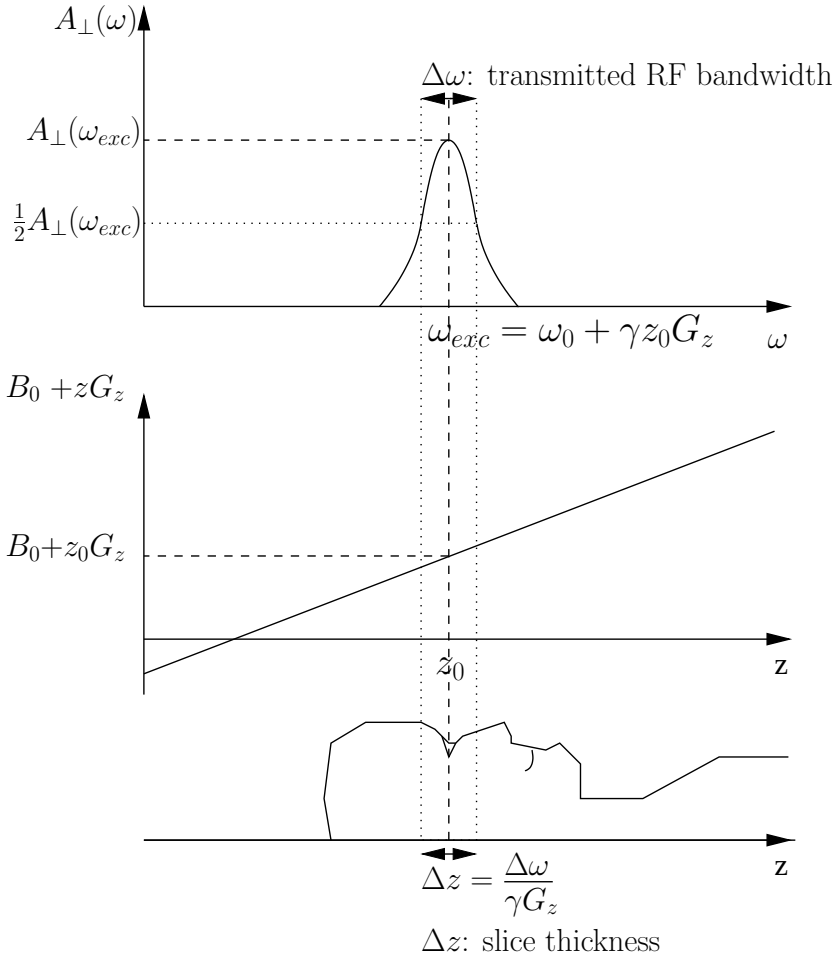


Figure 2.11: The gradient $\mathbf{G_z}$ permits to select a slice whose spins respect the resonance condition. The slice thickness is related to the transmitted RF bandwidth.

First, a magnetic field gradient $\mathbf{G_z}$ is created using a pair of gradient coils which axis is along the z-axis in our example. When only $\mathbf{B_0}$ and this gradient are applied, the spins precess at an angular velocity which depends on the spin position along the z-axis: $\omega(z) = \gamma B_0 + \gamma z G_z$. As a consequence, due to the Larmor resonance, an RF wave of angular velocity $\omega_{exc} = \omega_0 + \gamma z_0 G_z$ will excite the spins of a specific plane perpendicular to the z-axis and of equation $z = z_0$. Therefore, only the spins of this slice will see their magnetization flip. This slice selection is represented by fig. 2.11, which shows the relation between the slice thickness Δz and the transmitted RF bandwidth $\Delta\omega$ of the RF excitation pulse. This bandwidth corresponds to the full width at half maximum of the magnitude spectrum $A_\perp(\omega)$. As the gradient $\mathbf{G_z}$ permits to select a slice, it is also called the **slice**

selection gradient and is written $\mathbf{G_{SS}}$. It is not necessarily applied in the direction given by $\mathbf{B_0}$. This gradient $\mathbf{G_{SS}}$ is produced at the same time as the RF excitation pulse and lasts a limited time.

Once the slice has been selected, it is necessary to encode the MRI signal spatially in this slice using a 2D encoding with two phase and frequency encoding gradients. The **phase encoding gradient** $\mathbf{G_\Phi}$ (or $\mathbf{G_y}$), produced by a second pair of coils, modifies the spins' phases along the y-axis: the voxel lines in the slice therefore become distinguishable from each other because of their different phases. The gradient $\mathbf{G_\Phi}$ therefore allows to differentiate the lines in the slice with respect to their spins' phases. Finally, a third pair of coils causes a **readout gradient**, also called **frequency encoding gradient**, $\mathbf{G}_\omega$ (or $\mathbf{G_x}$) which distinguishes the slice columns along the x-axis with respect to their spin frequencies. This gradient is applied during the MRI signal readout. At this instant, the signal measurement enables to obtain the information of all frequencies and only one

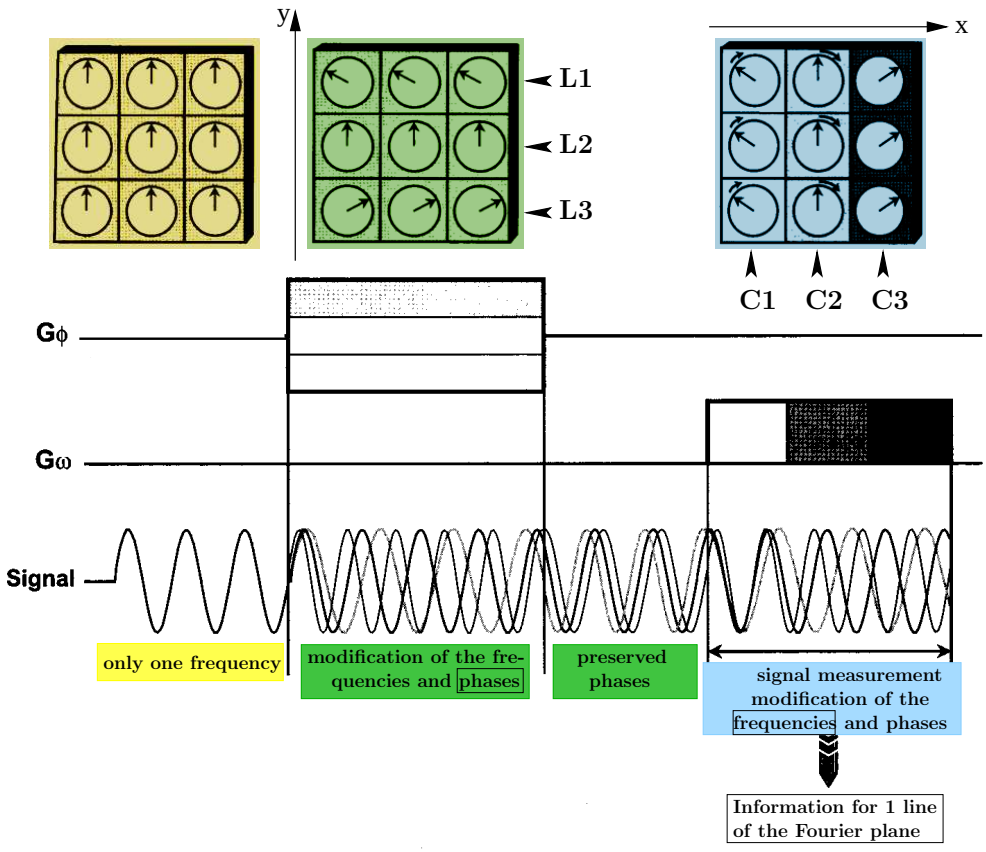


Figure 2.12: Effects of the gradients $\mathbf{G_\Phi}$ and $\mathbf{G}_\omega$ on the transverse magnetization signal: modification of the phases and frequencies possessed by the signals emitted at the different voxels of the slice. At the top, we can follow the schematic evolution of the magnetic moments in the slice voxels selected by $\mathbf{G_{SS}}$. The gradient $\mathbf{G_\Phi}$ permits to distinguish the different lines L1, L2, L3 in the slice according to their spin phases. The gradient $\mathbf{G}_\omega$ differentiates the columns C1, C2 et C3 according to their spin frequencies. Extracted from [Kastler et al. (2006)].

phase information, *i.e.* the frequential content of the line selected by $\mathbf{G_\Phi}$. This frequential content corresponds to a line in the Fourier plane. At each T_R, a new phase encoding gradient $\mathbf{G_\Phi}$ is applied which permits to fill in one new line of the Fourier plane. The repetition of this process enables to fill in the whole Fourier plane (in a sampled way). Fig. 2.12 sums up this MRI signal encoding process using the three gradients. The slice columns and lines encoded define together the **matrix** of pixels which will constitute the image. This image will be a representation of the **field of view** (FOV) that has been scanned by the magnet.

However, a problem appears: each three gradients $\mathbf{G_{SS}}$, $\mathbf{G_\Phi}$ and $\mathbf{G}_\omega$ causes a dephasing of the spins with respect to their position. This dephasing endures after each gradient application. If this effect is the one expected from the phase encoding gradient $\mathbf{G_\Phi}$, it is not the one expected from $\mathbf{G_{SS}}$ and $\mathbf{G}_\omega$. The phases caused by these two gradients have to be canceled. For that, bipolar gradients are used. A bipolar gradient has a positive and a negative lobe. The positive lobe is twice as long as the negative lobe, so that it cancels the dephasing at the temporal center of the 90° RF pulse for $\mathbf{G_{SS}}$ and at the temporal center of the echo readout for $\mathbf{G}_\omega$, as shown in fig. 2.13.

Once the entire signal spatial encoding system is set, how is the final image reconstructed?

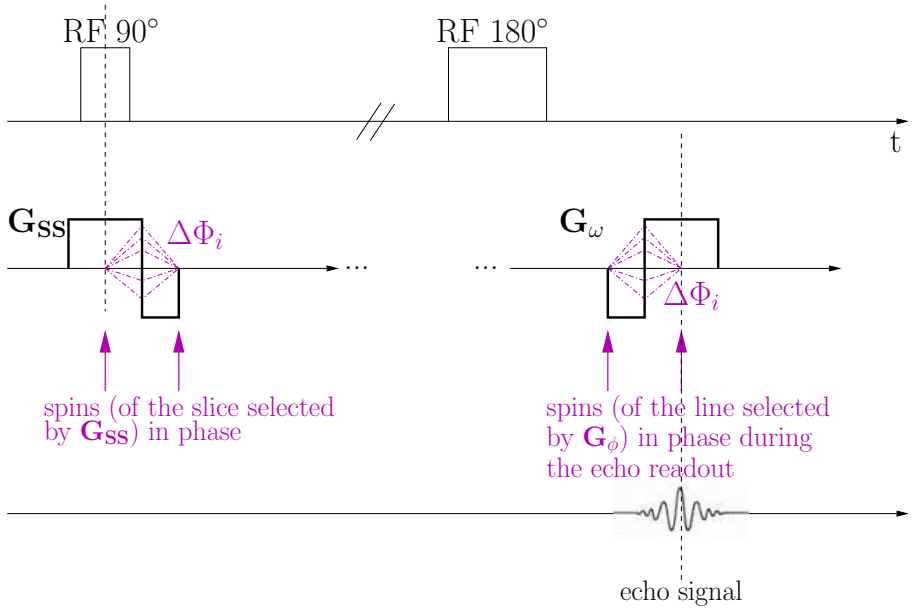


Figure 2.13: The slice selection gradient $\mathbf{G_{SS}}$ and the frequency encoding gradient $\mathbf{G}_\omega$ are designed as bipolar gradients. The gradient bipolarity allows, after the gradient application, to refocus the spins so that their phases present in the echo signal are only due to the phase encoding gradient $\mathbf{G_\Phi}$, which is not bipolar.

2.6 The image reconstruction

The magnetization signal emitted by the studied volume is measured by the receiving coil (or antenna). This coil measures a voltage $e(t)$, the expression of which is given by the Faraday induction law and by the principle of reciprocity [Haacke et al. (1999)]:

$$e(t) = -\frac{d\Phi(t)}{dt} = -\frac{d}{dt}\int_{\text{volume}} \mathbf{B}^{receive}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r},t) d^3\mathbf{r}, \tag{2.8}$$

with $\Phi(t)$ the magnetic flux induced in the coil by the magnetization, $\mathbf{B}^{receive}(\mathbf{r})$ the magnetic field created by the receiver coil at the position $\mathbf{r}$ for a unit current in the coil and $\mathbf{A}(\mathbf{r},t)$ the local magnetization. The x- and y-components of the field $\mathbf{B}^{receive}(\mathbf{r})$ can be written in terms of the receive magnetic field magnitude $B_\perp(\mathbf{r})$ and angle θ_B [Haacke et al. (1999)]:

$$\begin{cases} B_x^{receive}(\mathbf{r}) = B_\perp(\mathbf{r})\cos\theta_B, \\ B_y^{receive}(\mathbf{r}) = B_\perp(\mathbf{r})\sin\theta_B. \end{cases}$$

The authors of [Haacke et al. (1999)] calculated $e(t)$ for the case where no field gradient is applied. They showed that, when computing $e(t)$, the derivative of the e^{-t/T_2^*} factor can be neglected. It gives:

$$e(t) \propto \omega_0 \int_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(\mathbf{r},0) B_\perp(\mathbf{r}) \sin\left(\omega_0 t + \theta_B(\mathbf{r}) - \phi_0(\mathbf{r})\right) d^3\mathbf{r}, \tag{2.9}$$

with the phase $\phi_0(\mathbf{r})$ and the magnitude $A_\perp(\mathbf{r},0)$ determined by the initial RF pulse conditions. Now, when considering all gradient fields, the previous resonance angular frequency ω_0 has to be replaced by the new resonance angular frequency, which the spins precess with and which is given by: $\omega(x,y,z,t) = \gamma B_0 + \gamma z G_z(t) + \gamma x G_x(t) + \gamma y G_y(t)$. A demodulation followed by a low pass filtering applied to $e(t)$ leads to a data storage in two "coil components", a **real** one and an **imaginary** one. The resulting received complex signal is expressed such that:

$$s(t) \propto \omega_0 \int_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(\mathbf{r},0) B_\perp(\mathbf{r}) e^{i[(\Omega-\omega(\mathbf{r},t))t + \phi_0(\mathbf{r}) - \theta_B(\mathbf{r})]} d^3\mathbf{r}, \tag{2.10}$$

where Ω is the demodulation angular frequency. The transmitting and receiving RF coils are assumed to be sufficiently uniform, so that ϕ_0, θ_B and $B_\perp$ are independent of the position $\mathbf{r}$. The signal $s(t)$ can be written such that:

$$\boxed{s(t) = \omega_0 \Lambda B_\perp \int_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(\mathbf{r},0) e^{i[\Omega t - \phi(\mathbf{r},t)]} d^3\mathbf{r}}, \tag{2.11}$$

with Λ the factor of proportionality and $\phi(\mathbf{r},t) = -\int_0^t dt' \omega(\mathbf{r},t')$ the accumulated phase. When setting Ω equal to $\gamma B_0 + \gamma z G_z(t)$, the complex signal $s(t)$ becomes:

$$s(t) = \omega_0 \Lambda B_\perp \iint_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(x,y,z,0) e^{-i\gamma\left(x\int_0^t G_x(t')dt' + y\int_0^t G_y(t')dt'\right)} dxdy. \tag{2.12}$$

Let us write $k_x(t) = \frac{\gamma}{2\pi}\int_0^t G_x(t')dt'$ and $k_y(t) = \frac{\gamma}{2\pi}\int_0^t G_y(t')dt'$ defining the **Fourier plane**, also called **k space**. The following Fourier relation is then obtained:

$$\boxed{s(k_x(t),k_y(t)) = \omega_0 \Lambda B_\perp \iint_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(\mathbf{r},0) e^{-i2\pi(k_x(t)x+k_y(t)y)} dxdy}. \quad (2.13)$$

Incrementing the two frequency and phase encoding gradients $\mathbf{G_x}$ and $\mathbf{G_y}$ enables to sample the Fourier plane. Then, the double inverse fast Fourier transform of $s(k_x(t),k_y(t))$ performed for $t = T_E$ yields the final signal:

$$\boxed{s_{final}(x,y) = \omega_0 \Lambda B_\perp \int e^{-T_E/T_2^*(x,y,z)} A_\perp(x,y,z,0) dz}. \quad (2.14)$$

Then, the modulus of $s_{final}(x,y)$ allows to reconstruct the final MRI image in the spatial domain. This technique utilizing the inverse Fourier transform tool was developed by Richard Ernst and his team [Kumar et al. (1975)]. Fig. 2.14 represents the chronogram of a spin-echo sequence showing how the RF pulses and the gradients are put together in order to sample the Fourier plane. Finally, fig. 2.15 shows the image obtained in the Fourier plane, as well as the image obtained in the spatial domain after a double inverse Fourier transform.

To obtain the image of each slice of the scanned volume, the sequence in fig. 2.14 has to be reproduced for different slice selection gradients. To not excessively extend the sequence duration, we can take advantage of the dead time T_R-T_E to successively start the sequences for the other slices. This technique is called the **multislice technique**. However, the spin-echo sequence, even with the multislice technique, still lasts 15 min at least. The longer the sequence, the more laborious the exam, and this especially because the subject has to stay very still in the MRI machine, as the latter is very sensitive to movements. Indeed, any movement of the subject can create artifacts on the images. Staying still in the machine for the subject is easier if the sequence is short. Different imaging methods permit to reduce the sequence duration. One of them is parallel MRI.

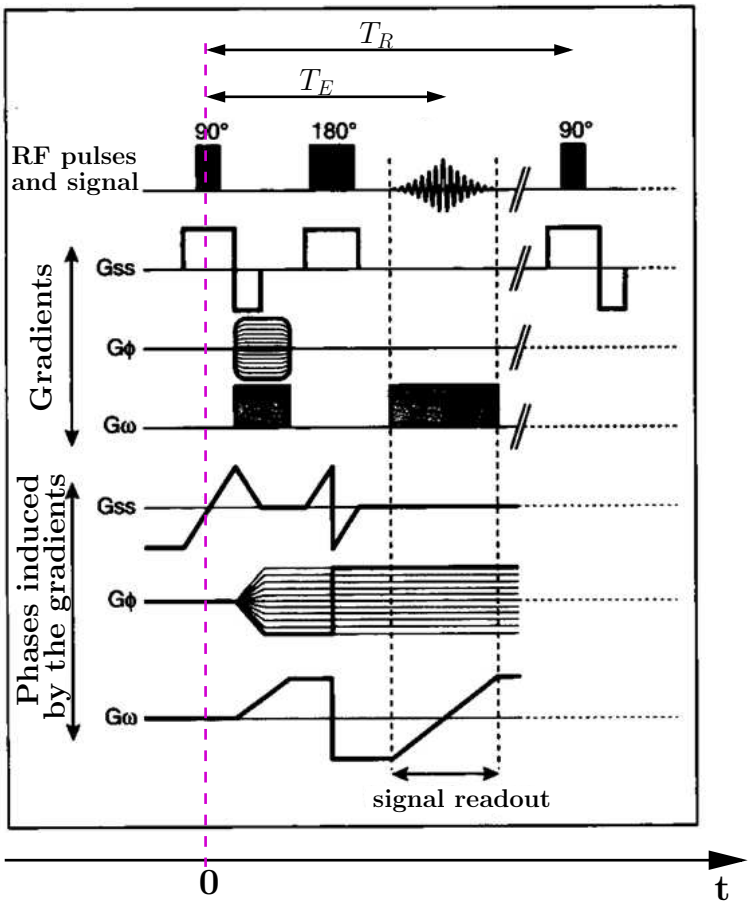


Figure 2.14: Spin-echo sequence chronogram. The figure shows how the RF pulses and the gradients are put together for a T_R duration. The phases induced by the gradients are also indicated. Extracted from [Kastler et al. (2006)].

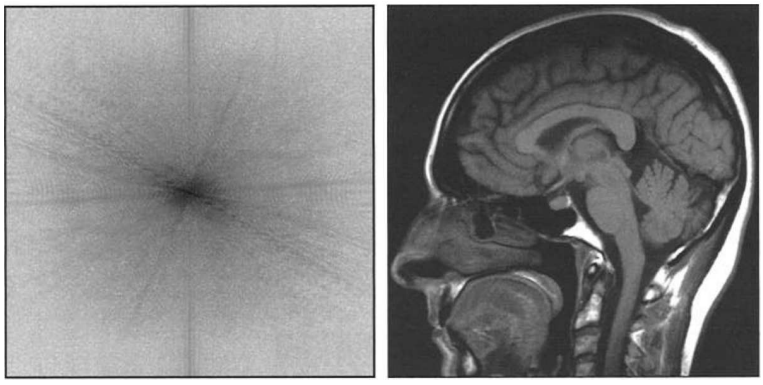


Figure 2.15: Magnitude in the **k space** (left) and in the spatial domain (right). To switch from the left image to the right one, a 2D inverse Fourier transform has to be applied. Extracted from [Kastler et al. (2006)].

2.7 Parallel MRI techniques

Parallel MRI consists of utilizing, for the MRI signal reception, a phased-array coil, *i.e.* an array composed by several surface coils or **channels** arranged in a network [Roemer et al. (1990)], as shown in fig. 2.16. Each channel is therefore getting the signal of the anatomical region in front of which it is located. After a demodulation followed by a low pass filtering applied on each channel signal, the real and imaginary components of each signal are obtained. These different complex signals, obtained in parallel by the coil array, are expressed such that:

$$m_l(k_x, k_y) = \iint s_{final}(x, y) c_l(x, y) e^{-i2\pi(k_x x + k_y y)} dx dy, \tag{2.15}$$

with $c_l(x, y)$ the sensitivity of the l^{th} coil measuring the signal $m_l(k_x, k_y)$. Then, these data $m_l(k_x, k_y)$ are combined together to form a global signal. This method permits to enhance the signal-to-noise ratio (SNR). Indeed, the latter is much higher than the SNR obtained using a volumic coil exploring the same FOV.

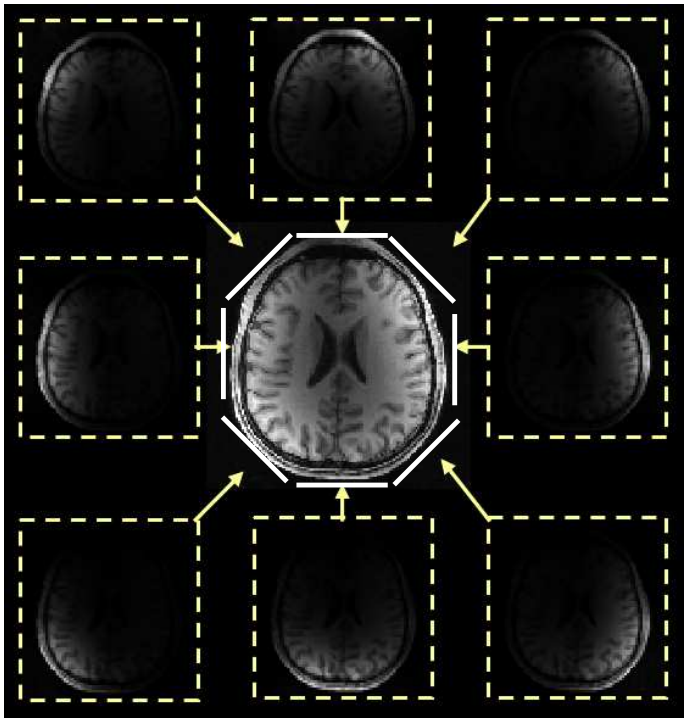


Figure 2.16: Phased-array head coil diagram with eight receiving coils represented by white lines around the centered image. Next to each coil, the signal received by the coil is shown. The central image corresponds to the recombined image. Fig. adapted from http://black.bme.ntu.edu.tw/courses/course_neuroimaging_fall10/slide_pmri_fhlin.pdf.

The aim of parallel MRI is to use the spatial information obtained from the different coils which receive the MRI signal in parallel. The idea is that the sensitivity encoding of each coil can partially replace the spatial encoding performed in classical MRI by the phase

and/or partition encoding gradients. Then the number of gradient encoding steps can be reduced which leads to an acceleration of the sequence. The first concept of parallel MRI appeared with [Carlson and Minemura (1993)], but the first practical implementation came out with [Sodickson and Manning (1997)] as the **simultaneous acquisition of spatial harmonics** (SMASH). The latter is a reconstruction method operating in the **k**-space. It linearly combines the sensitivities of the different coils to obtain spatial harmonics, such that:

$$\sum_{l=1}^{L} n_l^{(m)} c_l(x,y) \approx e^{im\Delta k_y y}, \tag{2.16}$$

with L the number of coils, $n_l^{(m)}$ the linear weights used to produce, with the sensitivities $c_l(x,y)$ of each l^{th} coil, the composite sensitivity profile $e^{im\Delta k_y y}$ with m an integer and $\Delta k_y = 2\pi/\text{FOV}$. These spatial harmonics $e^{im\Delta k_y y}$ enable to reconstruct the missing lines of the **k**-space. Indeed, eq. 2.15 becomes:

$$m_l(k_x,k_y) = \iint s_{final}(x,y) e^{im\Delta k_y y} e^{-i2\pi(k_x x + k_y y)} dxdy. \tag{2.17}$$

The spatial harmonics actually replace the gradient-induced modulation of the magnetization: less lines of the Fourier plane than usually are acquired and the missing lines are reconstructed using eq. 2.17. In practice however, this reconstruction algorithm requires specific coils which renders the method difficult to achieve.

A more practical solution than SMASH was developed by [Pruessmann et al. (1999)] and was called **sensitivity-encoding for fast MRI** (SENSE). In opposition to SMASH, SENSE reconstructs the missing lines of the FOV in the image domain. As shown in fig. 2.17 for an acquisition with an acceleration factor R of 2, when the subsampled **k**-space is reconstructed in the image domain, if the object to examine is bigger than the FOV, it yields an image with an aliasing artifact. Another consequence is that the FOV dimension along the y-axis is reduced by half. The SENSE algorithm "unfolds" the image directly using the sensitivity of each coil, such that, for each l^{th} coil:

$$m_l(x,y) = c_l(x,y) s_{final}(x,y) + c_l(x, y + \frac{\text{FOV}_y}{2}) s_{final}(x, y + \frac{\text{FOV}_y}{2}), \tag{2.18}$$

with $m_l(x,y)$ the signal intensity in a voxel in the sub-sampled image, in an aliased region and FOV_y the y-dimension of the FOV. The other notations are the same as before. Performing eq. 2.18 for each coil permits to reconstruct the global image without any aliasing.

After SENSE, other reconstruction solutions have been proposed, such as the **generalized autocalibrating partially parallel acquisition** (GRAPPA) algorithm [Griswold et al. (2002)]. The latter reconstructs the **k**-space missing lines in the Fourier domain, before the double inverse Fourier transform application, as in the SMASH method. Contrary to SMASH, it uses additionally acquired central **k**-space lines called auto-calibration signal (ACS) lines. The latter permit to automatically obtain the linear weights $n_l^{(m)}$, which are

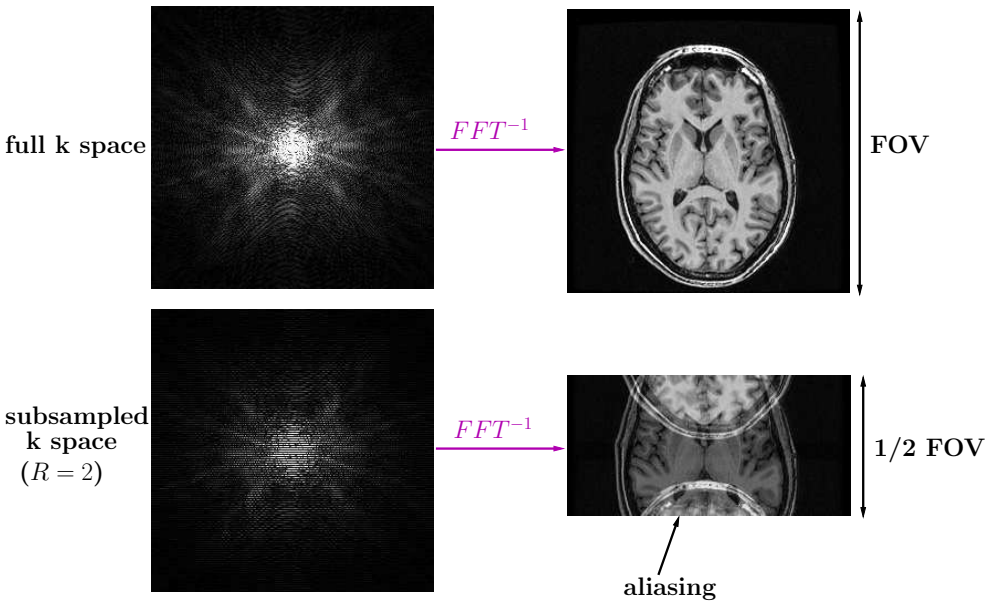


Figure 2.17: Subsampling of the **k** space with a sampling frequency more than twice the highest frequency in the image (top) (respect of the Nyquist-Shannon theorem), and then application of the double inverse Fourier transform to obtain the final image in the image domain. Subsampling of the **k** space (bottom) with an acceleration factor R of 2. The FOV is then divided by 2 along the y-axis. As the object to examine is bigger than the FOV, an artifact called aliasing appears in the image obtained after the double inverse Fourier transform. Image adapted from [Poupon (2008)].

necessary for the reconstruction of the missing lines. To determine these weights, multiple lines from all coils are used to fit an ACS line of a coil, following:

$$s_l^{\mathrm{ACS}}(k_y - m\Delta k_y) \approx \sum_{l=1}^{L} n_l^{(m)} s_l(k_y), \tag{2.19}$$

with $s_l^{\mathrm{ACS}}(k_y - m\Delta k_y)$ the ACS signal from the l^{th} coil for a line offset by $m\Delta k_y$ (m an integer and $\Delta k_y = 2\pi/\mathrm{FOV}$ as before). $s_l(k_y)$ is the acquired line at k_y from the l^{th} coil. Eq. 2.19 permits to give the linear weights $n_l^{(m)}$. From them, the missing lines can be determined, using a blockwise reconstruction following:

$$m_j(k_x, k_y - m\Delta k_y) = \sum_{l=1}^{L} \sum_{b=0}^{N_b-1} n_l^{(m)}(j,b) s_l(k_y - bR\Delta k_y), \tag{2.20}$$

where N_b is the number of blocks —a block being defined as a single acquired line and R-1 missing lines— and R is the acceleration factor. $n_l^{(m)}(j,b)$ is the linear weight defined for the block b and the j^{th} coil.

These parallel MRI reconstruction techniques are commonly used to save acquisition time. In the diffusion MRI modality, which we will present in the next chapter, parallel

MRI is very interesting. Indeed, acquisition times in diffusion MRI are usually long, yielding geometrical distortions caused by phase artifacts. Therefore, the use of parallel MRI permits to reduce these distortions, in addition of saving acquisition time. The main disadvantage of the parallel MRI reconstruction technique is that the SNR is reduced with respect to the non-accelerated sequence, at least by a factor $\sqrt{R}$ [Blaimer et al. (2004)]. For the SENSE reconstruction, it was established that [Pruessmann et al. (1999)]:

$$\mathrm{SNR}_{\mathrm{SENSE}} = \frac{\mathrm{SNR}_{\mathrm{FULL}}}{g\sqrt{R}}, \tag{2.21}$$

where $\mathrm{SNR}_{\mathrm{SENSE}}$ is the SNR on the SENSE reconstructed image and $\mathrm{SNR}_{\mathrm{FULL}}$ is the SNR on the image obtained with the non-accelerated sequence. g is the geometry factor which depends on the geometry of the coil array. Furthermore, we will see later that the use of parallel MRI changes the noise distribution in the data, compared to a single-coil acquisition. We will explain how the choice of the reconstruction method and the number of channels impact the noise distribution definition.

2.8 Conclusion of this chapter

We have presented the principle of magnetic resonance imaging, which opens the way to the generation of images offering contrasts corresponding to different brain tissues. We have explained how the spatial encoding works and how the image reconstruction is done, using an inverse Fourier transform. The steps of the signal reconstruction have an impact on the noise distribution, as we will see in chapter 5. Finally, we have mentioned the possibilities to accelerate the MRI sequences using parallel MRI. Again, the use of multiple coils to acquire the signal, as well as the reconstruction algorithms used in parallel MRI have also some effects on the noise corrupting the final image. We will come back to it later in the thesis.

This thesis is fully dedicated to diffusion MRI, for which noise issues are essential. Diffusion MRI is based on the same principles as MRI, with the addition of the measurement of diffusion inside the brain. It permits to produce images revealing the anatomical connectivity of the brain. Our noise removal work, which we will present later, was performed for those images.

Chapter 3

Knowledge about diffusion MRI (dMRI) used in this thesis

We have seen in the previous chapter that because brain tissues have different relaxation times, MRI can be used to obtain images of the brain with contrast. For this reason, when choosing the right MRI sequence parameters, it is possible to see anatomical brain structures. Is it possible to go further in the observation of anatomical structures? Can we have access to other contrasts and to more details on the brain tissue microstructure? Answers to these questions are and continue to be given by the diffusion MRI (dMRI) technique. dMRI relies on the sensitization of MRI to the random motion of water molecules in tissues. At the beginning of MRI, the Brownian motion due to the diffusion process of water molecules was considered as an artifact that researchers attempted to reduce. In the eighties however, the biophysical phenomenon behind these artifacts interested researchers, as they discovered that the random motion of the water molecules inside the brain permitted to obtain anatomical information at a microscopic scale which had never been achieved before. In this chapter, we first explain the biophysical phenomenon of water diffusion in tissues and how to sensitize MRI to the diffusion process. Then, we will see in details how mathematical models can be developed to encapsulate the diffusion process and how they can be exploited to infer the structural connectivity and to probe the tissue organisation at a cellular level. Finally, we concern ourselves with tractography, a tool whose aim is to represent WM fiber bundles in the most realistic way. This chapter is inspired in majority by courses given by Dr. Cyril Poupon on dMRI [Poupon (2010)] and by thesis chapters of [Poupon (1999a)], [Tuch (2002)], [Descoteaux (2008)] and [Marrakchi-Kacem (2011)].

3.1 dMRI basis principles

dMRI is an MRI technique which relies on the measurement of the brain water molecules' Brownian motion. At first, this Brownian motion was only seen as producing an artifact on NMR signals: Erwin Hahn noticed that the spin echo signal was attenuated because of the diffusion of the spins [Hahn (1950)]. The Bloch equations presented in the previous

chapter (eq. 2.2) were modified to account for the diffusion [Torrey (1956)]. A dozen years later, the pulse gradient spin echo sequence was proposed by [Stejskal and Tanner (1965)] enabling to measure the coefficients of the molecular diffusion from diffusion NMR. With the advent of MRI in the seventies, dMRI was also rendered possible. This new modality appeared in 1985 [LeBihan and Breton (1985); Merboldt et al. (1985); Taylor and Bushell (1985)]. An important clinical application of dMRI since its discovery is the early detection of **ischemia** and **cerebrovascular accidents** [Moseley et al. (1990); LeBihan et al. (1992)]. In such pathologies, an œdema appears and slows down the water diffusion which is visible on the dMRI images. Contrary to conventional MRI, dMRI reveals ischemic regions immediately and is consequently used in clinical routine today to this aim.

More than a diagnosis tool for cerebral ischemia, dMRI is currently the only non invasive way to explore *in vivo* the **anatomical connectivity** in a brain relying on the measurement of the brain water molecules' Brownian motion. Indeed, the water diffusion is modulated by the geometry of the WM fiber bundles. Consequently, dMRI permits to infer information on these connection structures. This method enables first to detect eventual anomalies of the nervous fibers and secondly to study the neuronal connections' functioning in cognitive science.

To understand how the **cerebral microstructure** is reconstructed with dMRI, we first describe the physical water diffusion phenomenon and how this phenomenon depends on the WM tissue geometry. Then, we explain how dMRI measures this water molecules' movement.

3.1.1 The physical water diffusion phenomenon

In this subsection, we explain the physical water diffusion phenomenon that the dMRI technique aims to measure.

In every environment, water molecules undergo the thermal agitation. They move and bang together randomly. In an environment without any obstacle or particular restriction, molecules behave following a **Brownian motion** [Brown (1828)]. It is the case for example for colouring agent molecules that we put in water (far before they go near the container walls). This Brownian motion can be modeled by a random walk, as shown in fig. 3.1 (a). A diffusion process, whatever the environment, is described by Fick's first law of diffusion [Fick (1855)] that links the colouring agent current density to their concentration in the environment:

$$J = -D\nabla C, \tag{3.1}$$

with J the molecules' current density (in molecules $\cdot s^{-1} \cdot m^{-2}$), D the diffusion coefficient of the molecules in the environment (in $m^2 \cdot s^{-1}$) and C the concentration of molecules (in molecules $\cdot m^{-2}$) (∇ is the nabla operator). The **diffusion coefficient** D is a measure of the molecules' capacity to diffuse depending on the environment, on the temperature and on the molecules in question.

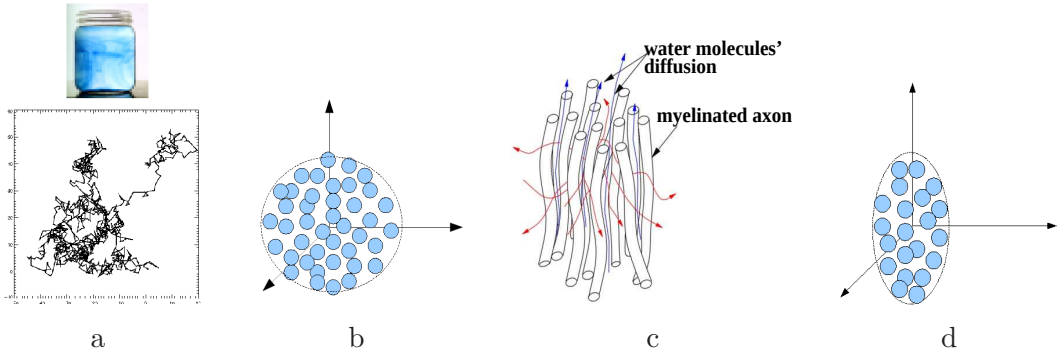


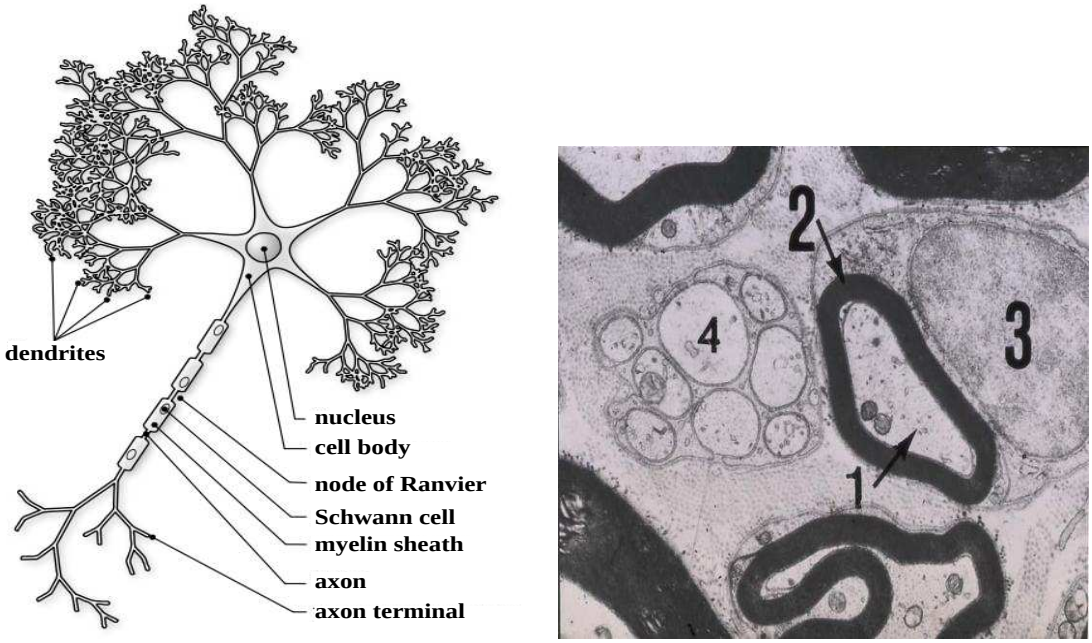
Figure 3.1: Representation of the diffusion phenomenon and of the mean free path in an environment without any obstacle (or whose obstacles have negligible effect like in a liquid) (a and b) and in the WM (c and d). In (a), the diffusion is free and therefore is an isotropic process, as indicated by (b), where the blue circles correspond to the colouring agent molecules. In (c), the diffusion is restricted inside axons and hindered in between. Consequently, the diffusion process is anisotropic in the WM, with favoured diffusion directions along axons (blue arrows) and inhibited diffusion directions perpendicular to them (red arrows). Fig. (a) extracted from Wikipedia, fig. (b) and (d) extracted from [Poupon (2010)] and fig. (c) extracted from [Poupon (1999a)].

Einstein showed that in an environment without any obstacle or restriction, D is linked to the mean square displacement travelled by molecules during the laps τ written as: $\langle \mathbf{r}^{\mathrm{T}}\mathbf{r} \rangle$, with $\mathbf{r}$ corresponding to the displacement vector of one molecule during τ, $\mathbf{r}^{\mathrm{T}}$ being its transpose and finally $\langle \cdot \rangle$ being the mean operator. This link is given by [Einstein (1905)]:

$$D = \frac{\langle \mathbf{r}^{\mathrm{T}}\mathbf{r} \rangle}{6\tau}. \tag{3.2}$$

In an environment where the diffusion process is isotropic, the mean free path of a particle (*i.e.* its mean displacement between two collisions) is the same whatever the movement direction. On the other hand, in biological tissues, the molecules' mobility depends on the tissue microstructure components such as cytoplasmic membranes, myelin, cytoskeleton and all cell organelles that constitute physical obstacles to the diffusion process (fig. 3.2). More particularly in WM, diffusion is restricted inside axons and hindered in between. Moreover, the process is favoured along axons and inhibited in the perpendicular directions. The water molecules' mean free path is therefore higher if the molecules' movement follows the fiber directions. Fig. 3.1 illustrates the modulation of the diffusion phenomenon by the WM fiber architecture. In this WM environment full of obstacles, we can measure the **apparent diffusion coefficient** (ADC), which is different from the true diffusion coefficient D because diffusion is here not only governed by Brownian motion, but also runs up against the environment architecture. The ADC value is around $3 \cdot 10^{-3} mm^2 \cdot s^{-1}$ in the CSF, $0.8 \cdot 10^{-3} mm^2 \cdot s^{-1}$ in the GM and between 0.2 and $1.2 \cdot 10^{-3} mm^2 \cdot s^{-1}$ in the WM (source: [http://www.irmresonance.over-blog.com/article-13469878.html]).

Diffusion is characterized by the water molecules' **displacement probability density function**, also called **propagator** and written as $P(\mathbf{r}, \tau)$. This function represents the



a Structure of a classical neuron

b Transverse section of a WM region

Figure 3.2: In (a), a neuron is represented. Fig. (a) extracted from Wikipedia. In (b), the transverse section of a WM region shows a myelinated axon (n°1) and an unmyelinated axon (n°4). The myelin is the black circle around the fibers (n°2) that permits their electric isolation. N°3 indicates a Schwann cell, which provides the myelination. Fig. (b) extracted from http://zarrouk.e-monsite.com/pages/content/le-tissu-nerveux.html

water molecule probability to move of a distance $\mathbf{r}$ during a time τ. This information contained in the propagator should allow to infer the tissue geometry at a microscopic level. We will see in section 3.2 that different local diffusion models were proposed to reconstruct the propagator. Before that, we first explain how to sensitize an MRI image to the diffusion process.

3.1.2 The diffusion phenomenon measurement with dMRI

The acquisition of diffusion weighted (DW) images requires modifications of the MRI spin-echo sequence, seen before in section 2.3, so that it is possible to measure the water molecules' motion. This leads to the **pulse gradient spin echo** (PGSE) sequence developed by [Stejskal and Tanner (1965)] and illustrated in fig. 3.3. Like the spin-echo sequence, it is made up of two RF pulses at 90° and 180°, the latter one permitting to cancel the dephasing of the spins due to environment inhomogeneities. The PGSE also contains the gradient $\mathbf{G_{SS}}$ permitting to select one slice of the examined object, as well as the phase encoding gradient $\mathbf{G_{\Phi}}$ and the frequency encoding gradient $\mathbf{G}_{\omega}$.

What distinguishes the PGSE sequence from the spin-echo sequence is that it contains two **diffusion gradients g(o)** oriented in the direction indicated by the unitary vector $\mathbf{o}$ used to tag the position of spins along the direction $\mathbf{o}$. These diffusion gradients are applied

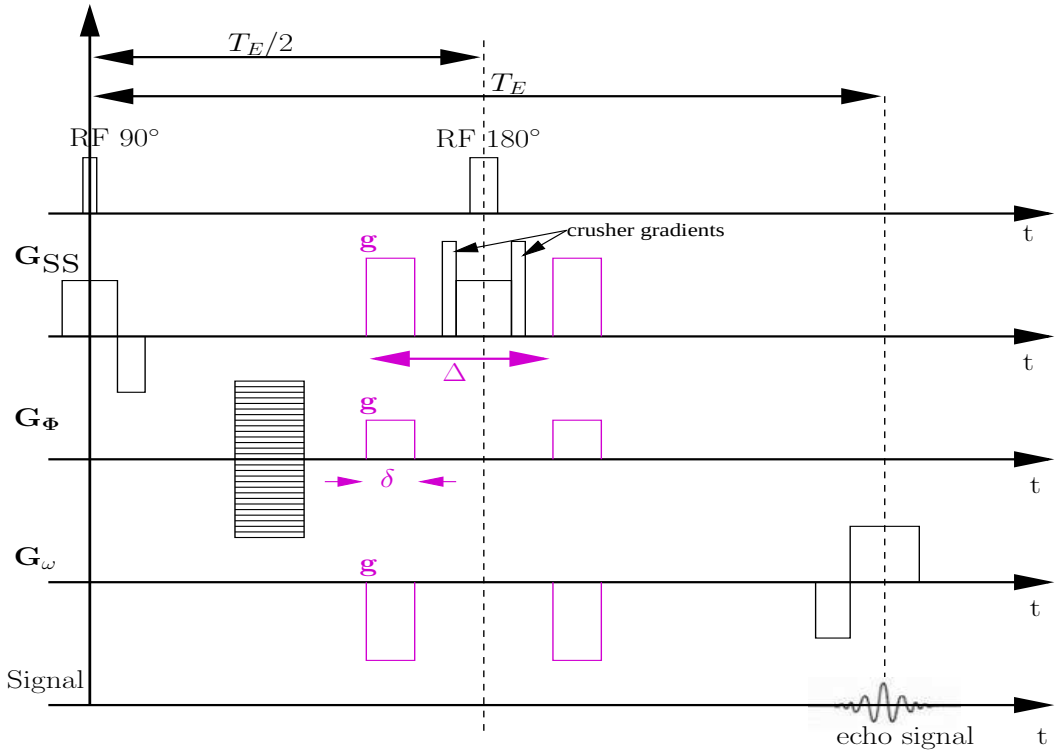


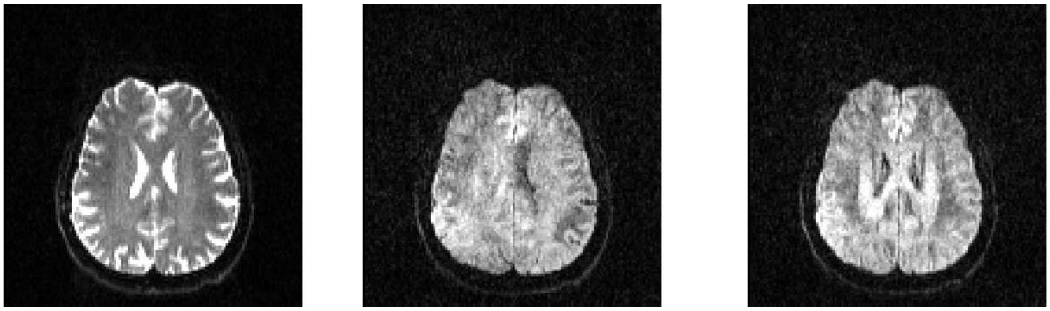
Figure 3.3: PGSE sequence. To measure the diffusion process, two diffusion gradients (in magenta), written as $\mathbf{g}$, are added, one before the 180° RF pulse and the other after. (We also added on the figure the crusher gradients, which are used to compensate the non ideal refocusing pulse to avoid artifacts due to unwanted signals. These crusher gradients are commonly used in MRI, and are not specific to dMRI.)

during a time δ. They are separated by the time Δ. The first diffusion gradient, applied before the 180° RF pulse, causes a dephasing of the spins that depends on their position. The dephasing of the motionless spins is cancelled by the second diffusion gradient after the 180° RF pulse. However, the dephasing of the spins that moved is not compensated and a dephasing remains that produces a reduced spin echo signal. Indeed, the measured DW signal S is lower than the T_2-weighted signal written as S_0, which is obtained without the application of any diffusion gradient. The measured signal S_0 corresponds to the magnitude of the signal $s_{final}(x, y)$ expressed in the previous chapter, as there was no diffusion gradient. For a spin-echo sequence, as there is a refocusing pulse, the field inhomogeneities are compensated. Therefore, the T_2^* signal in eq. 2.14 has to be replaced by the T_2 signal. Thus, the signal S_0 is proportional to $\exp[-T_E/T_2]$. The measured DW signal S corresponds to the term $|s_{final}(x, y)|$ obtained with the application of diffusion gradients. When considering the previously introduced diffusion propagator as a Gaussian function, we show thereafter that we can write:

$$S(b, \mathbf{g}) = S_0 \exp[-b\mathrm{ADC}(\mathbf{o})] \propto \exp[-T_E/T_2]\exp[-b\mathrm{ADC}(\mathbf{o})], \tag{3.3}$$

where b in $s \cdot mm^{-2}$ is the **diffusion weighting parameter**, also called *b-value* [LeBihan (1991)]. In the case of rectangular gradients: $b = \gamma^2 \|\mathbf{g}\|^2 \delta^2 (\Delta - \delta/3) = \gamma^2 \|\mathbf{g}\|^2 \delta^2 \tau$ (γ

being the proton gyromagnetic ratio seen in section 2.2 and $\|\cdot\|$ being the Euclidean norm operator), with $\tau = \Delta - \delta/3$ the **effective diffusion time** composed by a corrective term $\delta/3$ due to the diffusion occurring during the diffusion gradients' application times. In eq. 3.3, the parameter ADC($\mathbf{o}$) (in $mm^2 \cdot s^{-1}$) measures the mean mobility degree of the water molecules along the direction $\mathbf{o}$. Fig. 3.4 shows DW images, constructed from the signal S measured at each voxel of the FOV volume for two different diffusion orientations $\mathbf{o}$. When the molecules mainly diffuse along the diffusion orientation of interest, its corresponding signal becomes highly attenuated and the corresponding voxel appears dark.



T_2-weighted image | DW image $\mathbf{g}_i = (0.826, -0.121, 0.550)^{\mathrm{T}}$ | DW image $\mathbf{g}_i = (0.487, -0.137, 0.862)^{\mathrm{T}}$

Figure 3.4: T_2-weighted image and DW images for $b = 700s \cdot mm^{-2}$, with two different diffusion orientations $\mathbf{o}$, given by $\mathbf{g}_i$. These are images of a healthy human brain (as most of the brain images in the whole manuscript).

Such DW images are sufficient to diagnose ischemia. Indeed, when ischemia happens, the DW signal is instantaneously modified, thus enabling a very rapid detection of the suffering tissue [Moseley et al. (1990)]. dMRI is not limited to this particular medical detection. It can also explore the microscopic architecture of the brain, using a local modeling of the diffusion process.

3.2 Local modeling of the diffusion process

The water molecules' diffusion information is entirely contained in the diffusion propagator $P(\mathbf{r}, \tau)$, introduced in subsection 3.1.1. Is it possible to determine this function? It is precisely the aim of the local diffusion models. In this section, we first define the several functions containing information derived from the diffusion process and we introduce the notion of $\mathbf{q}$ space providing an adequate space to define local models. In particular, we present the historical **diffusion tensor imaging** (DTI) model, as well as the various **Q-Ball imaging** (QBI) and the **spherical deconvolution** (SD) models.

3.2.1 Definitions: dODF, cdODF, fODF, q space

With dMRI, the acquisition of the DW signal is performed using an adequate sampling of the diffusion propagator space and yields a large amount of DW volumes, each corresponding to a specific diffusion sensitization. These DW volumes are then used to infer a local model of the diffusion process within the brain. During the last twenty years, a plethora of local diffusion models were proposed to perform a synthesis operation on the DW data. These local models reconstruct different functions, all derived from the diffusion propagator.

The propagator $P(\mathbf{r}, \tau)$ is the proton displacement probability density function, with $\mathbf{r}$ the proton displacement vector and τ the effective diffusion time as defined previously. Inferring $P(\mathbf{r}, \tau)$ would provide the highly probable pathways of interest used by the protons in the brain. However, reconstructing the propagator itself requires a very long acquisition time which is not easy to realize in practice. Instead of considering the whole propagator, it can be easier to consider only a part of it. As we are interested in defining the routes of the protons, the necessary information can be obtained using the radial projection of the diffusion propagator. This radial projection was for a long time defined such that [Tuch (2004)]:

$$\Psi(\theta_i, \phi_i) = \int_0^\infty P(r, \theta_i, \phi_i, \tau) dr. \tag{3.4}$$

$\Psi(\theta_i, \phi_i)$ is called the **diffusion orientation distribution function** (dODF). Actually, this definition is erroneous. Indeed, in the spherical coordinate system, the propagator radial projection is rightly written such that [Tristán-Vega et al. (2009); Aganj et al. (2010)]:

$$\Psi_c(\theta_i, \phi_i) = \int_0^\infty P(r, \theta_i, \phi_i, \tau) r^2 dr, \tag{3.5}$$

with Ψ_c the **corrected dODF** (cdODF) (in opposition to Ψ). Both definitions were shown to provide interesting information about the angular orientations of the diffusion process.

From the distribution of the protons' diffusion process in the brain, the ultimate goal of researchers is to obtain the fiber geometry, which is at the origin of the anisotropic diffusion process. The **fiber orientation distribution function** (fODF), introduced by [Tournier et al. (2004)], contains the probability in 3D of the fiber bundles' orientations. In opposition to the dODF or to the cdODF, the fODF directly gives the routes of the fibers in the brain which constitute the object of interest in dMRI. It is important to distinguish the dODF, which constitutes the response of the water molecules to the fiber geometry of the brain, and the fODF, which delivers the most probable fiber geometry itself.

The dODF, cdODF and fODF can be reconstructed using the **q space**. The latter is similar to the **k** space used in MRI [Mansfield (1977); Ljunggren (1983)]. Here, the **q** vector is defined as: $\mathbf{q} = \gamma \delta \mathbf{g}$, with γ the proton gyromagnetic ratio and $\mathbf{g}$ the diffusion gradient, as defined previously. The **q** space is the dual space of the propagator space and

its $\mathbf{q}$ vector measures a displacement spatial "frequency". The $\mathbf{q}$ vector is also linked to the well-known diffusion weighting parameter b with the formula: $b = \tau\|\mathbf{q}\|^2$. Considering the diffusion gradients as Dirac delta functions, the authors of [Stejskal and Tanner (1965)] showed that we can link in the $\mathbf{q}$ space the dMRI measured signal S with the proton displacement probability density function $P(\mathbf{r}, \tau)$ (with $\mathbf{r}$ the proton displacement vector and τ the effective diffusion time as defined previously). Indeed, the normalized signal $E(\mathbf{q}, \tau) = S(\mathbf{q}, \tau)/S_0$ is proportional to 3D Fourier transform of the propagator, written as $\mathcal{F}[P(\mathbf{r}, \tau)]$:

$$\boxed{E(\mathbf{q}, \tau) = \int_{\Re^3} P(\mathbf{r}, \tau) \cdot e^{-2\pi i \mathbf{q}^{\mathrm{T}}\mathbf{r}} d\mathbf{r} = \mathcal{F}[P(\mathbf{r}, \tau)]}. \tag{3.6}$$

Most local diffusion models were established from this equation [Callaghan (1991)].

3.2.2 Local diffusion models' overview

The first local diffusion model proposed to make a synthesis of the large set of DW data was the **diffusion tensor imaging** (DTI) introduced by [Basser et al. (1994)]. This model delivers a volume containing in each voxel the probability distribution of the diffusion process using a 3×3 tensor. This model therefore gave an idea of how the routes taken by the protons in the brain were organized, without having recourse to histology. However, DTI has its flaw: it does not always accurately express the most likely directions of the diffusion process, as it assumes in a simplistic way that the water molecules' diffusion follows a Gaussian distribution and therefore allows the representation of a single fiber population. It cannot deal with several populations and efficiently represent complex fiber configurations such as crossings, kissings or splittings. To overcome this limitation, alternative local diffusion models were proposed that enable a higher angular resolution. They are called the **high angular resolution diffusion imaging** (HARDI) models and most of them are summarized in fig. 3.5.

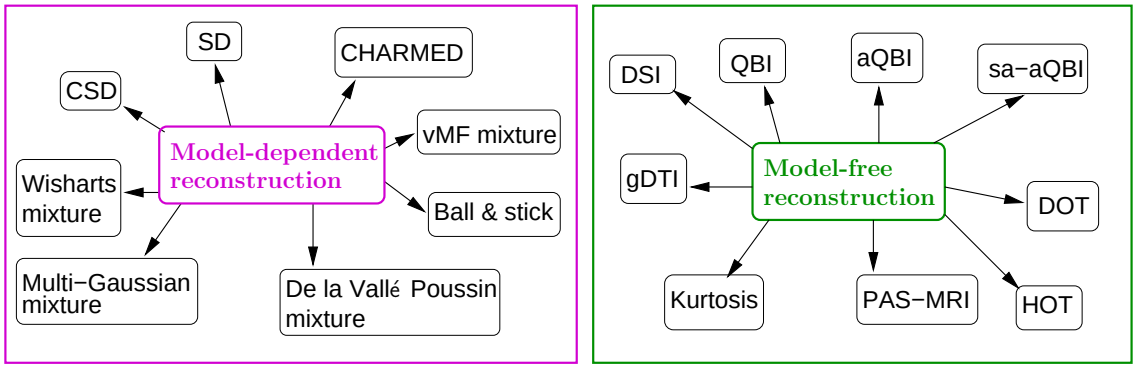


Figure 3.5: Outline of the different HARDI reconstruction methods in dMRI. Fig. inspired from [Descoteaux (2008)] and [Marrakchi-Kacem (2011)].

The aim of these HARDI models developed in the last fifteen years is to propose the

most realistic model of the diffusion process. All of these reconstruction methods rely on eq. 3.6 and are applied in the **q** space. Some of them make an *a priori* assumption on the diffusion propagator and are therefore called model-dependent, whereas other methods do not make any assumption and are therefore called model-free. Concerning the first group of methods, one of its historical models is the **multi-Gaussian mixture** (or multi-DTI) model, which generalizes the DTI model [Tuch (2002)]: it uses a Gaussian mixture to model the diffusion process. The resulting DW signal is then a combination of different signals due to one fiber bundle geometry. This model assumes that there is no water molecules' exchange between the diffusion compartments. In the same spirit, the **ball & stick** model [Behrens et al. (2003); Hosey et al. (2005)] uses two Gaussian distributions to model the diffusion process. The first one corresponds to a highly anisotropic diffusion process (referred as the stick, *e.g.* the fiber bundle) and the other one represents an isotropic diffusion process (referred as the ball). The **composite hindered and restricted model of diffusion** (CHARMED) [Assaf and Basser (2005)] proposes to model the diffusion process as a combination of an hindered diffusion compartment described by a Gaussian distribution and a restricted diffusion compartment modeled by a diffusion process inside a cylinder. For the other model-dependent reconstructions, we invite the reader to read section 4.3 from the thesis [Descoteaux (2008)] and the more recent paper of [Assemlal et al. (2011)].

Concerning the model-free reconstructions, the first historical model is the **diffusion spectrum imaging** (DSI) ([Callaghan et al. (1988); Wedeen et al. (2000, 2005)]), which uses a finite cartesian sampling on a grid restricted to a sphere in the **q** space. This model estimates the diffusion propagator from eq. 3.6 and computes the inverse Fourier transform of $E(\mathbf{q}, \tau)$ at each cartesian sample of the **q** space. While DSI allows to directly access the PDF, it requires a very long acquisition time (1h30 for a sampling including 515 points in the **q** space in case of a whole brain acquisition) which is unrealizable in a clinical routine. When restricting the reconstruction to the angular information corresponding to the distribution of the orientations of the diffusion process, it can be shown that a simple sampling over a sphere in the **q** space is sufficient. Lots of such single shell HARDI models were introduced in the past decade that all aimed at providing reliable angular profile of the diffusion process, while maintaining the number of samples as low as possible. One of the popular HARDI models is the **Q-Ball imaging** (QBI) [Tuch (2002)], which reconstructs the dODF presented previously. We will detail this model later in the chapter. Another popular model is the **persistant angular structure** (PAS-MRI) [Jansons and Alexander (2003)], which determines the radially PAS of the propagator. The peaks of the PAS enable to reconstruct the orientations of the diffusion process.

The outcome of the local diffusion models is inherently linked to the choice of the **q** space sampling. This is summarized in fig. 3.6, which shows the evolution of the sampling from the historical investigation of diffusion with NMR [Stejskal and Tanner (1965)] until the advanced dMRI techniques of today. The first HARDI model was the DSI, which requires many samples to deliver accurate estimations of the diffusion propagator. Then,

the single-shell HARDI models reduced the DW measurements to a sphere in the **q** space, as the information of interest was restricted to the radial projection of the propagator only. Several HARDI models were proposed with improvements over the years on the choice of the basis to represent the DW signal or on the *a priori* information to inject in the model. These improvements in the modeling decreased the amount of required samples and offered the possibility to go beyond the angular information. Therefore, other information than only the angular one was investigated about the diffusion propagator: the radial information contained in the propagator could be approached using some **multiple-shell** reconstruction methods [Özarslan et al. (2009); Assemlal et al. (2009); Kezele et al. (2010); Descoteaux et al. (2010)]. The latter proposed to estimate the propagator either for a same orientation in the **q** space on multiple shells or in a sparse manner on the multiple shells performing **hybrid** diffusion imaging (HYDI) [Alexander et al. (2006)]. Such multiple-shell models indicate like other models the main water molecules' diffusion orientations.

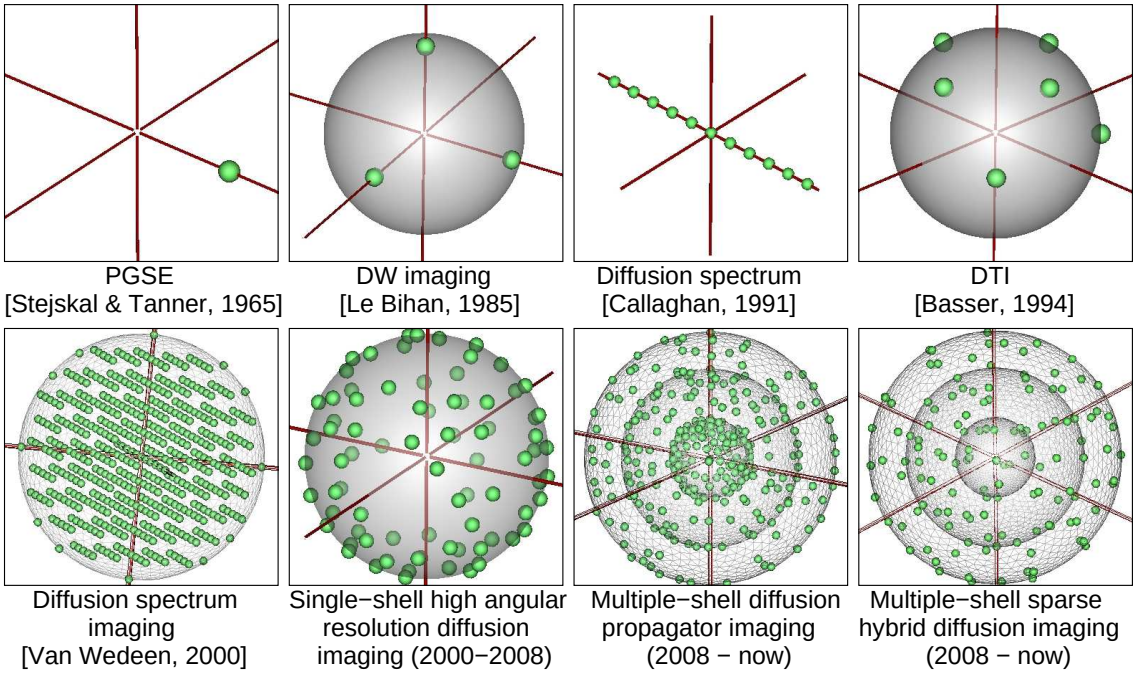


Figure 3.6: The different **q** space samplings from 1965 until now. Fig. extracted from [Poupon (2010)].

We will next limit ourselves to present the historical DTI model and the HARDI models of QBI and spherical deconvolution (SD), which were considered for real-time dMRI as they have all efficient reconstructions and are therefore easily adaptable to reconstructions of the tensor, dODF and fODF performed during the acquisition of the DW data.

3.2.3 The diffusion tensor imaging (DTI) model

The diffusion propagator equation 3.6, seen previously, can be analytically solved if we suppose the propagator to be Gaussian. The propagator is then written such that:

$$P(\mathbf{r},\tau) = \frac{1}{\sqrt{(4\pi\tau)^3|\mathbf{D}|}} \cdot e^{-\frac{1}{4\tau}\mathbf{r}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{r}} \tag{3.7}$$

where $\mathbf{D}$ is a 2^{nd} order tensor characterizing the diffusion process, introduced by Peter Basser *et al.* [Basser et al. (1994)]. When injecting eq. 3.7 in eq. 3.6, the DW signal S is then obtained such that:

$$S(\mathbf{q},\tau) = S_0 e^{-\tau\mathbf{q}^{\mathrm{T}}\mathbf{D}\mathbf{q}}. \tag{3.8}$$

$\mathbf{D}$ is a symmetric positive matrix that can be determined from a T_2 weighted image and from at least six DW images with typically a constant value for the diffusion gradient and six different diffusion orientations $\mathbf{o}$. Practically, because of noise, $\mathbf{D}$ requires more than six orientations. The eigensystem of the diffusion tensor $\mathbf{D}$ provides three eigenvalues λ_1, λ_2 and λ_3 corresponding to the three main directions of diffusivity $\mathbf{e_1}$, $\mathbf{e_2}$ and $\mathbf{e_3}$, such that $\lambda_1 > \lambda_2 > \lambda_3$. The eigenvector $\mathbf{e_1}$ gives the direction along which the water molecules' diffusion is the most important. The diffusion tensor can be represented by an ellipsoid whose three axes are aligned along $\mathbf{e_1}$, $\mathbf{e_2}$ and $\mathbf{e_3}$. In a voxel, if there is only one fiber bundle, the DTI rightly generates an ellipsoid whose main axis is parallel to the main diffusion direction (see a and b in fig. 3.7). However, in the case of a fiber bundles' crossing, the DTI results in what tends to be an oblate spheroid (*i.e.* $\lambda_1 = \lambda_2 >> \lambda_3$) or a sphere and thus is not able to bring to light the main diffusion directions (see c and d in fig. 3.7). Consequently, despite it is still popular in clinical applications, the DTI model is problematic because it cannot describe heterogeneous populations inside a voxel, and yet at a resolution of $8\mathrm{mm}^3$, there is between a third and two-thirds of the WM voxels that contain fiber bundle crossings [Descoteaux (2008)].

From the tensor, scalar indexes, which are rotationally invariant, can be calculated. The two most popular are the ADC ([Basser et al. (1994)]) and the **fractional anisotropy** (FA) ([Basser (1995)]), which can be computed from the eigensystem:

$$\mathrm{ADC} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}, \tag{3.9}$$

$$\mathrm{FA} = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \overline{\lambda})^2 + (\lambda_2 - \overline{\lambda})^2 + (\lambda_3 - \overline{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{3.10}$$

with $\overline{\lambda} = trace/3$, $trace$ being the trace of $\mathbf{D}$. An isotropic tissue has an FA near 0, whereas a very anisotropic tissue with only one main diffusion direction has an FA near 1. A color-encoded RGB (red-green-blue) map of the FA can also be reconstructed [Pierpaoli (1997)]. Fig. 3.8 shows an ADC, an FA and an RGB map. Important other scalar indexes like the **longitudinal diffusivity** $\lambda_{||} = \lambda_1$ and the **transverse diffusivity** $\lambda_{\perp} = (\lambda_2 + \lambda_3)/2$ can also be deduced to study the maturation of white matter fibers.

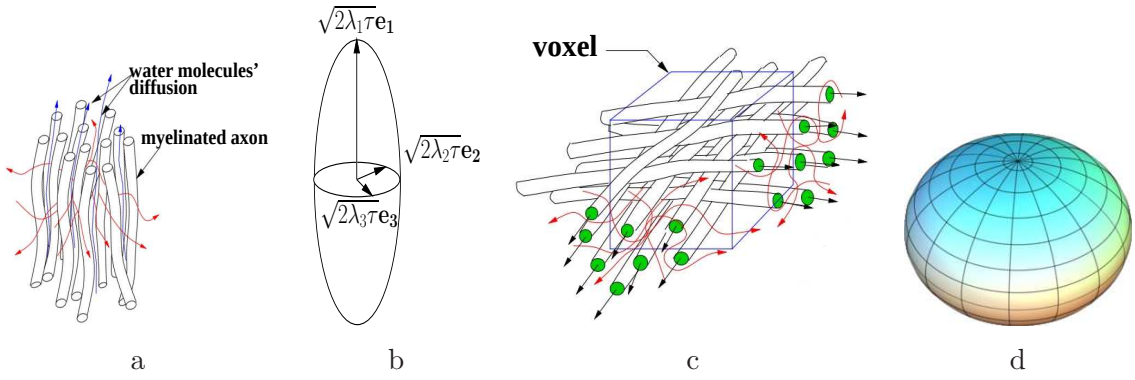


Figure 3.7: For the fiber bundle configurations in (a) and (c), the corresponding diffusion tensors are represented by an ellipsoid (b) and by an oblate spheroid (d), respectively. On the ellipsoid surface, the diffusion probability of the water molecules is the same. This is also true for the spheroid surface. The DTI correctly retrieves the main diffusion direction in the case of one fiber bundle in a voxel (b), but in the case of a bundle crossing, no one of the two main directions is retrieved (d). Fig. (a) and (c) are extracted from [Poupon (1999a)]. Fig. (d) is extracted from Wikipedia.

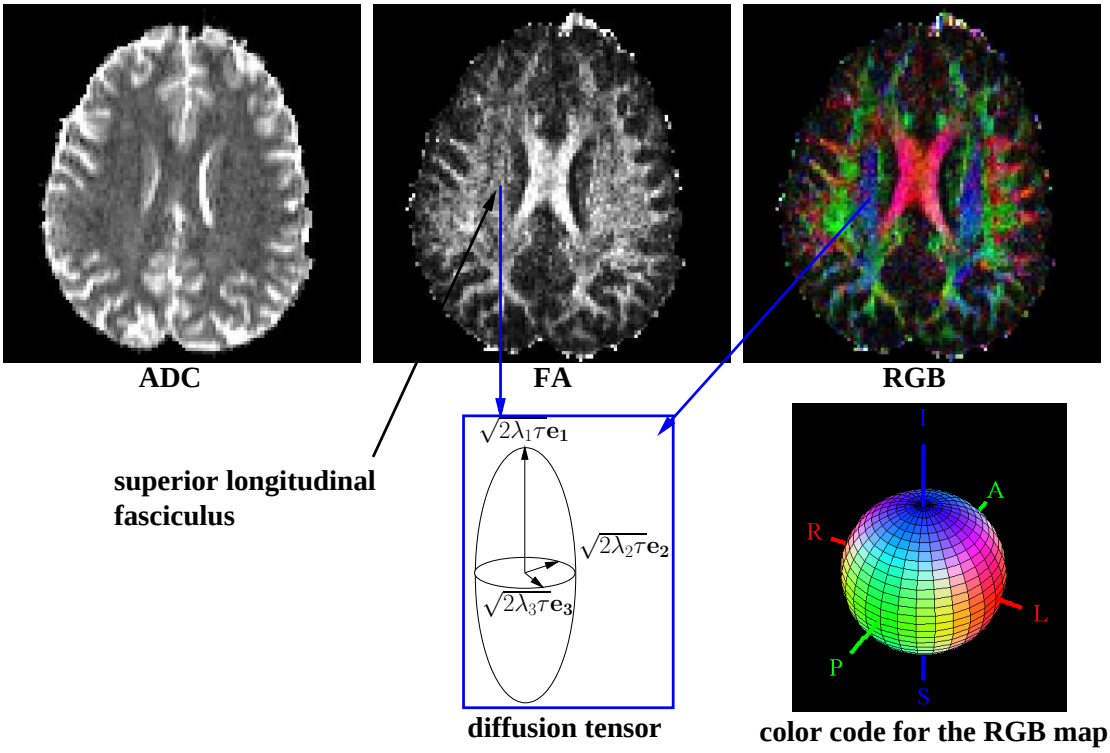


Figure 3.8: Popular maps of the DTI model.

As we have seen, the DTI model presents some limitations to describe the true diffusion phenomenon, but these limitations can be overcome by the use of HARDI models. The latter also stem from the diffusion propagator equation 3.6, but they do not make any strong

hypothesis on the propagator like the DTI model. As seen previously, single-shell HARDI models were developed to estimate with high angular resolution the radial projection of the diffusion propagator that contains the main diffusion orientations corresponding to the underlying main fiber directions. Such models rely on the decomposition of the DW signal over a single shell of the $\mathbf{q}$ space using adequate basis functions. The next subsection explains the decomposition on a modified spherical harmonics basis, which is used in the analytical QBI model.

3.2.4 Signal decomposition on a modified spherical harmonics basis

Fig. 3.9 shows an exemple of a sphere in the $\mathbf{q}$ space on which high angular single shell dMRI acquisitions are done. Every complex signal on the sphere can be decomposed on an orthogonal spherical harmonics (SH) basis adapted to the sphere. The modified SH are the equivalent of the harmonical decomposition in Fourier series. Here, as in signal processing, a decomposition of the DW signal S on a basis can be very useful to make a synthesis of the diffusion information. And yet, the signal S is real and symmetric (because in eq. 3.6 $P(\mathbf{r}, \tau)$ is real and symmetric). The works by [Frank (2002); Hess et al. (2006); Descoteaux et al. (2007)] consequently define a **modified SH basis**, which is real and symmetric, to decompose the signal S. For that, they define the index j from the degree $l = 0, 2, 4, ..., L$ and the order $m = -l, ..., 0, ..., l$ such that: $j(l, m) = (l^2 + l + 2)/2 + m$. The modified SH basis can then be written in the spherical coordinate system, where θ and ϕ represent the colatitude and the longitude, respectively, of the unitary vector $\mathbf{o}_i$ defining the orientation along which the diffusion is measured in the $\mathbf{q}$ space:

$$Y_j(\theta, \phi) = \begin{cases} \sqrt{2}\mathbf{Re}(Y_l^{|m|}(\theta, \phi)), & \text{if } m \leq 0, \\ Y_l^m(\theta, \phi), & \text{if } m = 0, \\ (-1)^{m+1}\sqrt{2}\mathbf{Im}(Y_l^m(\theta, \phi)), & \text{if } m \geq 0, \end{cases} \tag{3.11}$$

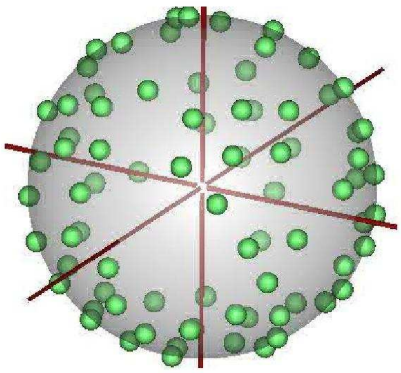


Figure 3.9: Sphere in the $\mathbf{q}$ space sampled in an homogeneous way according to the single shell HARDI method. The green points represent the samples. (extract from [Poupon (2010)])

where $\mathbf{Re}(Y_l^m(\theta,\phi))$ and $\mathbf{Im}(Y_l^m(\theta,\phi))$ represent the real and imaginary parts, respectively, of the SH $Y_l^m(\theta,\phi)$, which is defined by: $Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(cos(\theta))e^{im\phi}$, with P_l^m the associated Legendre function of the first kind with degree l and phase m. As the indexes l and m are defined for an index j designating the j^{th} SH Y_j, they are from now on written $l(j)$ and $m(j)$. The normalized signal $E = S/S_0$ can be decomposed on this new basis:

$$E(\theta_i,\phi_i) = \sum_{j=1}^{N} C_j^{\text{DW}} Y_j(\theta_i,\phi_i), \tag{3.12}$$

with $E(\theta_i,\phi_i)$ the DW signal normalized by S_0 measured for the diffusion orientation $\mathbf{o}_i$, the latter defined by (θ_i,ϕ_i), and with C_j^{DW} the j^{th} coefficient of the DW normalized signal on the SH $Y_j(\theta_i,\phi_i)$. The index i indicates that the measurement was done along the i^{th} diffusion orientation among the entire set of K directions used during the sequence. The parameter $N = \frac{(L+1)(L+2)}{2}$ corresponds to the number of modified SH, L being the maximum order used in the modified SH basis. We call $\mathbf{B}$ the basis matrix:

$$\mathbf{B} = \begin{pmatrix} Y_1(\theta_1,\phi_1) & Y_2(\theta_1,\phi_1) & \dots & Y_N(\theta_1,\phi_1) \\ \vdots & \vdots & \ddots & \vdots \\ Y_1(\theta_K,\phi_K) & Y_2(\theta_K,\phi_K) & \dots & Y_N(\theta_K,\phi_K) \end{pmatrix}. \tag{3.13}$$

Eq. 3.12 can then be written in a matrix form:

$$\mathbf{E} = \mathbf{B}\mathbf{C}^{\text{DW}}, \tag{3.14}$$

where $\mathbf{E} = [S(\theta_1,\phi_1)/S_0, \dots, S(\theta_i,\phi_i)/S_0, \dots, S(\theta_K,\phi_K)/S_0]^{\text{T}}$ is the DW normalized signal vector and $\mathbf{C}^{\text{DW}} = [C_1^{\text{DW}}, \dots, C_j^{\text{DW}}, \dots, C_N^{\text{DW}}]^{\text{T}}$ is the **coefficients' vector of the decomposition of E on the modified SH basis**.

This modified SH basis notion is essential for the following subsections because the analytical QBI models rely on it.

3.2.5 Q-Ball imaging (QBI)

In this subsection, we present the three QBI models, which appeared in this order: first the **numerical QBI**, then the **analytical QBI** (aQBI) and finally the **analytical QBI with the solid angle** (sa-aQBI).

The numerical QBI

The QBI method [Tuch (2004)] relies on the previously explained spherical sampling with a spherical radius fixed to a value q. The propagator is therefore reconstructed only from a sphere of the $\mathbf{q}$ space. Indeed, to estimate the fiber bundle orientations, only the angular information contained in the diffusion propagator is necessary. The advantage of this technique is that the DW data acquisition on a sphere is far more rapid than on the entire $\mathbf{q}$ space. Let us remind that the $\mathbf{q}$ vector is proportional to the diffusion weighting

parameter b: $b = \tau \cdot |\mathbf{q}|^2$. Concretely, a single shell acquisition dedicated to QBI requires the use of a quite high *b-value* (typically $b \geq 3000 s \cdot mm^{-2}$), and the use of a set of typically 60 to 200 diffusion directions $\mathbf{o}_i$, yielding an acquisition time of 15 to 60 min for a whole brain scan.

The QBI model is therefore only focused on the fiber bundle orientations and it is possible to estimate the previously introduced dODF of the water molecules written as Ψ. For that, the numerical QBI, as well as the aQBI, approximate it by the Funk-Radon transform (FRT) of the normalized DW signal E [Tuch (2004)]. More details on this approximation are given in appendix A (section 3.5). To obtain the FRT of the signal E, the value of E has to be obtained for each direction $\mathbf{o}_i$, and then, in the numerical QBI method, a numerical interpolation is done to complete the values of E on the sphere of the $\mathbf{q}$ space. This approximation of the function Ψ at the sampled and interpolated sphere points permits to reconstruct the fiber bundle directions, especially when the bundles are crossing, whereas the DTI is not able to do it.

The numerical QBI model utilizes two angular information smoothing steps, one due to the FRT, the other due to the signal interpolation along the equators, which are perpendicular to the directions $\mathbf{o}_i$. This double smoothing damages the angular resolution. Another method, the aQBI, which relies on a decomposition of the DW signal in the SH space, permits to avoid the second smoothing step. Indeed it compresses the measured DW information on a relatively small number of harmonic coefficients (typically 15 to 45 coefficients depending on the chosen harmonic order), yielding an increase of the reconstruction method robustness.

The analytical QBI (aQBI)

While the numerical QBI proposes a numerical interpolation of the signal E, the aQBI uses an analytical equation of E relying on the decomposition of E on the modified SH basis that we detailed in subsection 3.2.4. Frank and Descoteaux showed that it is possible to reconstruct the function Ψ in a fast and robust way using this analytical signal decomposition [Frank (2002); Descoteaux et al. (2007)]. Let us remind the decomposition equation of the DW normalized signal E on the modified SH basis:

$$\mathbf{E} = \mathbf{B} \cdot \mathbf{C}^{\mathrm{DW}}. \tag{3.15}$$

In a similar way, the vector $\mathbf{C}^{\mathrm{ODF}}$ of the **coefficients of the decomposition of Ψ on the modified SH basis** is written as:

$$\begin{bmatrix} \Psi(\theta_1, \phi_1) \\ \vdots \\ \Psi(\theta_i, \phi_i) \\ \vdots \\ \Psi(\theta_K, \phi_K) \end{bmatrix} = \mathbf{B} \cdot \mathbf{C}^{\mathrm{ODF}}. \tag{3.16}$$

In the numerical QBI, the FRT of the signal E has to be calculated to approximate the function Ψ. In the aQBI, the authors of [Descoteaux et al. (2007)] show the link between the decomposition of the function Ψ and the one of the signal E:

$$\mathbf{C}^{\mathrm{ODF}} = \mathbf{P}\mathbf{C}^{\mathrm{DW}}, \tag{3.17}$$

with $\mathbf{P}$ the N×N Funk-Hecke matrix, whose diagonal elements are equal to $2\pi P_{l(j)}(0)\ \forall j \in [\![1; N]\!]$, where the functions $P_{l(j)}$ correspond to the Legendre polynomials of degree $l(j)$ evaluated at 0.

To reconstruct the function Ψ, the coefficients $\mathbf{C}^{\mathrm{DW}}$ have to be determined to give access to the coefficients $\mathbf{C}^{\mathrm{ODF}}$ with eq. 3.17. The vector $\mathbf{C}^{\mathrm{DW}}$ can be estimated using the **least squares** method. Indeed, if we account for noise in the dMRI measurement, eq. 3.15 becomes: $\mathbf{M_E} = \mathbf{B}\mathbf{C}^{\mathrm{DW}} + \boldsymbol{\epsilon}$, where $\mathbf{M_E}$ represents the ratio $1/S_0 \cdot \mathbf{M}$, with $\mathbf{M}$ the measured noisy DW signal —S_0 being the noise-free T_2-weighted signal— and $\boldsymbol{\epsilon}$ the vector of the error due to the noise. Then, the term $\left(\mathbf{M_E} - \mathbf{B}\mathbf{C}^{\mathrm{DW}}\right)^{\mathrm{T}}\left(\mathbf{M_E} - \mathbf{B}\mathbf{C}^{\mathrm{DW}}\right)$ has to be minimized. The least squares solution yields the estimate $\hat{\mathbf{C}}^{\mathrm{DW}}$ of the noise-free vector $\mathbf{C}^{\mathrm{DW}}$ such that:

$$\hat{\mathbf{C}}^{\mathrm{DW}} = (\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B})^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot \mathbf{M_E} = \mathbf{B}^{\dagger} \cdot \mathbf{M_E}, \tag{3.18}$$

where $(\cdot)^{\dagger}$ corresponds to the Moore-Penrose pseudo-inverse operator. To improve the estimation of the coefficients $\mathbf{C}^{\mathrm{DW}}$, Maxime Descoteaux proposed to use a **Tikhonov regularization** in the least squares estimation. Now, the coefficients $\hat{\mathbf{C}}^{\mathrm{DW}}$ have to minimize $\left(\mathbf{M_E} - \mathbf{B}\mathbf{C}^{\mathrm{DW}}\right)^{\mathrm{T}}\left(\mathbf{M_E} - \mathbf{B}\mathbf{C}^{\mathrm{DW}}\right) + \lambda \mathbf{C}^{\mathrm{DW}\mathrm{T}}\mathbf{L}\mathbf{C}^{\mathrm{DW}}$, with λ the regularization factor and $\mathbf{L}$ the N×N Laplace-Beltrami smoothing matrix, which is diagonal with diagonal elements equal to $l(j)^2\,(l(j)+1)^2\,\forall j \in [\![1; N]\!]$. The solution is then given by:

$$\hat{\mathbf{C}}^{\mathrm{DW}} = (\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B} + \lambda\mathbf{L})^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot \mathbf{M_E}, \tag{3.19}$$

where $\lambda\mathbf{L}$ constitutes the regularization term. The latter permits to smooth the solution and thus to diminish the number of negative peaks, which appear in the DW signal recomposed from the coefficients $\mathbf{C}^{\mathrm{DW}}$. There are even more negative peaks if the maximum SH order L used in the modified SH basis is high. Consequently, the regularization enables to take advantage of higher SH terms, which offer more details on the signal, while suppressing the terms creating the negative peaks. In [Descoteaux (2008)], the value $\lambda = 0.006$ is specified to be an optimal value for separating configurations with one fiber from configurations with two fibers on a large range of signal-to-noise ratios (SNR) and of *b-values*. If λ is fixed to 0, eq. 3.19 simplifies to eq. 3.18. Then, when having obtained the coefficients $\hat{\mathbf{C}}^{\mathrm{ODF}}$ using eq. 3.17 with $\hat{\mathbf{C}}^{\mathrm{DW}}$, the dODF can be reconstructed such that:

$$\hat{\boldsymbol{\Psi}} = \mathbf{B}_{\mathrm{out}} \cdot \hat{\mathbf{C}}^{\mathrm{ODF}}, \tag{3.20}$$

with $\hat{\boldsymbol{\Psi}} = [\hat{\Psi}(\theta_1, \phi_1), \ldots, \hat{\Psi}(\theta_i, \phi_i), \ldots, \hat{\Psi}(\theta_{K_{\mathrm{out}}}, \phi_{K_{\mathrm{out}}})]^{\mathrm{T}}$ and $\mathbf{B}_{\mathrm{out}}$ a $\mathrm{K_{out}}$×N matrix being the matrix of the modified SH basis as the K×N matrix $\mathbf{B}$, unless it contains $\mathrm{K_{out}}$ rows corresponding to $\mathrm{K_{out}}$ output orientations, with usually $\mathrm{K_{out}} \geq$K.

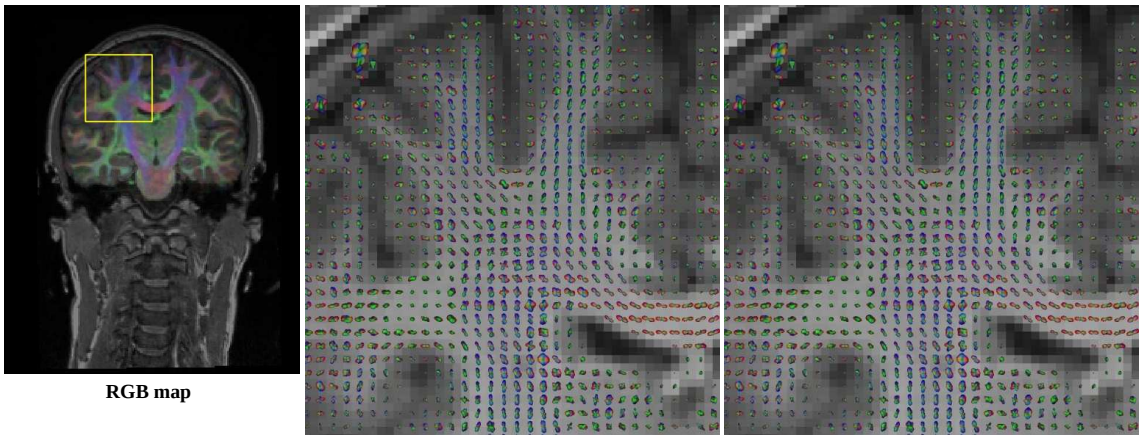


Figure 3.10: Single shell HARDI results at $b = 3000s \cdot mm^{-2}$ with the numerical QBI and the aQBI. On the left, the RGB map depicts the main diffusion directions obtained with DTI, *i.e.* the maxima of the eigenvalues of the tensor at each voxel (red along the x-axis, green along the y-axis, blue along the z-axis). A red ROI is shown inside. In the ROI, the functions Ψ are estimated with the numerical QBI (middle) and the aQBI (right). The aQBI gives an angular precision on the diffusion orientations which is equivalent to the precision given by the numerical QBI. Moreover, the aQBI lightens the computation time and is free from any error due to a numerical interpolation.

Concretely, this analytical estimation of the function Ψ is as precise as the one obtained using the numerical QBI reconstruction, as shown in fig. 3.10. The main advantage of the aQBI is to reduce the computation time thanks to the DW signal representation using only a few set of coefficients and to be free from any error due to a numerical interpolation along the equator. The main mathematical relations, which are essential for the rest of the thesis, are summed up in table 3.1.

Table 3.1: Summary of the notations and main relations of the aQBI.

Analytical decomposition of the normalized signal E on the modified SH basis:

• $\mathbf{S} = [S(\theta_1, \phi_1), \ldots, S(\theta_i, \phi_i), \ldots, S(\theta_K, \phi_K)]^{\mathrm{T}}$ the K×1 vector of the DW noise-free signals.
$\mathbf{M} = [M(\theta_1, \phi_1), \ldots, M(\theta_i, \phi_i), \ldots, M(\theta_K, \phi_K)]^{\mathrm{T}}$ the K×1 vector of the measured (and therefore noisy) DW signals.
$\mathbf{M_E} = [M(\theta_1, \phi_1)/S_0, \ldots, M(\theta_i, \phi_i)/S_0, \ldots, M(\theta_K, \phi_K)/S_0]^{\mathrm{T}}$ the K×1 vector of the measured DW signals normalized by the noise-free T_2-weighted signal S_0.

• $\mathbf{C}^{\mathrm{DW}} = [C_1^{\mathrm{DW}}, \ldots, C_j^{\mathrm{DW}}, \ldots, C_N^{\mathrm{DW}}]^{\mathrm{T}}$ the N×1 coefficients' vector of the decomposition of the vector $\mathbf{E}$ on the modified SH basis, the latter defined as:

$$Y_j(\theta, \phi) = \begin{cases} \sqrt{2}\mathbf{Re}(Y_{l(j)}^{|m(j)|}(\theta, \phi)), & \text{if } m \leq 0 \\ Y_{l(j)}^{m(j)}(\theta, \phi), & \text{if } m = 0 \\ (-1)^{m(j)+1}\sqrt{2}\mathbf{Im}(Y_{l(j)}^{m(j)}(\theta, \phi)), & \text{if } m \geq 0, \end{cases}$$

• $\boxed{\mathbf{E} = \mathbf{B}\mathbf{C}^{\mathrm{DW}}}$, with

$$\mathbf{B} = \begin{pmatrix} Y_1(\theta_1, \phi_1) & Y_2(\theta_1, \phi_1) & \ldots & Y_N(\theta_1, \phi_1) \\ \vdots & \vdots & \ddots & \vdots \\ Y_1(\theta_K, \phi_K) & Y_2(\theta_K, \phi_K) & \ldots & Y_N(\theta_K, \phi_K) \end{pmatrix}$$ the matrix of the modified SH for which $N = \frac{(L+1)(L+2)}{2}$ is the number of SH used, L being the maximum SH order.

Obtaining the coefficients of the dODF Ψ on the modified SH basis:

• The estimation of $\mathbf{C}^{\mathrm{DW}}$ from $\mathbf{M_E}$ is written as: $\hat{\mathbf{C}}^{\mathrm{DW}} = (\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B} + \lambda \mathbf{L})^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot \mathbf{M_E}$, with λ the regularization factor and $\mathbf{L} = \begin{pmatrix} \ddots & & \\ & l(j)^2(l(j)+1)^2 & \\ & & \ddots \end{pmatrix}$ the N×N Laplace-Beltrami smoothing matrix.

• Finally, the estimation of the coefficients' vector of the dODF Ψ on the modified SH basis is such that: $\boxed{\hat{\mathbf{C}}^{\mathrm{ODF}} = \mathbf{P} \cdot \hat{\mathbf{C}}^{\mathrm{DW}} = \mathbf{P} \cdot (\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B} + \lambda \mathbf{L})^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot \mathbf{M_E}}$, with

$$\mathbf{P} = \begin{pmatrix} \ddots & & \\ & 2\pi P_{l(j)}(0) & \\ & & \ddots \end{pmatrix}$$ the N×N Funk-Hecke matrix.

• And the dODF Ψ is then determined by: $\hat{\mathbf{\Psi}} = \mathbf{B}_{\mathrm{out}} \cdot \hat{\mathbf{C}}^{\mathrm{ODF}}$, with
$\hat{\mathbf{\Psi}} = [\hat{\Psi}(\theta_1, \phi_1), \ldots, \hat{\Psi}(\theta_i, \phi_i), \ldots, \hat{\Psi}(\theta_{K_{\mathrm{out}}}, \phi_{K_{\mathrm{out}}})]^{\mathrm{T}}$.

David Tuch defined, similarly to the FA, the concept of **generalized fractional anisotropy** (GFA) ([Tuch (2004)]), which extends to the aQBI model:

$$\text{GFA} = \sqrt{\frac{\sum_{i=1}^{K} K\left(\Psi(\theta_i,\phi_i) - \langle\Psi\rangle\right)^2}{(K-1)\sum_{i=1}^{K}\Psi(\theta_i,\phi_i)^2}} = \sqrt{1 - \frac{{C_0^{\text{ODF}}}^2}{\sum_{j=1}^{N}{C_j^{\text{ODF}}}^2}} \tag{3.21}$$

with K the number of orientations along which the dODF Ψ is reconstructed. This GFA has the advantage to present high values in voxels containing fiber crossings, whereas the FA, which relies on the DTI, presents low values in such voxels, because of the DTI inability to describe more than one fiber bundle population in a voxel.

The analytical QBI with the solid angle (sa-aQBI)

The sa-aQBI aims at determining the cdODF Ψ_c, which we rewrite the expression through the cone of the infinitesimal solid angle $d\Omega$:

$$\Psi_c(\theta_i,\phi_i) = \int_0^\infty P(r,\theta_i,\phi_i,\tau)r^2 dr. \tag{3.22}$$

The authors of [Aganj et al. (2010)] showed that, after mathematical developments from eq. 3.6, Ψ_c is written in function of the normalized noise-free signal $E(\mathbf{o}_i) = S(\mathbf{o}_i)/S_0$ —$\mathbf{o}_i$ being the orientation defined by (θ_i,ϕ_i)— such that:

$$\Psi_c(\mathbf{o}_i) = \frac{1}{4\pi} + \frac{1}{16\pi^2}\mathcal{G}\{\nabla_b^2 ln\left(-lnE(\mathbf{o}_i)\right)\}, \tag{3.23}$$

with $\mathcal{G}$ designating the FRT and ∇_b^2 designating the Laplace-Beltrami operator. As for the aQBI, the estimate of the coefficients' vector of the decomposition of the measured vector $\mathbf{M_E}$ on the modified SH basis is written as: $\hat{\mathbf{C}}^{\text{DW}} = (\mathbf{B}^{\text{T}}\cdot\mathbf{B} + \lambda\mathbf{L})^{-1}\cdot\mathbf{B}^{\text{T}}\cdot\mathbf{M_E}$. However, the estimate of the coefficients' vector of the decomposition of the cdODF on the modified SH basis written as $\hat{\mathbf{C}}^{\text{ODF}_c}$ is different from the aQBI vector $\hat{\mathbf{C}}^{\text{ODF}}$:

$$\begin{cases} \hat{\mathbf{C}}^{\text{ODF}_c}(j) = \frac{1}{2\sqrt{\pi}} & \text{if } j = 1, \\ \hat{\mathbf{C}}^{\text{ODF}_c}(j) = \frac{1}{16\pi^2}\mathbf{P}\cdot\mathbf{N}\cdot(\mathbf{B}^{\text{T}}\cdot\mathbf{B} + \lambda\mathbf{L})^{-1}\cdot\mathbf{B}^{\text{T}}\cdot ln\left((-ln\mathbf{M_E})\right) & \text{if } j > 1, \end{cases} \tag{3.24}$$

with $\mathbf{N}$ the N×N diagonal matrix with diagonal elements equal to $-l(j)\left(l(j)+1\right)\forall j \in [\![1;N]\!]$ and $\mathbf{M_E}$ the measured (therefore noisy) normalized signal defined such that $\mathbf{M_E} = 1/S_0\cdot\mathbf{M}$, with S_0 the noise-free T_2 weighted signal as before. This technique relying on eq. 3.24 requires to calculate $ln\left((-ln\mathbf{M_E})\right)$. And yet, the latter term is unstable for M_E close to 0 or 1, values which amplify eventual errors present in M_E. To avoid such instabilities, [Aganj et al. (2010)] proposed to keep the values of M_E away from the unstable regions [0,δ_1] and [1-δ_2,1], with $\delta_1 = 10^{-3}$ and $\delta_2 = 10^{-3}$. They defined the function $f(M_E)$, such

that:

$$f(M_E) = \begin{cases} \delta_1/2, & \text{if } M_E < 0 \\ \delta_1/2 + M_E^2/(2\delta_1), & \text{if } 0 \leq M_E < \delta_1 \\ M_E, & \text{if } \delta_1 \leq M_E < 1 - \delta_2, \\ 1 - \delta_2/2 - (1 - M_E)^2/(2\delta_2), & \text{if } 1 - \delta_2 \leq M_E < 1 \\ 1 - \delta_2/2, & \text{if } 1 \leq M_E. \end{cases} \tag{3.25}$$

Then, they replaced M_E by $f(M_E)$ in eq. 3.24 for the computation of the coefficients $\hat{\mathbf{C}}^{\mathrm{ODF}_c}$. Once these are calculated, it is straightforward to obtain the estimate of the cdODF similarly to the aQBI model: $\hat{\mathbf{\Psi}}_\mathbf{c} = \mathbf{B}_{\mathrm{out}} \cdot \hat{\mathbf{C}}^{\mathrm{ODF}_c}$, with $\hat{\mathbf{\Psi}}_\mathbf{c} = [\hat{\Psi}_c(\theta_1, \phi_1), \ldots, \hat{\Psi}_c(\theta_i, \phi_i), \ldots, \hat{\Psi}_c(\theta_{K_{\mathrm{out}}}, \phi_{K_{\mathrm{out}}})]^{\mathrm{T}}$.

In the aQBI model, a consequence of the r^2 term omission in the radial propagator projection is that the reconstructed dODF requires an artificial normalization, called min-max normalization [Tuch (2004)], which leads to the normalized dODF written as Ψ_{norm}: $\Psi_{norm}(\mathbf{o}_i) = \frac{\Psi(\mathbf{o}_i) - \min\Psi(\mathbf{o}_i)}{\max\Psi(\mathbf{o}_i) - \min\Psi(\mathbf{o}_i)}$, $\min\Psi(\mathbf{o}_i)$ and $\max\Psi(\mathbf{o}_i)$ designating the minimum and maximum values, respectively, of the dODF Ψ. On the contrary, the sa-aQBI does not require any normalization. However, in our implementation of the sa-aQBI, we usually use a max normalization to improve the visualization of cdODF and avoid the presence of very small to very large cdODF: $\Psi_{c_{norm}}(\mathbf{o}_i) = \frac{\Psi_c(\mathbf{o}_i)}{\max\Psi_c(\mathbf{o}_i)}$. The question is now to know if the sa-aQBI yields better results than the aQBI. The sa-aQBI detects simulated fiber crossings, whose angles are too small to be detected by the aQBI [Aganj et al. (2010)]. On simulated and real data, the sa-aQBI produces sharper and more precise dODFc than the aQBI results in anisotropic regions, while generating more spherical cdODF than the aQBI results in isotropic regions. However, the r^2 term in the cdODF increases the noise impact on the cdODF in comparison to the ODF of the aQBI model. This yields quite

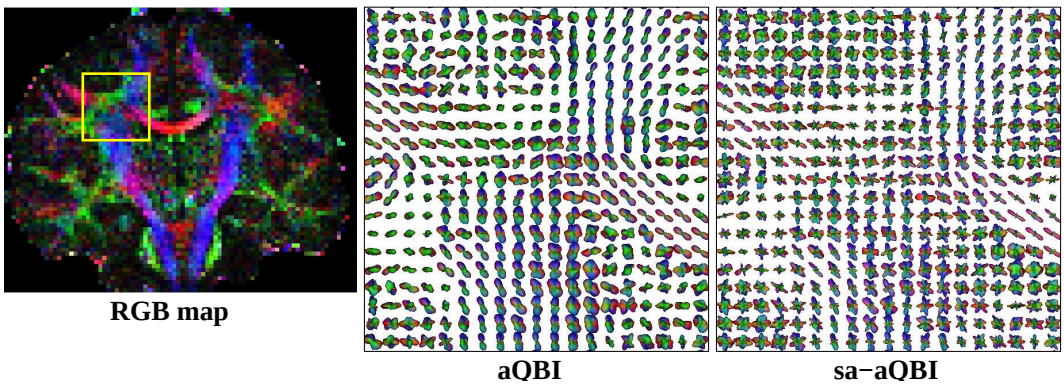


Figure 3.11: Single shell HARDI results at $b = 3000s \cdot mm^{-2}$ with aQBI and sa-aQBI with the maximum SH order L set to 8. On the left, the ROI is shown in yellow in the RGB map. In the ROI, the functions Ψ and Ψ_c are estimated with the aQBI (middle) and the sa-aQBI (right), respectively.

noisier cdODF than the aQBI dODF. Fig. 3.11 illustrates these both positive and negative effects.

Another method enabling to produce sharper results than the aQBI is the SD that we show hereinafter.

3.2.6 Using spherical deconvolution (SD) to get the fODF

The dODF and cdODF are each a measurement of the probability distribution of the water molecules' main diffusion orientations. And yet, we are not fundamentally interested by this diffusion process characteristic, but we use it above all to indirectly explore the underlying structural tissue organization, making the hypothesis that a link exists between the fiber distribution, which we are interested in, and the diffusion process anisotropy due to the fiber configuration. While assuming an impulse response of a homogeneous fiber bundle to the diffusion process, the reconstruction of the fODF, function introduced in subsection 3.2.1, was proposed using spherical deconvolution (SD) of the DW signal [Anderson and Ding (2002); Tournier et al. (2004); Jian and Vemuri (2007)]. A constrained super-resolved SD was also proposed [Tournier et al. (2007)]. In [Descoteaux (2008)], the SD of the dODF by the dODF for a single fiber (impulse dODF response) was proposed. This model considers that the dODF is the result of the convolution of the fODF by the response of a simple bundle to the diffusion process, as illustrated in fig. 3.12. Thus, the SD of the dODF directly yields the fODF (indicated at the right bottom in fig. 3.12), which more precisely describes the bundle orientations. From a technical point of view, the choice of a modified SH basis leads to a very simple convolution or deconvolution operation in the harmonic coefficients' space:

$$\Psi_f(\theta_i, \phi_i) = \sum_{j=1}^{N} \frac{C^{\mathrm{ODF}}(j)}{R(j)} Y_j(\theta_i, \phi_i), \tag{3.26}$$

with $R(j) = 2\pi \int_{-1}^{+1} P_{l(j)}(t) R'(t) dt$, where R' is the convolution kernel corresponding to the impulse response of an homogeneous fiber bundle to the diffusion process. R' is supposed to be a tensor with eigenvalues $[\lambda_1, \lambda_2, \lambda_2]$, such that $\lambda_1 >> \lambda_2$; the tensor is then represented by an elongated spheroid, called prolate spheroid, and is expressed as: $R'(t) = \frac{1}{8\pi b\sqrt{\lambda_2^2 \lambda_1}} \frac{1}{\sqrt{(\lambda_2/\lambda_1 - 1)t^2 + 1}}$.

It has to be kept in mind that this indirect method to determine the WM fiber bundles' structure is not totally reliable because it does not account for neither the membranes' permeability, with or without any membrane's channel, nor the myelin, which accelerates the information transmission in axons. That is why other local models combinated with new dMRI sequences continue to be developed in the aim to create a distant microscope to deeply explore the cerebral microstructure.

The fODF presents a better angular resolution than the dODF for the fiber tractography, a technique permitting to virtually reconstruct the brain fiber bundles. Indeed, tractography generates virtual fibers approximating the real brain fiber bundle configura-

tion. We explain in the next section the streamline deterministic tractography tool, which is one of the tractography algorithms the local models lead to.

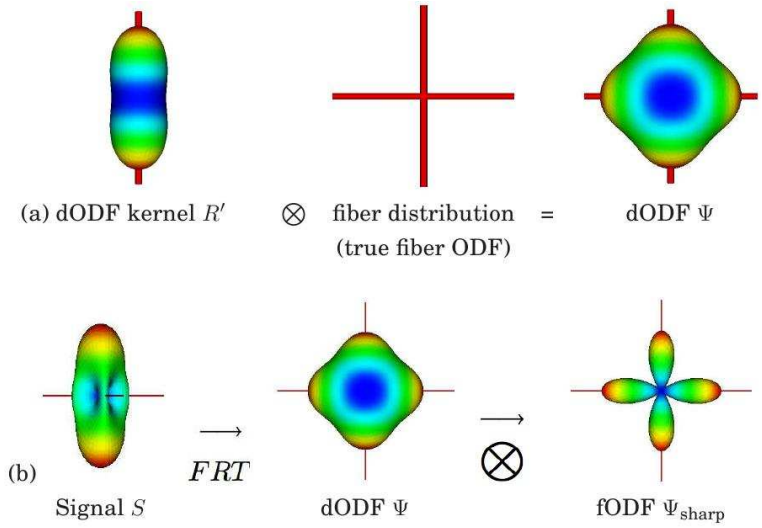


Figure 3.12: (a) The convolution product between the dODF corresponding to a simple fiber bundle ("dODF kernel R'" in the fig.) and the unknown fiber orientation distribution ("fiber distribution" in the fig.) gives the dODF corresponding to this unknown fiber configuration ("dODF" in the fig.). (b) After having obtained the dODF using the FRT of the signal E, the fODF, written as Ψ_{sharp} in the fig., can be determined using a SD. (extract from [Descoteaux (2008)]).

3.3 Tractography techniques

Tractography aims at reconstructing *in vivo* and *non invasively* the anatomical brain connectivity. To this aim, it builds pathways from the diffusion propagator field assumed to correspond to the underlying WM anatomy. Pathways reconstructed with tractography do not approximate individual axons, but an optimal pathway in the sense of the diffusion process. This numerical reconstruction relies on the local models' results, for example the dODF in the aQBI model or the fODF in the SD model. From these results, the tractography algorithm follows the main diffusion or fiber directions and creates therefore routes of fibers that connect the different brain areas. It aims at reconstructing the tracts for the whole brain as presented in fig. 3.15.

The first class of tractography algorithms to appear in the literature was the class of **streamline deterministic tractography**. This streamline method proposes to simply propagate a line following the voxel to voxel fiber directions given by the local diffusion model. The streamline deterministic tractography was developed in different research teams. This method does not depend on any specific local diffusion model and works from any ODF model. The authors of [Mori et al. (1999); Conturo et al. (1999); Poupon (1999a); Basser et al. (2000); Westin et al. (2002); Lazar and Alexander (2003)] among others applied the algorithm from DTI fields, whereas [Tuch et al. (2002); Perrin et al. (2005);

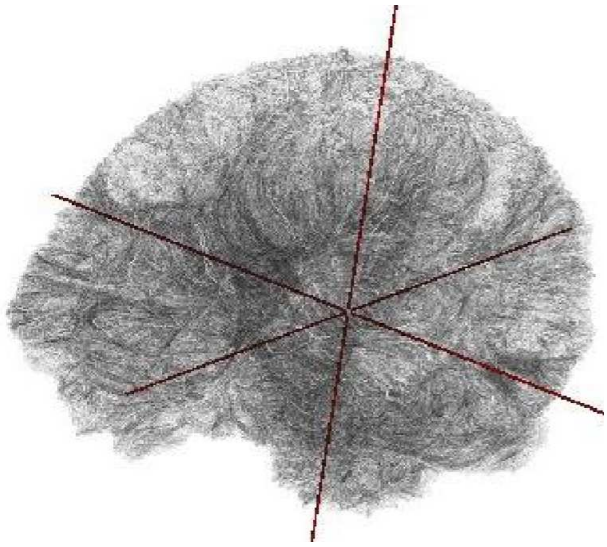


Figure 3.13: Tracts in the whole brain.

Bergmann et al. (2007); Descoteaux et al. (2009); Chao et al. (2008)] among others applied it on fields of HARDI models. As we applied the algorithm on the fODFs obtained using the SD model, we will now explain the principle of the class of streamline deterministic tractography with the fODF application example [Descoteaux et al. (2009)].

First, the technique arranges many seeds (one or several per voxel) in a region of interest (ROI) —the whole brain for example— and aims at propagating digital fibers through the seeds. To build a digital fiber, the streamline deterministic tractography algorithm considers a seed P in some voxel and estimates its fODF using interpolation. Then, it looks for the seed P' such that PP' is the fiber direction, which is the most likely, as illustrated in fig. 3.14:

$$P' = P + \Delta r \cdot \arg\max_{\mathbf{o}_i} \left(\Psi_f[P](\mathbf{o}_i)\right), \tag{3.27}$$

with Δr the step, which is a constant value defined by the user (usually $\Delta r < 0.25 \times \text{res}_v$ with res_v the voxel resolution) and $\arg\max_{\mathbf{o}_i} \left(\Psi_f[P](\mathbf{o}_i)\right)$ the most likely orientation vector corresponding to the highest $\Psi_f[P](\mathbf{o}_i)$. The curve is thus updated step by step. One stopping criterion is that the curvature cannot be greater than a threshold: the curve angle threshold is usually fixed to 60°. Another stopping criterion is that the fiber has to stay in a region where the GFA is higher than a threshold usually fixed to 0.1 [Descoteaux et al. (2009)]. This will force the digital fiber to stay in the WM. This criterion is controversial, as even if the main part of axones are located in the WM, their extremities are located in the GM. That is why other tractography masks, relying on the T_1-weighted data, have been proposed [Perrin et al. (2008); Guevara et al. (2011)].

The streamline deterministic tractography method is rapid —the algorithm parallelized on four CPU lasts 6 min for a whole brain with 1 seed per voxel— and is easy to use, but it is sensitive to initialisation choices and also to noise, which can lead to erroneous fibers: if

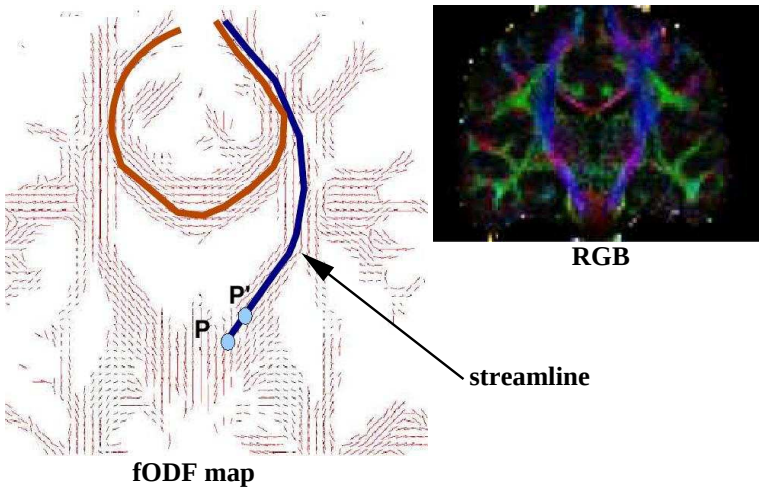


Figure 3.14: Principle of the streamline deterministic tractography done from a fODF map. The tool searches the point P' such that the fiber direction given by PP' is the most likely. Two digital fibers in the corpus callosum (CC) and in the longitudinal fasciculus are drawn in blue and red, respectively.

a tract direction given by the local model contains an error, then this error is propagated in the next streamilnes. Moreover, this tool is not always able to detect crossing fiber bundles as it searches for a unique most likely direction given by eq. 3.27.

Another class of tractography techniques, called **geodesic tractography**, was proposed in [Parker et al. (2002); Jbabdi et al. (2004); Lenglet et al. (2004); Jbabdi et al. (2008)]. It consists of constructing geodesics in a metric space corresponding to the water diffusion space. A geodesic is the pathway linking one point to the other and minimizing the distance between the two points in the metric space. A difficult aspect of the method is to find a metric space, in which geodesics represent tracts. Moreover, the use of geodesics for tractography assumes that the unknown tracts are geodesics and vice-versa with respect to the metric space derived from the local diffusion model. However, this assumption is not always valid: if a geodesic exists to connect any two points in the brain, the two points are not necessarily connected by a tract.

Streamline probabilistic tractography techniques [Parker and Alexander (2003); Perrin et al. (2005); Chao et al. (2008); Berman et al. (2008); Descoteaux et al. (2009)] make another class of tractography algorithms. They allow not to follow systematically the optimum direction in order to deal with putative errors in the angular profiles of ODFs. To choose the next pathway segment, a statistical sampler is used inside an aperture cone centered on the most probable direction given by the local model. This method cannot propagate errors such as in the streamline deterministic tractography, as there is a new statistical sampling at each tracking step, forbidding any determinisitic error propagation. The advantages of this techniques are its robustness to noise and its success in passing the crossing bundles, compared to the deterministic approaches. The disadvantage of the method is its time computation cost.

Techniques regrouped under the term of **bayesian probabilistic tractography** [Behrens et al. (2003); Friman et al. (2006); Jbabdi et al. (2007); Behrens et al. (2007); Zhang et al. (2007); Morris et al. (2008); Melie-García et al. (2008)] estimate a global connectivity using a Bayesian framework relying on a Markov Chain Monte Carlo algorithm to infer the distribution of the connectivity between regions. The result is a map containing probabilistic connectivity maps between the different brain regions.

Finally, the class of **global tractography** techniques appeared in [Poupon (1999a); Cointepas et al. (2002); Fillard et al. (2009); Kreher et al. (2008); Reisert et al. (2011)]. Instead of attempting to build a fiber tract step by step following the pathway, this technique considers each tract as a composition of tract segments to be moved in competition to obtain a whole set of tracts using a global energy minimization. This model stems from the so-called *spin-glass model* introduced in [Poupon (1999a); Cointepas et al. (2002); Fillard et al. (2009)]. The method relies on a priori information about the anatomy of WM fibers such as, for instance, the low curvature characterizing the geometry of these fibers. This last tractography algorithm better detects the fiber tracts, but requires a long computation time.

Fig. 3.15 compares global, probabilistic and deterministic approaches. It is visible that the global tractography outperforms the other methods in terms of fiber tract inference and connections retrieved. However, both global and probabilistic techniques require high computation times compared to deterministic tractography techniques. We will address in chapter 8 the feasibility of performing tractography in real-time to see the tracts being estimated and refined during the ongoing acquisition. For this real-time objective, we will have strong constraints about the time of the tractography algorithm. Therefore, we will restrict our investigations to streamline deterministic tractography as a first step in inferring connectivity in real-time.

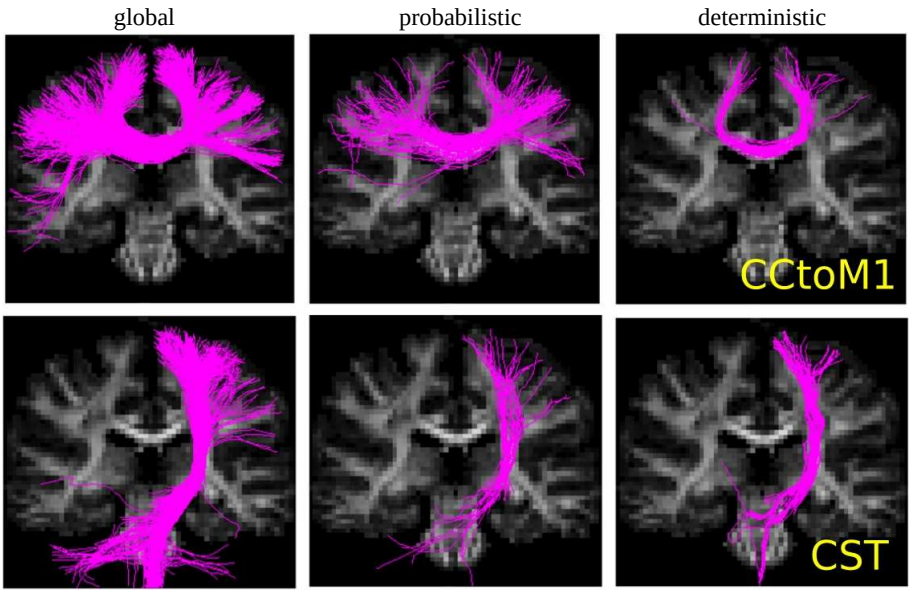


Figure 3.15: Comparison between the global, probabilistic (Gibbstracker) and deterministic tractography techniques. The top line shows the tracts linking the corpus callosum (CC) to the left motor cortex (CCtoM1). The bottom line shows the corticospinal tracts to the left motor cortex (CST). Fig. extracted from [Reisert et al. (2011)].

3.4 Conclusion of this chapter

In this chapter, we gave an overview of dMRI. We introduced the principle of sensitization of the NMR signal using the pulsed gradient spin echo sequence. Then, we introduced the notion of propagator, ODF and **q**-space and we provided a quick summary of the plethora of local diffusion models, focusing on the diffusion tensor, Q-ball and spherical deconvolution models. Finally, we made an overview of the various classes of tractography techniques available in the literature.

All these tools are the basis of the framework we will introduce in the next chapter aiming at offering a real-time environment to perform dMRI analysis immediately during ongoing scans. In the following chapters, we will present the main contributions of this thesis consisting of revisiting all the introduced dMRI algorithms to match this real-time objective.

3.5 Appendix A: The Funk-Radon transform (FRT) of the DW signal: an approximation of the dODF Ψ

The FRT of a function f for a unitary given vector $\mathbf{u}$ is written such that:

$$\mathcal{G}[f(\mathbf{w})](\mathbf{u}) = \int_{|\mathbf{w}|=1} \delta(\mathbf{u}^{\mathrm{T}}\mathbf{w}) f(\mathbf{w}) d\mathbf{w} \tag{3.28}$$

with δ the Dirac delta function. With this formula, it can be seen that the FRT of the function f is maximum in the perpendicular direction to the direction where f is maximum.

David Tuch showed that the FRT of the normalized signal E for the diffusion orientation $\mathbf{o}$ can be expressed as ([Tuch (2004)]):

$$\mathcal{G}[E(\mathbf{q},\tau)](\mathbf{o},q_0) = 2\pi q_0 \int_{\mathcal{E}^2} P(r,\theta,z) J_0(2\pi q_0 r) r dr d\theta dz, \tag{3.29}$$

where q_0 is the radius of the acquisition sphere in the $\mathbf{q}$ space; $P(r,\theta,z)$ is the diffusion propagator at the point, whose coordinates are (r,θ,z); J_0 is the Bessel function of order 0, shown in fig. 3.16 (right). The FRT of E is thus equal to the integral of the diffusion propagator P on the equator $\mathcal{E}^2$, whose plane is perpendicular to $\mathbf{o}$, as shown in fig. 3.16 (left).

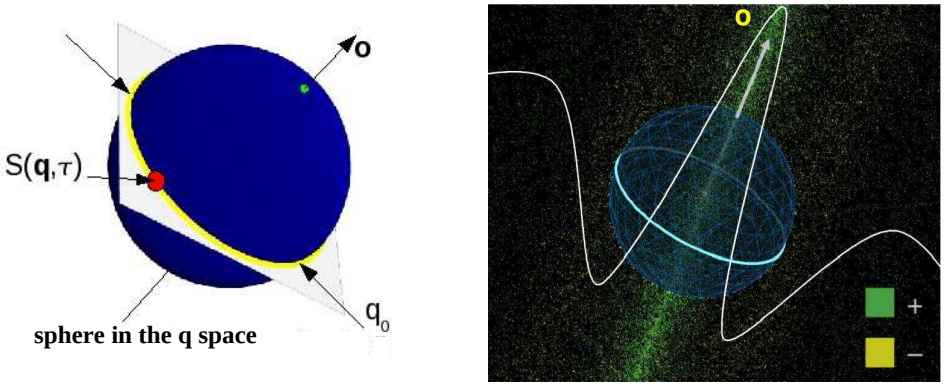


Figure 3.16: Left: the FRT of the signal E, which is measured for the direction $\mathbf{o}_i$, is equal to the integral of the diffusion propagator $P(r,\theta,z)$ on the sphere equator represented by a yellow circle. The equator plane is perpendicular to the diffusion orientation $\mathbf{o}$. Right: representation of the Bessel function of order 0, written as J_0 in the space of the displacements $\mathbf{r}$. (extract from [Tuch (2004))]

The FRT of the signal E is an approximation of the dODF aimed to be reconstructed in the diffusion local models. Indeed, in the case of a single shell acquisition on the sphere of radius q_0 previously introduced, the dODF can be written as:

$$\Psi(\mathbf{o}) = \int_{\mathcal{E}^2} P(r,\theta,z)\delta(r)\delta(\theta) r dr d\theta dz, \tag{3.30}$$

with the direction $\mathbf{o}$ along the z-axis (see appendix 7.7 from [Descoteaux (2008)]). The Dirac delta function can be approximated by a Bessel function of order 0 named J_0 (on

the right in fig. 3.16). Therefore, the FRT of E given by eq. 3.29 can approximate the dODF expressed in eq. 3.30. The mathematical details are given in the appendix 7.7 from [Descoteaux (2008)].

Chapter 4

State-of-the-art report on real-time dMRI (rtdMRI)

After having presented the dMRI modality in the previous chapter, we now tackle the real-time (RT) research field, which aims at processing data almost immediately after their acquisition. In this chapter, we focus on an RT project applied to dMRI. The contributions of this thesis entirely take place within this project. What are the concepts and goals of this project? And which methods have been investigated to achieve this innovative MRI workflow? This chapter first proposes an introduction to the *raisons d'être* of the RT project. It then explains the algorithmic methodology, which is adapted to the project constraints and enables the project's achievement. The dMRI sequences have also to be modified to fulfill the RT purpose with an optimization of the orientation set. Finally, the hardware architecture has to be reconsidered so that the project works. We end the chapter with the results of the RT project which were obtained at its beginning in 2008, as well as with some improved results. This chapter is inspired by the publication by [Poupon et al. (2008b)], from which it uses explanations, notations and figures. The chapter relies also a lot on the publication by [Deriche et al. (2009)] for the main part of the mathematical developments, as well as on [Chui and Chen (1987)] and [Kay (1993)].

4.1 Introduction

Running an rtdMRI sequence on a magnet means running a typical dMRI sequence with an adaptation for an RT processing: each DW acquired volume shall immediately be processed in the aim of getting and seeing the processing result before the acquisition of the next volume [Poupon et al. (2008b)]. The big difference with a classical dMRI sequence is therefore that the DW volume processing is no longer a *post*-processing step done after the complete acquisition —meaning when the exam is over—. On the contrary, it is an RT processing, which therefore **happens during the exam itself**. As a typical dMRI sequence consists of a succession of DW volume acquisitions for different diffusion sensitizations, the RT processing of a DW volume, for which diffusion is measured along a

given orientation, has to happen during the repetition time T_R, before the next acquisition along another orientation. To illustrate this explanation, let us consider the rtdMRI technique applied to DTI. The processing result can be the FA map for example. During the RT sequence, the console will first display the initialized FA map after the 6^{th} diffusion measurement has been done —because six orientations at least are necessary to reconstruct the diffusion tensor, as explained in 3.2.3, page 43—. Then, the console will update and refine this FA map each time a new DW volume has been acquired (*i.e.* for another orientation). The example with the FA map is shown in fig. 4.1. At the 14^{th} iteration, we notice that the map quality is already good, when less than 34% of the complete acquisition has been done. At the last iteration, the result is the same as the one obtained by a post-processing technique. rtdMRI can of course be applied to other dMRI maps, like the ADC, RGB or dODF maps, as we will see it in section 4.2. With these first explanations, we saw what does rtdMRI produce. What are now the advantages of such a RT workflow?

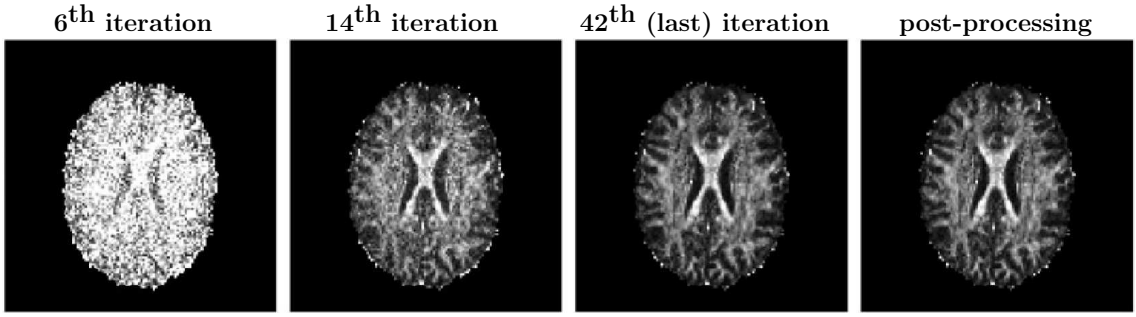


Figure 4.1: rtdMRI results (three images on the left) on an FA map built from DTI and comparison with the post-processing result (last image on the right). The FA map is first initialized after six iterations (because six orientations at least are necessary to reconstruct the diffusion tensor). Then, the map is refined gradually with the next iterations. At the 14^{th} iteration, the map quality is already good, when less than 34% of the complete acquisition has been done. Fig. extracted from [Poupon et al. (2008b)].

There are different motivations for such a RT workflow. From the patient point of view, the optimization of the sequence, which produces a reduced scan duration, results in **more comfort for the patient**. Indeed, the patient has to stay motionless during the whole MRI exam which can be constraining. For claustrophobic patients, the exam is also painful. Therefore, if the exam is shorter, it is easier for the patient. A RT workflow also decreases the number of non-diagnostic results and therefore generates **less waste of data**, as the quality of the data are verified during their acquisition. The presence of an artifact can be instantaneously detected and the data can either be corrected in RT or removed and acquired again. It is worth mentioning that concerning the antenatal and neonatal studies, there is today 75% of the exams which is thrown away because of the uncontrolled motion of the foetus or the child. A similar problem exists for exams of epileptic, Parkinson's and Huntington's patients who cannot control their movements.

From the methodological point of view, such a RT workflow gives the possibility to **correct in RT** for artifacts and to **restart an acquisition** when necessary. Concerning motion artifacts, online motion correction techniques have been proposed [Zaitsev et al. (2006); Quin et al. (2009); Ooi et al. (2009); White et al. (2010); Aksoy et al. (2011); Maclaren et al. (2012)]. With the RT process, the acquisition or processing parameters can be modified during the acquisition to refine the results. That is what is called **feedback**. Furthermore, the number of scans can be optimized. For example, the physician can decide to stop the acquisition if the image quality is sufficient for him to deliver his diagnosis, as in the case of fig. 4.1, at the 14^{th} iteration for example.

All these advantages apply for the particular case of dMRI. Concerning the RT detection of motion, methods dedicated to dMRI have been developed in [Caruyer et al. (2010, 2011)] leading to two possible strategies: either the acquisition, for which motion occurs, has to be redone, or the acquisition plan has to be modified. With rtdMRI, it is also possible for the physician to deliver a diagnosis during the exam. When a patient has ischemia, it is essential to detect it as soon as possible to reduce the time between the beginning of ischemia and the medical treatment. An immediate knowledge about the state of the fibers where ischemia occurred could even orientate a specific treatment. Furthermore, the time saving advantage brought by this RT workflow is particularly interesting for local diffusion models requiring a long acquisition time compared to clinical routines. The use of such models is easier with rtdMRI, which permits to use only as measurements as necessary and to stop the acquisition once the results are sufficient. Finally, it is also possible with this RT process to retro-propagate an information about the fiber orientation to better choose the next orientations along which the diffusion will be measured.

The low cost of high parallel computing systems with respect to the cost of MRI systems could in the future yield novel information technology architectures enabling direct interactions between acquisition, pre- and post-processing that would improve the workflow of MRI. Not only could the errors of acquisition be corrected online, but the diagnosis may also be done in RT. The topics of this thesis is fully dedicated to propose such a RT environment with an application to dMRI.

In the next section, we will make a review of the available techniques to deal with incremental influence.

4.2 Incremental frameworks

4.2.1 Review of incremental frameworks

The first attempt to solve a general estimation problem and extract a signal of interest x in a signal-plus-noise setting such as $y = f(x) + \epsilon$, with y the measurement, f a function and ϵ the noise, was performed with the least-squares estimator [Gauss (1809); Legendre (1805)]. This method can be used recursively for stationary signals and is then called the **recursive (or sequential) least squares** approach. The least squares technique minimizes a weighted linear least squares cost function, corresponding to the sum of the

squares of the errors, to estimate the signal x from the data y. No probabilistic assumptions are made about the measured data. This method is valid for Gaussian, as well as non-Gaussian noise. However, the performance of the estimator are dependent on the noise properties. According to the Gauss-Markov theorem, the least squares estimator is optimal only if the function f is linear, with a white noise ϵ.

The next incremental framework appeared with the **Wiener filter**, corresponding to a **sequential linear minimum mean square error** (sequential LMMSE) estimator [Wiener and Hopf (1931); Kolmogorov (1941b,a); Wiener (1949)]. The latter is a linear noise filter dedicated to stationary signals. It minimizes the mean square error (MSE) defined for an estimator $\hat{x}$ such that: $\text{MSE}(\hat{x}) = E[(\hat{x} - x)^2]$, with $E[\cdot]$ the expectation operator. The Wiener filter contributed to give rise to the well-known and often used **Kalman filter** (KF) [Kalman (1960); Kalman and Bucy (1961)], for which some equivalent ideas can be found in the earlier works of [Thiele (1880); Swerling (1959)]. This filter was the first to be adapted to non-stationary signals. Based on a dynamic system between the measurements (*observations*) and the unknown parameters (*state vector*), the KF predicts the evolution of the state vector with respect to the previous measurements. Moreover, it also corrects and refines the dynamic result, each time a new measurement is performed. The KF is based on a linear system with a zero-mean Gaussian assumption concerning the noise corrupting the system. It can easily be adapted to a Gaussian noise with a non-zero mean. Other filters extend the KF to non-Gaussian, non-linear filtering, like the **extended KF** (EKF) described in [Jazwinski (1970); Gelb et al. (1974); Anderson and Moore (1979)], which locally linearises the evolution model using Taylor series expansions. However, the EKF is only reliable for almost linear models and may diverge in other cases [van der Merwe et al. (2000)]. To overcome the limitation of the EKF, the **unscented KF** (UKF) was proposed [Julier et al. (1995); Julier and Uhlmann (1997); Wan and van der Merwe (2000)]. The latter relies on an *unscented transformation*, which uses a set of sampled points to configure the means and covariances of probability distributions. The UKF was shown to outperforms the EKF, but the UKF can in general only be applied on models driven by Gaussian noises [van der Merwe et al. (2000)].

In parallel to the developments of derived Kalman filters, **point-mass approaches** were developed for the state estimation of discrete-time nonlinear non-Gaussian stochastic systems [Bucy (1969); Bucy and Senne (1971)]. In opposition to the EKF and UKF, these methods are called global, as they approximate the probability density functions (PDFs) of the states in a large area of the state space [Šimandl et al. (2006)]. These methods are based on the discretization of the state space by a grid. They estimate the densities only at a finite fixed set of points. Such techniques are dedicated to small state spaces, as they are computationally demanding.

Another famous approach to the estimation of dynamic nonlinear systems is the **sequential Monte Carlo approach**, also called as **particle filter** or **bootstrap filter**, which combines a Monte Carlo sampling method with recursive Bayesian filtering [Gordon et al. (1993); Kong et al. (1994); Liu and Chen (1995, 1998); Pitt and Shephard (1999);

Doucet et al. (2000)]. Following the idea of the point-mass approach, the sequential Monte Carlo technique uses a Monte Carlo simulation to describe the densities through a set of samples. For more details on the different types of particle filters, we invite the reader to read [Chen (2003)].

Gaussian mixture approximations represent another solution for nonlinear, non-Gaussian filtering [Sorenson and Alspach (1971); Alspach and Sorenson (1972); Masreliez (1975); Plataniotis et al. (1997); Ito and Xiong (2000)]. The idea is to use a weighted sum of Gaussian densities to approximate another density function. The resulting Gaussian mixture can then be injected in a bank of KFs or EKFs to be run in parallel. The **mixture of Kalman filters** proposed by [Chen and Liu (2000)] proposes to use a Gaussian mixture model with a bank of EKFs, in which each EKF is run with a Monte Carlo sampling approach.

From these different estimators dedicated to dynamic systems, other methods combining different techniques, *e.g.* the unscented particle filter which associates the UKF and a particle filter, were developed and we invite the reader to read [Chen (2003)] for more details on these hybrid approaches.

The rtdMRI framework proposed by [Poupon et al. (2008b)] is based on a KF. The next subsection presents this incremental estimator in details.

4.2.2 The Kalman Filter (KF)

The KF, invented by Rudolf Kalman in 1960 [Kalman (1960)], is a sequential estimator enabling to update, and even predict, from observations measured over time, the results of a known linear model. It works in a recursive way and minimizes the mean square error of the estimation: it is a *sequential minimum mean square error* (sequential MMSE) estimator. The model, on which it relies, can be static or dynamic, but has to be linear. This filter is used in very various fields, some of which are: object tracking, computer-assisted navigation, weather forecast, production trends and cycles in economy. It is a very rapid tool enabling RT studies, as the KF estimations are produced almost instantaneously after the measurement is taken.

In our dMRI framework, the common dMRI results are maps of the brain architecture that are static at the macroscopic scale. They represent what we want to infer in RT. Consequently, in our case, the linear model required for the KF comes down to:

$$\boxed{\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}}, \text{ with } \begin{cases} \mathbf{y}\text{: the K}\times 1 \textbf{ observation vector}, \\ \mathbf{A}\text{: the K}\times\text{N design matrix}, \\ \mathbf{x}\text{: the N}\times 1 \textbf{ state vector}, \\ \boldsymbol{\epsilon}\text{: the K}\times 1 \text{ vector of } i.d. \text{ noises distributed according to } \mathcal{N}(0, \mathbf{R}), \end{cases} \tag{4.1}$$

where the vector $\mathbf{y}$ is filled with a new observation at each iteration i, i going from 1 to K. $\mathbf{A}$ is the design matrix and $\mathbf{x}$ is the searched state vector. The *i.d.* noises are

independently distributed random variables, which are, in the KF framework, normally distributed according to the probability distribution function (PDF) $\mathcal{N}(0, \mathbf{R})$ with mean equal to zero and with K×K covariance matrix $\mathbf{R} = E\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\right]$. As the noise channels are independent, the matrix $\mathbf{R}$ is diagonal with its diagonal elements written as $R_i \; \forall i \in [\![1; K]\!]$ ($[\![1; K]\!]$ meaning the intervall of integers varying from 1 to K). In eq. 4.1, $\boldsymbol{\epsilon}$ therefore represents a zero-mean Gaussian noise [Kay (1993)]. If, additionally, $\mathbf{R}$ can be written as $\mathbf{R} = R \cdot \mathbf{I}$, with R a constant and $\mathbf{I}$ the identity matrix, then the noise becomes white Gaussian distributed. The hypothesis of a zero-mean Gaussian noise required for the KF constitutes a limitation for the estimator, which is not optimal when noise is not Gaussian distributed. This limitation has to be well kept in mind and we will further deal with it when considering the RT noise correction issue.

The goal of a RT study is to estimate the state vector $\mathbf{x}$ gradually with the measurements performed over time. The state vector has to be updated each time the vector $\mathbf{y}$ is filled with a new observation. In our static dMRI framework, the design matrix, as well as the state vector are static. As a consequence, the Kalman equations' system enabling to estimate the state vector at the i^{th} iteration does not contain any prediction step. The KF procedure in this static case simplifies into a Wiener filter and is written such that:

$$\begin{cases} \text{innovation: } \nu_i = y_i - \mathbf{a}_i \hat{\mathbf{x}}_{i-1}, \\ \text{innovation covariance: } s_i = R_i + \mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathbf{T}}, \\ \text{gain: } \mathbf{k}_i = s_i^{-1} \cdot \mathbf{P}_{i-1} \mathbf{a}_i^{\mathbf{T}}, \\ \text{state estimate at iteration } i\text{: } \hat{\mathbf{x}}_i = \hat{\mathbf{x}}_{i-1} + \nu_i \mathbf{k}_i, \\ \text{estimation error covariance at iteration } i\text{: } \mathbf{P}_i = \mathbf{P}_{i-1} - \mathbf{k}_i \mathbf{a}_i \mathbf{P}_{i-1}, \end{cases} \tag{4.2}$$

with $\mathbf{a}_i = [\mathbf{A}_{i1}, ..., \mathbf{A}_{iN}]$ the i^{th} row of the known matrix $\mathbf{A}$. Let us now explain how to read this equation system 4.2, which delivers at iteration i the **state vector estimate** $\hat{\mathbf{x}}_i$, which we are interested in. In our system, the estimation error at iteration i is equal to $\mathbf{x} - \hat{\mathbf{x}}_i$ with covariance $\mathbf{P}_i = E\left[\left(\mathbf{x} - \hat{\mathbf{x}}_i\right)\left(\mathbf{x} - \hat{\mathbf{x}}_i\right)^{\mathrm{T}}\right]$. During the KF process, the trace of the **estimation error covariance** is minimized. The trace value corresponds to the *mean square error*. This quantity at each iteration is therefore an indicator of the procedure progress. The **innovation**, also called residual, corresponds to the difference between the actual observation y_i and the estimated observation $\mathbf{a}_i \hat{\mathbf{x}}_{i-1}$. To obtain the innovation covariance, let us now express the innovation as $\nu_i = y_i - \mathbf{a}_i \mathbf{x} + \mathbf{a}_i \left(\mathbf{x} - \hat{\mathbf{x}}_{i-1}\right)$. In the latter equation, the term $y_i - \mathbf{a}_i \mathbf{x}$ corresponds to the noise ϵ_i whose covariance is $R_i = E\left[\epsilon_i \epsilon_i^{\mathrm{T}}\right]$. The term $\mathbf{x} - \hat{\mathbf{x}}_{i-1}$ corresponds to the estimation error whose covariance is $\mathbf{P}_i$. Therefore, using the linear combination property of the variance, we easily calculate the variance of ν_i and obtain the **innovation covariance** expression given by the equation system 4.2. Finally, to calculate the state estimate, as well as the estimation error covariance, we need to compute the **gain**. The latter is calculated so that it minimizes the trace of the estimation error covariance [Maybeck (1979)]. A Bayesian demonstration of the KF equations can be done relying on this statement [Kay (1993)]. It is also possible to obtain

the system given by 4.2 using a *weighted least squares* (WLS) approach, which minimizes the cost function given by $(\mathbf{y} - \mathbf{Ax})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y} - \mathbf{Ax})$ [Chui and Chen (1987); Deriche et al. (2009)]. More details on the KF equations' derivation using a WLS approach are given in appendix A (section 7.6) at the end of this chapter. The gain then permits to estimate the state vector at iteration i with respect to the estimate at the previous iteration. The same is done to obtain the estimation error covariance. Fig. 4.2 summarizes the process of the KF, as we use it (*i.e.* like a Wiener filter).

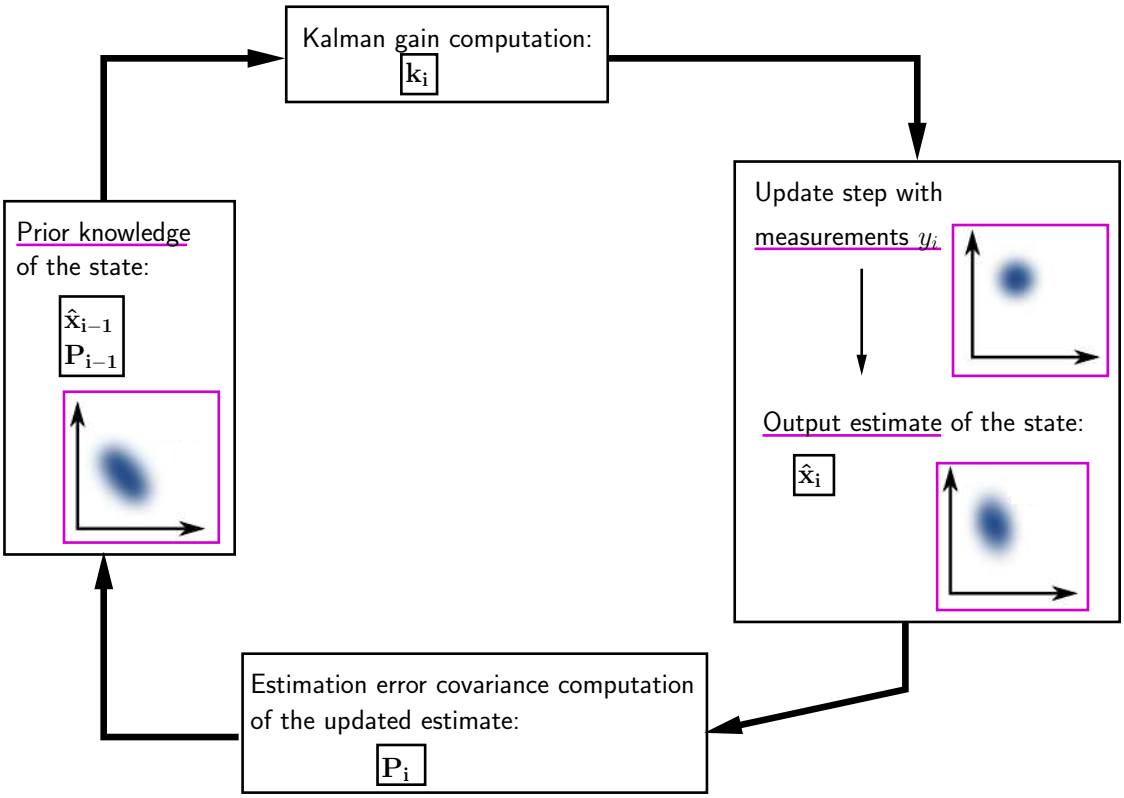


Figure 4.2: Diagram of the KF used as a Wiener filter. Fig. adapted from en.wikipedia.org and ccar.colorado.edu. The estimation of the state is refined at each new iteration i.

Finally, to utilize the equation system 4.2, an initialization of the unknown $\hat{\mathbf{x}}$ and $\mathbf{P}$ is required. A good initialization favors the KF convergence. In the KF theory, the initial state vector $\mathbf{x}_0$ is supposed to follow a normal distribution with mean $\boldsymbol{\mu_x}$ and variance $\mathbf{V_x}$. This leads to set: $\hat{\mathbf{x}}_0 = \mu_\mathbf{x}$ and $\mathbf{P}_0 = \mathbf{V_x}$. However, concretely, when no information on the distribution of $\mathbf{x}_0$ is known, then it is common to set: $\hat{\mathbf{x}}_0 = \mathbf{0}$ and $\mathbf{P}_0 = V\mathbf{I}$, with V big enough so that the KF has little trust in the initial estimation. In this case where no prior probabilistic information is known, the KF is not Bayesian anymore and behaves like a classical *sequential WLS* estimator. If, additionally, $\mathbf{R} = \mathbf{I}$, then the cost function to minimize with the least squares approach becomes $(\mathbf{y} - \mathbf{Ax})^{\mathrm{T}} (\mathbf{y} - \mathbf{Ax})$ and the KF behaves no longer as a sequential WLS estimator, but as a *sequential ordinary least squares* (OLS) estimator. In case of noise variance inhomogeneities (heteroscedasticity),

the OLS estimator is less accurate than the WLS estimator because it gives the same weight to every measurement and thus does not account for the noise variance inequalities towards the iterations. These OLS and WLS approaches are valid even if the ϵ_i are not independent noises distributed according to zero-mean Gaussian PDFs. However, they may not be optimal in this case. On the contrary, if the ϵ_i are *i.d.* noises according to zero-mean Gaussian PDFs, the OLS estimator, in case of homoscedasticity, and the WLS estimator in case of heteroscedasticity, are said to be efficient because they then correspond to the *minimum variance unbiased estimator* [Kay (1993)]. Consequently, if the independent zero-mean Gaussian noise condition is required for the KF, it is also recommended, for a good efficiency in the estimation, when the KF is simply used as a sequential WLS estimator.

Using these equations given by 4.2 enables to estimate the state vector $\mathbf{x}$ and to update it each time a new observation is done. While the vector $\mathbf{y}$ is filled with measurements, the estimation of $\mathbf{x}$ is refined. Even if the KF as we use it here simplifies into a LSE, we kept the KF notation all along this thesis. This KF tool can be applied on all voxels of an acquired volume in dMRI, with the requirement of having put in place a linear model adapted to dMRI that respects the KF constraints, like the independent zero-mean Gaussian noise condition.

4.3 Adapting dMRI models to the Kalman framework

In this section we first introduce the frame adopted in this thesis and we highlight the major noise issue to be considered in the KF context. We then explain how to adapt the dMRI models of the DTI, aQBI and sa-aQBI reconstruction algorithms to the Kalman framework. The models presented here are fundamental for the rtdMRI workflow. They constitute the basis for a KF application, as seen in the previous section. **Let us remind here that we write the raw (therefore noisy) and the noise-free DW signal M and S, respectively. We will also now write the raw and noise-free T_2-weighted signal M_0 and S_0, respectively.**

4.3.1 Brief introduction on the frame adopted in this thesis with a focus on noise issues

Before beginning to apply the KF on the DTI, aQBI and sa-aQBI diffusion models, it is essential to globally consider our real framework and think about possible dMRI issues that could interfere with the KF ideal frame. Indeed, when wanting to apply some theoretical solution to a real situation, there are often real parameters that need to be considered because they can perturb the ideal approach. This constitutes the confrontation between the theory and its experimental use.

When we think of what could interfere with the KF frame presented in the previous section, we consider the artifacts occuring during a dMRI exam. There are many possible artifacts. In this thesis however, we will mainly focus on one artifact: the **noise**

issues, which directly impact the KF application. Concerning the susceptibility artifacts (geometrical distortions) which happen in regions in between two tissues having different magnetic susceptibilities, a $\mathbf{B_0}$ field map is acquired before the dMRI sequence to apply a posteriori a correction on the DW data [Jezzard and Balaban (1995)]. For small motion of the subject, the manufacturer already provided corrections based on online inference of the motion using registration techniques and reorientation of the next slice position accordingly. In the case of dMRI, Eddy currents have to be corrected too when a single spin echo sequence is used. However, it can easily be combined with motion correction. More details are provided in appendix A, page 249, where we describe how the real data used in our study were acquired. For a good understanding of the following subsections, it is necessary to discuss the basis of noise analysis, even though more details are given in chapter 5, section 5.1. We perform here a noise analysis considering the DW signal M, but the same analysis could be performed on any measured signal coming from the MRI scanner. Let us express, for a given voxel and a given diffusion orientation, the DW measured signal M with respect to the corresponding noise-free signal S:

$$M = S + \epsilon, \tag{4.3}$$

where ϵ is the acquisition noise. On most clinical MRI systems, M is acquired using one of the parallel MRI techniques described in chapter 2, section 2.7. With parallel MRI, each channel of the phased-array coil receives a signal and all signals coming from all channels have to be combined together to build the final signal M. A way to do the combination is to use the **sum of squares** (SoS) reconstruction, which is defined such that [Constantinides et al. (1997)]:

$$M = \sqrt{\sum_{c=1}^{n} [(S_{r_c} + \epsilon_{r_c})^2 + (S_{i_c} + \epsilon_{i_c})^2]}, \tag{4.4}$$

with c the index designating a channel among the n channels of the phased-array coil. S_{r_c} and S_{i_c} are the real and imaginary parts, respectively, of the complex signal S_c received by the channel c. Similarly, ϵ_{r_c} and ϵ_{i_c} are the real and imaginary parts, respectively, of the complex noise ϵ_c present on the channel c. There are other combination methods than the SoS, but in this thesis, we will only consider the latter, as it was the one used in the experiments. In the noise literature, both ϵ_{r_c} and ϵ_{i_c} are assumed to be independent Gaussian processes with a mean equal to zero and a variance σ^2 [Henkelman (1985)]. We will detail the reasons of this hypothesis which are linked to the noise origins in section 5.1. From this assumption, it is possible to infer the PDF followed by M: it is a **noncentral $\boldsymbol{\chi}$** (nc-χ) distribution [Constantinides et al. (1997)]. This PDF followed by M is defined with three parameters: the noise-free DW signal S, the variance σ^2 of the noise present on each channel of the receiving coil and the number of channels n. It is a function, somewhat looking like a Gaussian distribution, but with a certain positive skewness making the PDF asymmetrical compared to a Gaussian distribution. The nc-χ distribution is called a Rician distribution when $n = 1$. For more details, see section 5.1. The nc-χ distribution

can be approximated by a **Gaussian of mean S and variance σ^2** when the SNR, defined as SNR$= S/\sigma$, is **moderate to high** and when the number of channels n is **low** (like $SNR \geq 4$ for $n = 1$) [Henkelman (1985); Sijbers et al. (1998)]. Indeed, fig. 4.3 shows the convergence of $E(M)$ on S, as well as the convergence of the noise standard deviation of M, written as σ_M, on σ for high SNR values. Considering now the noise term $\epsilon = M - S$, its mean is $E(\epsilon) = E(M) - E(S)$, using the linearity of the expectation operator. Therefore $E(\epsilon)$ converges on 0. The noise standard deviation of ϵ, written as σ_ϵ, is equal to σ_M and therefore converges on σ. As the number of channels n increases, the approximation of a zero-mean Gaussian noise with standard deviation σ is less true, as shown in fig. 4.3. Therefore, we have to keep in mind that this approximation does not enable to account for the true noise nature, especially at low SNR values or with high values of n.

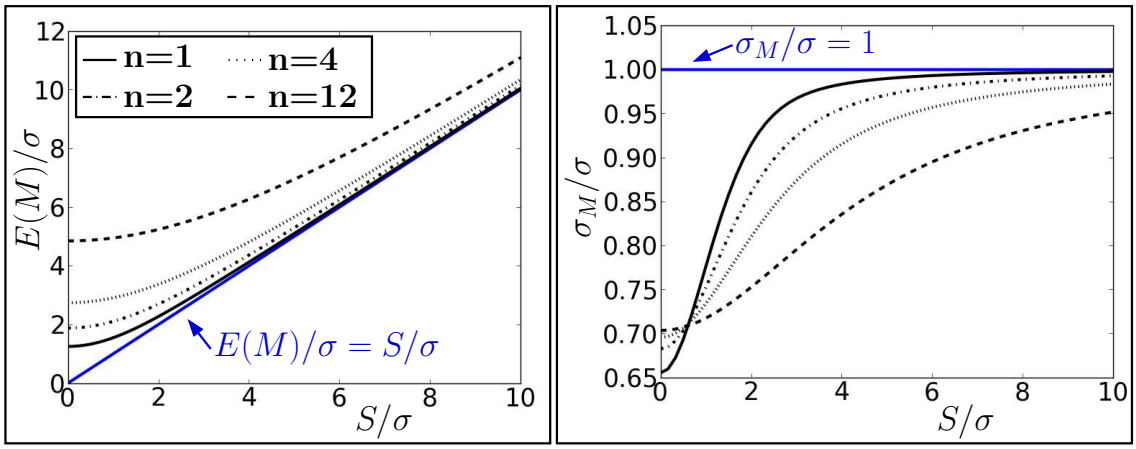


Figure 4.3: Mean and noise standard deviation of the PDF followed by M

With all this information, let us now tackle the KF model application in dMRI. First, we will explain the adaptation of the DTI model to a KF.

4.3.2 DTI Kalman framework

DTI, introduced in subsection 3.2.3, proposes to consider water molecules' diffusion as a Gaussian process and expresses the noise-free DW signal $S(\mathbf{o}_i)$ —$\mathbf{o}_i$ corresponding to the diffusion orientation given by $[o_x, o_y, o_z]^T$— such that:

$$S(\mathbf{q}_i, \tau) = S_0 e^{-b_i \mathbf{o}_i^T \mathbf{D} \mathbf{o}_i}, \tag{4.5}$$

where b_i is the *b-value* at iteration i, the iteration corresponding to the measurement along the orientation $\mathbf{o}_i$. Let us notice that this equation is written for a considered voxel: let us keep in mind that all equation variables are voxel-wise, except b_i and $\mathbf{o}_i$. $\mathbf{D}$ is the diffusion tensor, the eigenvalues of which are required to generate the conventional DTI maps (FA, MD, PD, TD). The coefficients of $\mathbf{D}$ are the unknowns. Now, if we assume that we have access to the noise-free T_2 weighted signal S_0, we can then concretely apply eq. 4.5 to the

measured DW signal $M(\mathbf{o}_i)$, such that:

$$M(\mathbf{o}_i) = S_0 e^{-b_i \mathbf{o}_i^T \mathbf{D} \mathbf{o}_i} + \mu_i, \tag{4.6}$$

where μ_i corresponds to the error and corresponds to the acquisition noise if we assume that there is no other artifact. The acquisition noise μ_i is assumed to be independent from an iteration i to another. As seen in the previous subsection, we know that μ_i is distributed according to a PDF that makes $M(\mathbf{o}_i)$ follow a nc-χ distribution.

In the KF, $\mathbf{D}$ corresponds to the state vector $\mathbf{x}$ of section 4.2. To transform eq. 4.6 in a linear equation, the natural logarithm of $\frac{M_0}{M(\mathbf{o}_i)}$ has to be expressed [Poupon et al. (2008b)]:

$$\boxed{ln\left(\frac{S_0}{M(\mathbf{o}_i)}\right) = b_i \mathbf{o}^T \mathbf{D} \mathbf{o}_i - ln\left(1 + \frac{\mu_i}{S_0} e^{b_i \mathbf{o}_i^T \mathbf{D} \mathbf{o}_i}\right)}. \tag{4.7}$$

From this eq. 4.7, we can deduce the following linear model:

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}, \tag{4.8}$$

where the observations' vector is written as $\mathbf{y} = [y_1, ..., y_K]^T$ with $y_i = ln\left(\frac{S_0}{M(\mathbf{o}_i)}\right)$ for iteration i corresponding to the measurement along the diffusion orientation $\mathbf{o}_i$. The state vector $\mathbf{x}$ corresponds to the **six coefficients of the symmetric diffusion tensor**: $\mathbf{x} = [D_{xx}, D_{xy}, D_{xz}, D_{yy}, D_{yz}, D_{zz}]^T$. The design K×6 matrix $\mathbf{A}$ has rows designated by $\mathbf{a}_1, ..., \mathbf{a}_K$, whose expression is (here for the i^{th} row): $\mathbf{a}_i = b_i[o_{i_x}^2, 2o_{i_x}o_{i_y}, 2o_{i_x}o_{i_z}, o_{i_y}^2, 2o_{i_y}o_{i_z}, o_{i_z}^2]$. Finally, the K×1 noise vector $\boldsymbol{\epsilon}$ is expressed such that: $\epsilon_i = -ln\left(1 + \frac{\mu_i}{S_0} e^{b_i \mathbf{o}_i^T \mathbf{D} \mathbf{o}_i}\right)$ for iteration i. ϵ_i can also be written as: $\epsilon_i = ln\left(\frac{S_0}{M(\mathbf{o}_i)}\right) - ln\left(\frac{S_0}{S(\mathbf{o}_i)}\right) = ln\left(S(\mathbf{o}_i)\right) - ln\left(M(\mathbf{o}_i)\right)$. To satisfy the conditions imposed by the KF on the linear model, the noises ϵ_i have to be *i.d.* according to zero-mean Gaussian distributions. The noises μ_i of eq. 4.6 are indeed *i.d.* and so are also the noises ϵ_i. However, μ_i is not Gaussian distributed, as seen in the previous subsection, and the distribution followed by ϵ_i should be studied to know whether it is a Gaussian PDF or not. Nevertheless, for the authors of [Poupon et al. (2008b)], the first goal was not to perfectly account for the noise nature, but to propose a RT model for DTI. Therefore, they assumed that the noises ϵ_i were distributed according to $\mathcal{N}(0, \mathbf{R})$, with $\mathbf{R} = \mathbf{I}$. As they initialized the KF with no prior probabilistic information ($\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{P}_0 = V\mathbf{I}$ with $V = 10^6$), their KF method simplifies into a sequential OLS estimator. However, it is worth mentioning that [Salvador et al. (2005)] studied the distribution followed by ϵ_i, when $n = 1$, therefore in the case of a Rician noise, and they obtained the three following results on the noise vector $\boldsymbol{\epsilon}$:

- $E[\boldsymbol{\epsilon}] \approx 0$.
- $\mathbf{R} = R_i \cdot \mathbf{I}$ with $R_i \approx (\sigma_i / S(\mathbf{o}_i))^2$, where σ_i^2 is the variance of the Gaussian noises on each real and imaginary parts of the receiving channel.
- For SNR values higher than 5, the noise distribution is approximately a zero-mean Gaussian PDF with variance $\mathbf{R}$.

The expression of R_i in the second item can be mathematically confirmed using the propagation of uncertainty. Let us define z_i, corresponding to the noise-free y_i, and expressed as $z_i = ln\left(\frac{S_0}{S(\mathbf{o}_i)}\right)$. The error on z_i, written as δz_i, can be indeed related to the error on $S(\mathbf{o}_i)$, written as $\delta S(\mathbf{o}_i)$, through: $\delta z_i = \frac{\partial z_i}{\partial S(\mathbf{o}_i)}\delta S(\mathbf{o}_i) = -\frac{\delta S(\mathbf{o}_i)}{S(\mathbf{o}_i)}$. Applying the variance on this result gives us: $\mathrm{Var}\left(\delta z_i\right) = \mathrm{Var}\left(\delta S(\mathbf{o}_i)\right)/S(\mathbf{o}_i)^2$. $\mathrm{Var}\left(\delta S(\mathbf{o}_i)\right)$ corresponds to the variance of the nc-χ distribution followed by $M(\mathbf{o}_i)$ and it is different from σ_i. However, as seen in the previous subsection, when the SNR is high, the nc-χ distribution can be approximated by a Gaussian PDF of mean $S(\mathbf{o}_i)$ and variance σ_i^2. Using this and as $\mathrm{Var}\left(\delta z_i\right) = R_i$, we obtain the approximation $R_i \approx \left(\sigma_i/S(\mathbf{o}_i)\right)^2$. Applying these results in the DTI Kalman framework, the authors of [Casaseca-de-la-Higuera et al. (2012)] improved the KF model developed for DTI by [Poupon et al. (2008b)]: indeed, the latter model permits to use a KF, which behaves as a sequential OLS estimator, whereas the KF model in [Casaseca-de-la-Higuera et al. (2012)] behaves, with $\mathbf{R} = R_i \cdot \mathbf{I}$, as a sequential WLS estimator. This WLS estimator is more accurate, as the results obtained by Pablo Casaseca-de-la-Higuera *et al.* confirm it. For more details, we invite the reader to read their publication. The DTI Kalman framework with both assumptions on $\mathbf{R}$ (for an OLS or a WLS estimation) is summarized in table 4.1.

To apply this KF model, we also need to know the noise-free T_2-weighted signal S_0, which appears in eq. 4.7. We can either approximate it directly by the measured T_2-weighted signal M_0 or we can apply a correction on M_0 and use the resulting corrected T_2-weighted signal $\hat{S}_0$ in replacement of S_0 in eq. 4.6. The results shown later in this chapter were obtained following the first solution, *i.e.* replacing S_0 by M_0 in eq. 4.6, as the authors of [Poupon et al. (2008b)] and [Casaseca-de-la-Higuera et al. (2012)] did, assuming that $M_0 = S_0$.

Table 4.1: The DTI Kalman framework

$\boxed{\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}}$, with for iteration i:

$$\begin{cases} y_i = ln\left(\frac{S_0}{M(\mathbf{o}_i)}\right) \\ \mathbf{a}_i = b_i[o_x^2, 2o_xo_y, 2o_xo_z, o_y^2, 2o_yo_z, o_z^2] \\ \mathbf{x} = [D_{xx}, D_{xy}, D_{xz}, D_{yy}, D_{yz}, D_{zz}]^T \\ \epsilon_i \text{ supposed to follow } \mathcal{N}(0, \mathbf{R}) \end{cases}$$

Definition of the noise covariance matrix $\mathbf{R}$:

$\mathbf{R}$ is diagonal, such that:
- either $\mathbf{R} = \mathbf{I}$ following [Poupon et al. (2008b)],
- or $\mathbf{R} = R_i\mathbf{I}$ with $R_i \approx (\sigma_i/S(\mathbf{o}_i))^2$ following [Casaseca-de-la-Higuera et al. (2012)]

Initialization for the KF equations' system 4.2:

$$\begin{cases} \mathbf{x}_0 = \mathbf{0} \\ \mathbf{P}_0 = V\mathbf{I} \text{ with } V = 10^6. \end{cases}$$

The equations' system given by 4.2 can be applied on the model described in this table.

We will see in subsection 4.5.1 that this set up system achieves its goal of generating DTI maps in real-time. What are now the aQBI and sa-aQBI Kalman frameworks?

4.3.3 The aQBI and sa-aQBI Kalman frameworks

Here, we develop all aQBI and sa-aQBI model frameworks, which we use as input of a KF. We begin with the aQBI models.

aQBI Kalman framework

Contrary to DTI, aQBI does not assume that the water molecules' diffusion follows a Gaussian distribution. A consequence of that is that aQBI permits to better infer fiber bundles' crossings (for more details see subsection 3.2.5). For this reason, this technique is well-spread in the research world, but is actually not commonly used in clinical protocols because of its too long required acquisition time. However a RT process enabling to give the aQBI results almost instantaneously, while the patient is still in the magnet, could seriously be considered for a clinical exam protocol. Indeed, rtdMRI permits to make as many acquisition measurements as is necessary, with online processing and visualization,

and thus to gain a lot of acquisition time, as explained in section 4.1. The possibility of RT aQBI is demonstrated in [Poupon et al. (2008b)], from which we repeat some explanations.

As seen in subsection 3.2.4, we can write the noise-free DW normalized signals' vector $\mathbf{E} = [S(0)/S_0, ..., S(K)/S_0]^T$ as:

$$\mathbf{E} = \mathbf{B} \cdot \mathbf{C}^{\mathrm{DW}}, \tag{4.9}$$

with $\mathbf{B}$ the matrix of the modified SH basis defined with eq. 3.13. $\mathbf{C}^{\mathrm{DW}}$ is the N×1 coefficients' vector of the decomposition of the noise-free DW signals' vector on the modified SH basis. Now, we can apply eq. 4.9 to the measured DW normalized signals' vector $\mathbf{M_E} = [M(0)/S_0, ..., M(K)/S_0]^T$ such that:

$$\boxed{\mathbf{M_E} = \mathbf{B} \cdot \mathbf{C}^{\mathrm{DW}} + \boldsymbol{\epsilon}}, \tag{4.10}$$

where $\boldsymbol{\epsilon}$ represents the vector containing the acquisition noise normalized by the noise-free T_2-weighted signal, assuming that there is no other artifact corrupting the DW signal. As mentioned in the previous subsection, the acquisition noise is independent from an iteration i to another. Its PDF makes $M(\mathbf{o}_i)$ follow a nc-χ distribution. For the moment, we will approximate the noise PDF by a zero-mean Gaussian PDF, as required by the KF model constraints summed up by eq. 4.1. The authors of [Poupon et al. (2008b)] worked with a sequential OLS estimator, *i.e.* $\mathbf{R} = \mathbf{I}$, but we could also work with a sequential WLS estimator with $\mathbf{R} = R_i\mathbf{I}$ where $R_i = (1/S_0^2)\mathrm{Var}\left(\delta S(\mathbf{o}_i)\right)$. Using as before the approximation of the nc-χ distribution followed by $M(\mathbf{o}_i)$ by a Gaussian PDF of mean $S(\mathbf{o}_i)$ and variance σ_i^2, we could approximate R_i by σ_i^2/S_0^2. We will come back later on the noise considerations, which are here not much respected. Eq. 4.10 enables therefore to apply the KF on the observations' vector $\mathbf{M_E}$, with the design matrix $\mathbf{B}$ and the state vector $\mathbf{C}^{\mathrm{DW}}$.

Eq. 4.10 constitutes the first way to apply the KF with the vector $\mathbf{C}^{\mathrm{DW}}$ being the state vector. It is also possible to apply the KF in a second way with $\mathbf{C}^{\mathrm{ODF}}$ being the state vector. For that, we use eq. 3.17, which we rewrite here: $\mathbf{C}^{\mathrm{ODF}} = \mathbf{P}\mathbf{C}^{\mathrm{DW}}$. The KF model linear equation is then written as:

$$\boxed{\mathbf{M_E} = \left(\mathbf{B}\mathbf{P}^{-1}\right)\mathbf{C}^{\mathrm{ODF}} + \boldsymbol{\epsilon}.} \tag{4.11}$$

The noise vector is the same as previously.

To apply this second KF model, as well as the first one, we need to know the noise-free T_2-weighted signal S_0, as it appears in the observations' vector definition. As for the DTI Kalman framework, we can either approximate it directly by the measured T_2-weighted signal M_0 or apply a correction on M_0 and use the resulting corrected T_2-weighted signal $\hat{S}_0$ in replacement of S_0 in eq. 4.11. The results shown later in this chapter were obtained following the first solution.

In both cases, whether the state vector is $\mathbf{C}^{\mathrm{DW}}$ or $\mathbf{C}^{\mathrm{ODF}}$, the KF equations are derived by minimizing the cost function given by $(\mathbf{y} - \mathbf{A}\mathbf{x})^{\mathrm{T}}\,\mathbf{R}^{-1}\,(\mathbf{y} - \mathbf{A}\mathbf{x})$ (as done in appendix A (section 7.6)). But, it is also possible to apply the KF using the Tikhonov regularization

introduced in chapter 3, subsection 3.2.5. Concerning this, the approach in [Poupon et al. (2008b)] consisted to work from a linear model containing the regularization. From eq. 3.19, which we rewrite here: $\hat{\mathbf{C}}^{\mathrm{DW}} = (\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B} + \lambda \mathbf{L})^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot \mathbf{M_E}$, the authors of [Poupon et al. (2008b)] applied the Moore-Penrose pseudo-inversion operator and obtained:

$$\mathbf{M_E} = \mathbf{B}^{+} \cdot \mathbf{C}^{\mathrm{DW}} + \boldsymbol{\epsilon}', \tag{4.12}$$

with $\mathbf{B}^{+} = \left(\left(\mathbf{B^T} \cdot \mathbf{B} + \lambda \mathbf{L} \right)^{-1} \cdot \mathbf{B^T} \right)^{\dagger}$. The vector $\boldsymbol{\epsilon}'$ is, unlike $\boldsymbol{\epsilon}$, impacted by the regularization appearing in the design matrix through the term $\lambda \mathbf{L}$. However, as explained in [Deriche et al. (2009)], the k^{th} row of the design matrix defined as $\mathbf{B}^{+}$ depends on $\mathbf{B}$ and therefore on all K gradient acquisition directions. This means that, to work well, this system requires to know in advance the diffusion directions that will be used during the RT sequence. And yet, this is not the desired behavior for the RT project. Moreover, the authors of [Deriche et al. (2009)] showed that the KF model relying on eq. 4.12 is indeed suboptimal at every estimation step except for the last one. They proposed another way to inject the regularization in the KF. Their method is based on the incorporation of the regularization in the cost function, which becomes: $(\mathbf{y} - \mathbf{A}\mathbf{x})^{\mathrm{T}} \mathbf{R}^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x}) + \lambda \mathbf{x}^{\mathrm{T}} \mathbf{L} \mathbf{x}$. The latter expression leads to a derivation of the KF equations which is similar as in appendix A, unless it contains the regularization term. This derivation can be found in [Deriche et al. (2009)], which proves that the resulting KF equations are exactly equal to the system given by 4.2. The difference with the KF without regularization lies in the initialization equation of the estimation error covariance $\mathbf{P}$ which becomes: $\mathbf{P}_0 = ((1/V)\mathbf{I} + \lambda \mathbf{L})^{-1}$ with $V = 10^6$. Even though the difference between the KF equations without or with regularization appears in the initialization only, this change is reported in the KF equations' system given by 4.2 and therefore still impacts the whole KF process. The two possible equations on which we can apply the KF, without or with regularization, are resumed in table 4.2.

We will see in subsection 4.5.2 that this set up system achieves its goal.

Table 4.2: The two aQBI Kalman frameworks

<table>
<tr><td colspan="2">$\boxed{\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}}$, with for the iteration i:</td></tr>
<tr><td>with state vector $\mathbf{x} = \mathbf{C}^{\mathrm{DW}}$</td><td>with state vector $\mathbf{x} = \mathbf{C}^{\mathrm{ODF}}$</td></tr>
<tr><td>$\begin{cases} y_i = M_E(\mathbf{o}_i) \\ \mathbf{A} = \mathbf{B} \\ \mathbf{x} = \mathbf{C}^{\mathrm{DW}} \\ \epsilon_i \text{ supposed to follow } \mathcal{N}(0, \mathbf{R}) \end{cases}$</td><td>$\begin{cases} y_i = M_E(\mathbf{o}_i) \\ \mathbf{A} = \mathbf{B} \cdot \mathbf{P}^{-1} \\ \mathbf{x} = \mathbf{C}^{\mathrm{ODF}} \\ \epsilon_i \text{ supposed to follow } \mathcal{N}(0, \mathbf{R}) \end{cases}$</td></tr>
<tr><td colspan="2">Definition of the noise covariance matrix $\mathbf{R}$
$\mathbf{R}$ is diagonal and defined such that:
- either $\mathbf{R} = \mathbf{I}$ following [Poupon et al. (2008b)],
- or $\mathbf{R} = R_i\mathbf{I}$ with $R_i \approx \sigma_i^2/S_0^2$.</td></tr>
<tr><td colspan="2">Initialization for the KF equations' system 4.2:
without regularization
$\begin{cases} \mathbf{x}_0 = \mathbf{0} \\ \mathbf{P}_0 = V\mathbf{I} \text{ with } V = 10^6. \end{cases}$
with regularization
$\begin{cases} \mathbf{x}_0 = \mathbf{0} \\ \mathbf{P}_0 = ((1/V)\mathbf{I} + \lambda\mathbf{L})^{-1} \text{ with } V = 10^6 \text{ ([Deriche et al. (2009)])}. \end{cases}$</td></tr>
<tr><td colspan="2">The equations' system given by 4.2 can be applied on the models described in this table.</td></tr>
</table>

sa-aQBI Kalman framework

Concerning sa-aQBI, it is also interesting to make it KF-feasable with a linear model. Let us remind that the sa-aQBI relies on the following system:

$$\begin{cases} \mathbf{C}^{\mathrm{ODF}_c}(j) = \frac{1}{2\sqrt{\pi}} & \text{if } j = 1, \\ \mathbf{C}^{\mathrm{ODF}_c}(j) = \frac{1}{16\pi^2}\mathbf{P} \cdot \mathbf{N} \cdot (\mathbf{B}^{\mathrm{T}} \cdot \mathbf{B} + \lambda\mathbf{L})^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot ln\left((-ln\mathbf{E})\right) & \text{if } j > 1, \end{cases} \tag{4.13}$$

with the K×1 vector $\mathbf{E}$ defined such that its ith element, corresponding to the diffusion orientation $\mathbf{o}_i$, is $E(\mathbf{o}_i) = S(\mathbf{o}_i)/S_0$. All other notations in 4.13 are given in subsection 3.2.5. Let us remind the notation $\mathbf{M_E}$ defined such that $M_E(\mathbf{o}_i) = M(\mathbf{o}_i)/S_0$. The sa-aQBI adapted linear model for applying the KF was first introduced in [Caruyer et al.

(2010)]. It is expressed such that:

$$\boxed{ln\left(-ln\mathbf{M_E}\right) = \mathbf{B}\cdot\mathbf{C}^{\mathrm{SA}} + \boldsymbol{\epsilon}}, \tag{4.14}$$

where $\mathbf{C}^{\mathrm{SA}}$ is the coefficients' vector of the decomposition of $ln\left(-ln\mathbf{E}\right)$ on the modified SH basis. $\boldsymbol{\epsilon}$ is the vector containing $ln\left(-ln\mu_i\right)/S_0\ \forall i \in [\![1;K]\!]$, with μ_i the acquisition noise vector at iteration i. As for the previous models, we approximate the noise PDF by a zero-mean Gaussian distribution, as required by the KF model constraints. We can either work with a sequential OLS estimator, *i.e.* $\mathbf{R} = \mathbf{I}$, or work with a sequential WLS estimator with $\mathbf{R} = R_i\mathbf{I}$ where $R_i \approx \frac{\sigma_i^2}{S(\mathbf{o}_i)^2 ln(S(\mathbf{o}_i)/S_0)^2}$ ([Caruyer et al. (2011)]). The expression of R_i can be explained using the propagation of uncertainty and using the approximation of the nc-χ distribution followed by $M(\mathbf{o}_i)$ by a Gaussian PDF of variance σ_i^2, as it was done for the DTI adapted linear model. Again, as for some previous models, to apply this KF model, we need to know the noise-free T_2-weighted signal S_0 to be able to compute $\mathbf{M_E}$. As before, we can either approximate it directly by the measured T_2-weighted signal M_0 or apply a correction on M_0 and use the resulting corrected T_2-weighted signal $\hat{S}_0$ in replacement of S_0 in eq. 4.14. The results shown later in this chapter were obtained following the first solution, as the authors of [Caruyer et al. (2011)] did. Finally, as for the aQBI, regularization can be incorporated in the KF model. The difference between the non-regularized and the regularized models only comes from the initialization, as resumed in table 4.3. From the estimated state vector $\hat{\mathbf{C}}^{\mathrm{SA}}$, it is then possible to estimate the coefficients of the cdODF $\hat{\mathbf{C}}^{\mathrm{ODF}_c}$, such that:

$$\begin{cases} \hat{\mathbf{C}}^{\mathrm{ODF}_c}(j) = \frac{1}{2\sqrt{\pi}} & \text{if } j = 1, \\ \hat{\mathbf{C}}^{\mathrm{ODF}_c}(j) = \frac{1}{16\pi^2}\mathbf{P}\cdot\mathbf{N}\cdot\hat{\mathbf{C}}^{\mathrm{SA}}(j) & \text{if } j > 1. \end{cases} \tag{4.15}$$

With this sa-aQBI adapted linear model, we come to the end of the section on the setting up of linear models adapted for rtdMRI. Among the DTI, aQBI and sa-aQBI linear models adapted to a KF application, we remark that only the aQBI proposed models incorporate the authentic acquisition noise: the noise vector $\boldsymbol{\epsilon}$ is then only constituted by the acquisition noise. On the contrary, for the DTI linear model, the noise vector $\boldsymbol{\epsilon}$ is defined such that: $\epsilon_i = -ln\left(1 + \frac{\mu_i}{S_0}e^{b_i\mathbf{o}_i^T\mathbf{D}\mathbf{o}_i}\right)$ for iteration i. Concerning the sa-aQBI linear model, it is defined such that: $\epsilon_i = ln\left(-ln\mu_i\right)/S_0$ for iteration i. This means that for the DTI and sa-aQBI models, the knowledge on the acquisition noise has to be adapted to the noise term ϵ_i of the corresponding model. Especially the PDF followed by ϵ_i does not correspond to the one followed by the acquisition noise and has to be recalculated. For this chapter, this specific details are not of much importance as we will approximate in all three cases the noise PDF by a zero-mean Gaussian distribution. But this information about noise will be used when incorporating adapted noise correction in the KF process (see chapter 7).

To concretely apply these theoretical KF-feasable models, it is necessary to have the practical set up adapted to the RT objective. The next section concerns other aspects, which have to be considered, so that the RT project works well.

Table 4.3: The sa-aQBI Kalman framework

$\boxed{\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}}$, with for iteration i:

$$\begin{cases} y_i = ln\left(-ln\frac{M(\mathbf{o}_i)}{S_0}\right) \\ \mathbf{A} = \mathbf{B} \\ \mathbf{x} = \mathbf{C}^{\mathrm{SA}} \\ \epsilon_i \text{ supposed to follow } \mathcal{N}(0, \mathbf{R}) \end{cases}$$

Definition of the noise covariance matrix $\mathbf{R}$:

$\mathbf{R}$ is diagonal, such that:
- either $\mathbf{R} = \mathbf{I}$ following [Poupon et al. (2008b)],
- or $\mathbf{R} = R_i\mathbf{I}$ with $R_i \approx \frac{\sigma_i^2}{S(\mathbf{o}_i)^2 ln(S(\mathbf{o}_i)/S_0)^2}$ following [Caruyer et al. (2011)]

Initialization for the KF equations' system 4.2:

without regularization

$$\begin{cases} \mathbf{x}_0 = \mathbf{0} \\ \mathbf{P}_0 = V\mathbf{I} \text{ with } V = 10^6. \end{cases}$$

with regularization

$$\begin{cases} \mathbf{x}_0 = \mathbf{0} \\ \mathbf{P}_0 = ((1/V)\mathbf{I} + \lambda\mathbf{L})^{-1} \text{ with } V = 10^6 \text{ ([Deriche et al. (2009)])}. \end{cases}$$

The equations' system given by 4.2 can be applied on the model described in this table.

4.4 RT constraints

This section begins with the explanation of the necessity to optimize the diffusion gradients' orientation set with the goal of performing an optimized rtdMRI sequence from an RT point of view. Then, the section is focused on hardware considerations and issues of the RT project.

4.4.1 Otimization of the diffusion gradients' orientation set

This section focuses on the diffusion gradients' orientation set used during a rtdMRI sequence. When performing dMRI, the diffusion of the water molecules is probed along several orientations in the space, using various diffusion gradients in the dMRI sequence. Thus, during the sequence, each new measurement brings new information about the dif-

fusion 3D geometry. The orientation set contains all orientations, for which a diffusion measurement is performed using dMRI. It would be of course ideal to measure diffusion in all the possible directions, but this is not realistic, because of the too long acquisition time. Concretely, the best way to optimally cover the angular possibilities seems to have a uniform orientation set. Nevertheless, as the diffusion process is symmetric, both orientations $\mathbf{o}_i$ and $-\mathbf{o}_i$ generate the same diffusion result. Following this analysis, the authors of [Jones et al. (1999)] proposed an optimized orientation set based on a slightly modified electrostatic repulsion model. In their model, each gradient direction is represented by an antipodal pair of identical electric unit charges to precisely account for the diffusion process symmetry. The charges are placed on the surface of a sphere and forced to stay on it. Each pair is allowed to pivot and all pairs repulse each other. The optimized —in the sense of most uniform— set of diffusion gradients is then found when minimizing the systems' energy defined as the sum of the electrostatic-like repulsive potentials between every pair of charges. This global energy, for K orientations played during the sequence, is written as:

$$E_K(\mathbf{o}_1, ..., \mathbf{o}_K) = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} E^0(\mathbf{o}_i, \mathbf{o}_j), \tag{4.16}$$

with

$$E^0(\mathbf{o}_i, \mathbf{o}_j) = \frac{1}{\|\mathbf{o}_i - \mathbf{o}_j\|} + \frac{1}{\|\mathbf{o}_i + \mathbf{o}_j\|} \tag{4.17}$$

being the electrostatic-like potential representing the energy between the two orientations $\mathbf{o}_i$ and $\mathbf{o}_j$. This model is easy to solve for particular numbers of orientations. The solution for six orientations is shown in fig. 4.4 with a regular icosahedron. But, when such a perfect solution is not known, the optimization algorithm by Derek Jones *et al.* can be run to find the orientations' configuration yielding the minimum of energy based on some criterion. Another approach by [Papadakis et al. (2000)] has been proposed to find the optimal orientations' configuration. It is also based on the electrostatic repulsion model and determines the optimized set when maximizing the distance between any pair.

Concerning the particular case of rtdMRI, for which the sequence can be stopped at any time if the user decides it, the chronological order, in which the orientations are used, is of paramount relevance. Indeed, along which direction should diffusion be measured at the beginning, and then later until the end of the exam, so that the RT orientation set is roughly uniform, should the acquisition be finished before completion? Several works have answered this question. The publication of [Poupon et al. (2008b)] presents the method by [Dubois et al. (2006)]. The authors of [Deriche et al. (2009)] compared their own approach with the techniques by Jessica Dubois *et al.* and [Cook et al. (2007)]. All methods were shown to be quite equivalent and we invite the reader to read the article by [Deriche et al. (2009)] for more details on the several methods and their comparison. As we performed RT experiments for this thesis with the technique by Jessica Dubois *et al.*, we will only describe the latter.

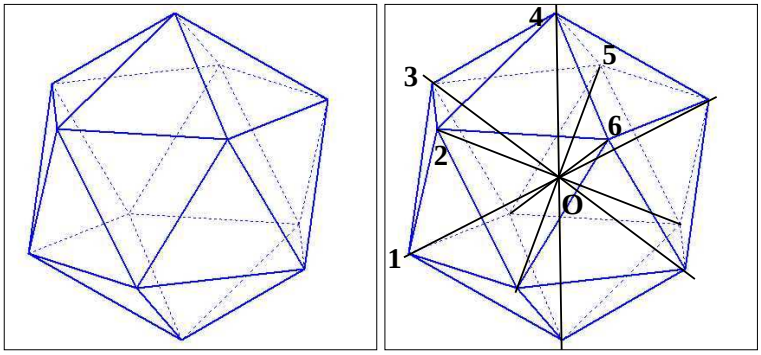


Figure 4.4: Optimized orientation set with six orientations. The same regular icosahedron is shown on the left and on the right, but the six optimal diffusion are added on the right. The set is obtained strictly uniform. Adapted fig. from http://www.ilemaths.net/img/forum_img/0271/forum_271542_1.jpg.

The idea of the optimization solution of [Dubois et al. (2006)] is not only to create a final uniform orientation set whith all orientations planned to be played during the sequence, but also to generate ordered **time-intermediate subsets** of orientations which are **approximately uniform**. Consequently, if the exam is stopped before completion, even if all orientations have not been played, the resulting played orientation set will be approximately uniform. To realize this idea, the authors of [Dubois et al. (2006)] defined the repulsive potential between two orientations $\mathbf{o}_i$ and $\mathbf{o}_j$, such that:

$$E(\mathbf{o}_i, \mathbf{o}_j) = \alpha_{ij} E^0(\mathbf{o}_i, \mathbf{o}_j), \tag{4.18}$$

with α_{ij} representing the interaction coefficient between the two orientations. The global energy to minimize therefore becomes:

$$E_K(\mathbf{o}_1, ..., \mathbf{o}_K) = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \alpha_{ij} E^0(\mathbf{o}_i, \mathbf{o}_j), \tag{4.19}$$

This parameter α_{ij} modulates the interaction between $\mathbf{o}_i$ and $\mathbf{o}_j$, depending if they are far or not from the acquisition beginning. In the model by [Dubois et al. (2006)], orientation subsets are defined such that: a subset $(S+1)$ includes the previous subset (S). Orientations belonging together to the smallest subset are the most constrained to be uniform and their interaction is maximal: $\alpha_{ij} = 1$. On the other hand, if $\mathbf{o}_i$ and $\mathbf{o}_j$ have in common to belong together only to the last subset, then their interaction is minimum, with $\alpha_{ij} = \alpha_{\min}$. In [Poupon et al. (2008b)], $\alpha_{\min}$ is set to 0.5. Consequently, minimizing eq. 4.19 favours approximately uniform time-intermediate subsets. However, this criterium is different from the minimization of eq. 4.16. Therefore, one risk of the method is that it will not generate the optimal uniform solution for the final global orientation set. But a suboptimal solution may be sufficient.

To better understand the effects of this algorithm, let us take an example of a total set of 42 orientations, illustrated in fig. 4.5. The first subset is defined such that it contains the

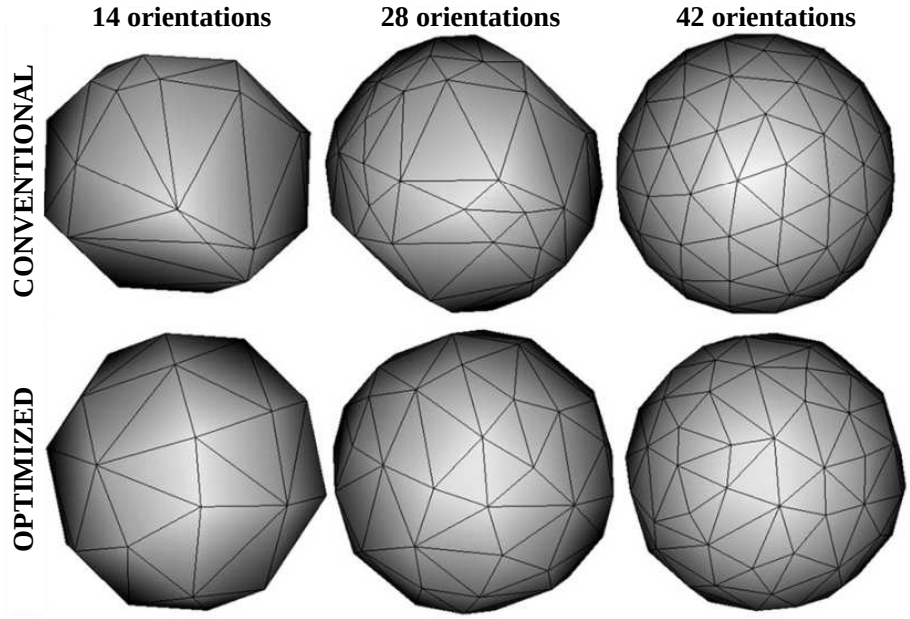


Figure 4.5: Comparison between the conventional (top line) and optimized (bottom line) orientation subsets of a sequence with 42 orientations played in total. The optimized intermediate subsets (at 14 and 28 orientations played) are more uniform than the corresponding conventional subsets. Consequently, the optimized subsets have to be used for RT scans, which may be stopped at any time. Fig. extracted from [Poupon et al. (2008b)].

first 14 orientations played. The second subset contains the first subset with in addition 14 other more orientations played. Finally the last subset is equal to the total set with the 42 orientations. A high constraint is put on the first 14 orientations so that the first subset is uniform. The constraint will then be a bit released for the next orientations coming (second subset), and even more for the last orientations (third subset), but still sufficient to build a final total set of 42 orientations which is roughly uniform. Fig. 4.5 clearly shows that the intermediate orientation sets are more uniform with the optimized solution of [Dubois et al. (2006)], than with the conventional method. However, the final set is more uniform with the latter method. Nevertheless, from a RT perspective as explained at the beginning of this subsection, the optimized solution is much more desirable.

The next subsection tackles the hardware architecture set up for the RT project.

4.4.2 RT hardware architecture

To run an MRI sequence and perform the data processing in RT, the hardware architecture has to be well organized.

Before tackling the RT adapted hardware architecture, let us begin with the presentation of the simpler hardware organization when there is no RT goal. When a classical MRI sequence is played in the MRI scanner, the acquired data follow the flow indicated by the black arrows in fig. 4.6: the data first go through the reconstruction unit (delivered

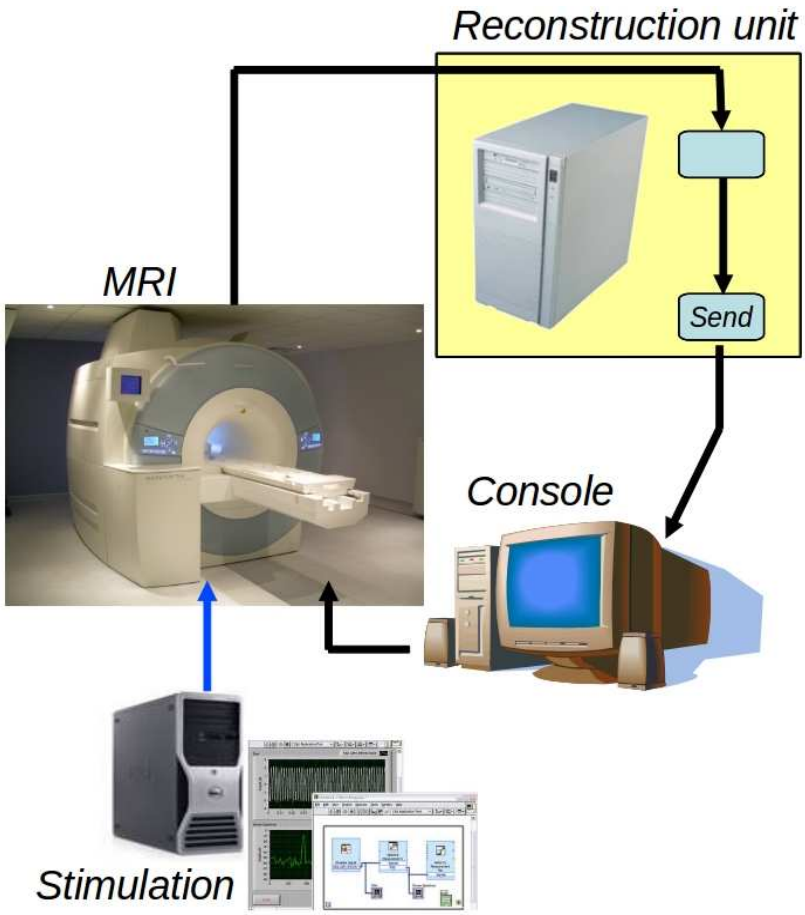


Figure 4.6: The classical hardware architecture with no RT goal. The MRI machine acquires data, which are sent to the manufacturer reconstruction unit for some rapid processing. Then, the data are sent to the console for visualization by the MRI technologist. Stimuli can be used in case of an fMRI sequence. Fig. adapted from [Poupon and Riff (2009)].

by the manufacturer with the magnet itself). This reconstruction unit performs several processings such as parallel reconstruction, retrospective motion correction that remains limited to simple tasks due to the limited number of processing units available (8 CPUs on the Tim Trio). Such processing tasks can be of course deactivated by the user, if wanted. Then, the unit sends the data to the console, on which the radiologic technologist visualizes them and considers if they are of good quality or not. Depending on his decision, he will command the start of the same sequence or of another MRI sequence on the magnet. Also, if it is fMRI, the MRI technologist also commands the start of stimuli. With this system, the data are acquired, then processed and finally stored to files. Once the exam is completed, these data can be post-processed on the operator console using the available tools provided by the manufacturer. Clearly, the post-processing of dMRI cannot be done online and there is no possibility to easily develop custom tools to be installed on the operator console. Therefore, the information technology architecture has to be revisited to allow high performance RT computing, such as the RT Kalman frameworks introduced

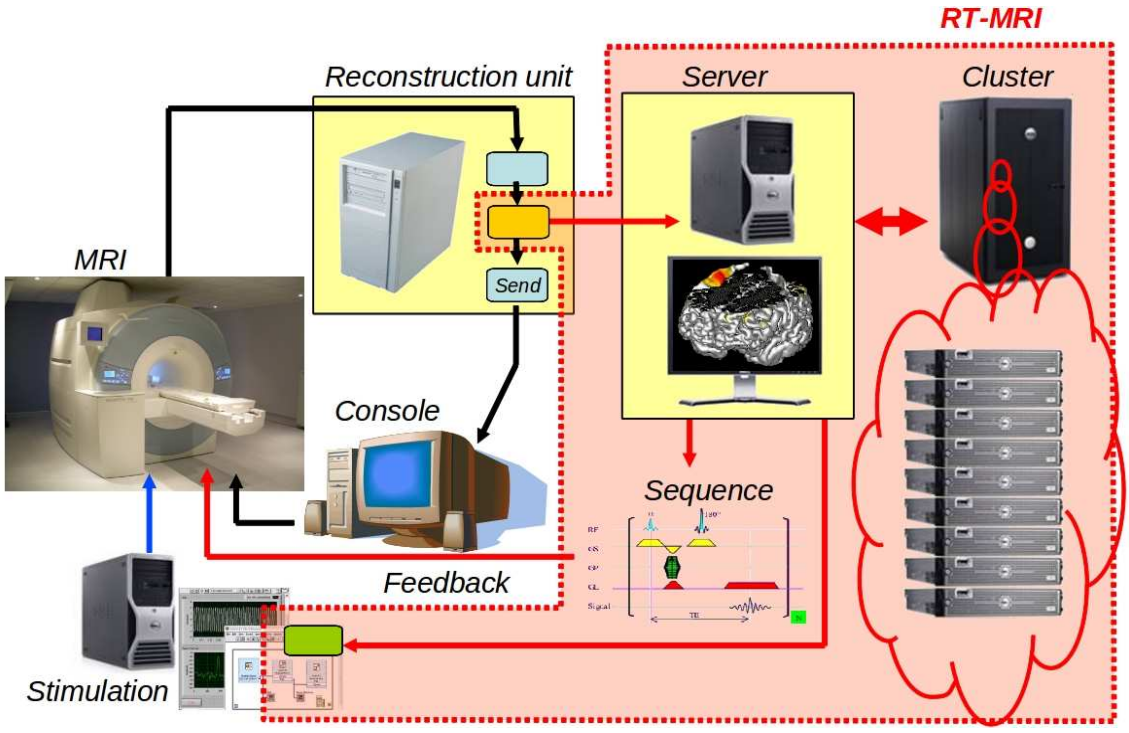


Figure 4.7: The hardware architecture adapted to the RT project. In addition to the system shown in fig. 4.6, the red block "RT-MRI" constitutes the hardware part enabling to perform RT MRI. Before the reconstruction unit sends the data to the console, these data are sent to a server, which commands the running of algorithms on the data. To reduce the processing time, these algorithms are parallelized and distributed on a cluster of 80 CPUs. In addition to this RT processing, some automatic feedback can be performed on the sequence parameters, as well as on the stimuli controle console in the case of RT fMRI. Fig. extracted from [Poupon and Riff (2009)].

in subsection 4.3.

This RT environment added to the MRI system is the red block "RT-MRI" of fig. 4.7 that was set up by Fabrice Poupon. It consists of a further loop that allows to send the data to a custom high performance computing architecture, to process them efficiently and to put the result back to the operator console. In order to meet the computational requirements, a **cluster** of 80 CPUs is used to take benefits from parallel and distributed implementations of post-processings on this cluster, yielding a huge reduction of the processing time. For the processing tasks such as noise removal presented later in the manuscript, the processing time was always obtained far below the T_R. The RT-processed data are then sent to the console for rendering during the exam.

Fig. 4.7 also shows the feedback possibilities allowed by the RT hardware system. From the console visualization, the user himself can decide to zoom in a region, to stop the acquisistion and play another sequence... This feedback is non-automatic. But there can also be automatic feedbacks, which can be decided at the server level: once the RT-processing is performed, the server can send modification commands concerning the MRI

sequence played in the magnet. For instance, the gradients can be modified with respect to a certain parameter in the results obtained at the iteration played. Concerning fMRI sequences, the stimuli can also be automatically adapted according to the results obtained.

This entire hardware system is innovative, as well as compatible with a clinical use. To perform methodological tests on the RT algorithms independently of an acquisition with a subject in the magnet, an RT MRI simulator was developed which mimics an MR acquisition from an already acquired dataset.

The next section will show the first results obtained with rtdMRI on real data. These results constitute a reference in this thesis, as we will often compare our results in part III of the thesis to these ones.

4.5 Proof of concept: RT local modeling

Here we show the first results obtained with rtdMRI on real data at $b = 3000s \cdot mm^{-2}$. For more details on the acquisition and preprocessing parameters for these data, see appendix A. In this section, we did not account for any noise information in the KF models: neither did we estimate a noise variance, nor did we estimate the noise-free DW signals which are necessary for the calculation of the covariance matrix $\mathbf{R}$ in the WLS cases. These results are therefore based on the OLS versions of the algorithms presented in section 4.3. We show the DTI results, as well as the aQBI and sa-aQBI results.

4.5.1 RT application for DTI

In this section, we show in fig. 4.8 some DTI results, namely the ADC, FA and RGB maps, obtained in RT at the 6^{th}, 10^{th}, 19^{th}, 39^{th} and 60^{th} iterations. The 19^{th}, 39^{th} and 60^{th} iterations correspond to each last iteration of the three subsets of orientations using the optimized algorithm of [Dubois et al. (2006)]. Therefore, at these three iteration indexes, the orientations used are almost uniformly distributed and the maps contain structural information, which was homogeneously depicted. As we can see in fig. 4.8, the maps, particularly the FA and RGB ones, are already of good quality at the 39^{th} iteration. No highly significant change in the maps is noticed after this 39^{th} iteration. We also see that the maps obtained at the last iteration are the same as the ones obtained with the post-processing algorithms.

The KF algorithm evolution in RT can also be analyzed in a quantitative way using the trace, at each iteration i, of the estimation error covariance matrix P_i. This quantity is minimized during the KF process and thus indicates its progress, as pointed out in section 4.2. Let us here notice that this trace is the same at each voxel, as the KF process is independent from any spatial consideration. Fig. 4.9 shows the evolution of this trace at one voxel during the RT process. We started the graph only at the 6^{th} iteration, as the diffusion tensor requires at least six measurements to be estimated. The curve significantly decreases at the beginning of the process and reaches a *plateau* around the 40^{th} iteration. This is consistent with our qualitative analysis on the images in fig. 4.8.

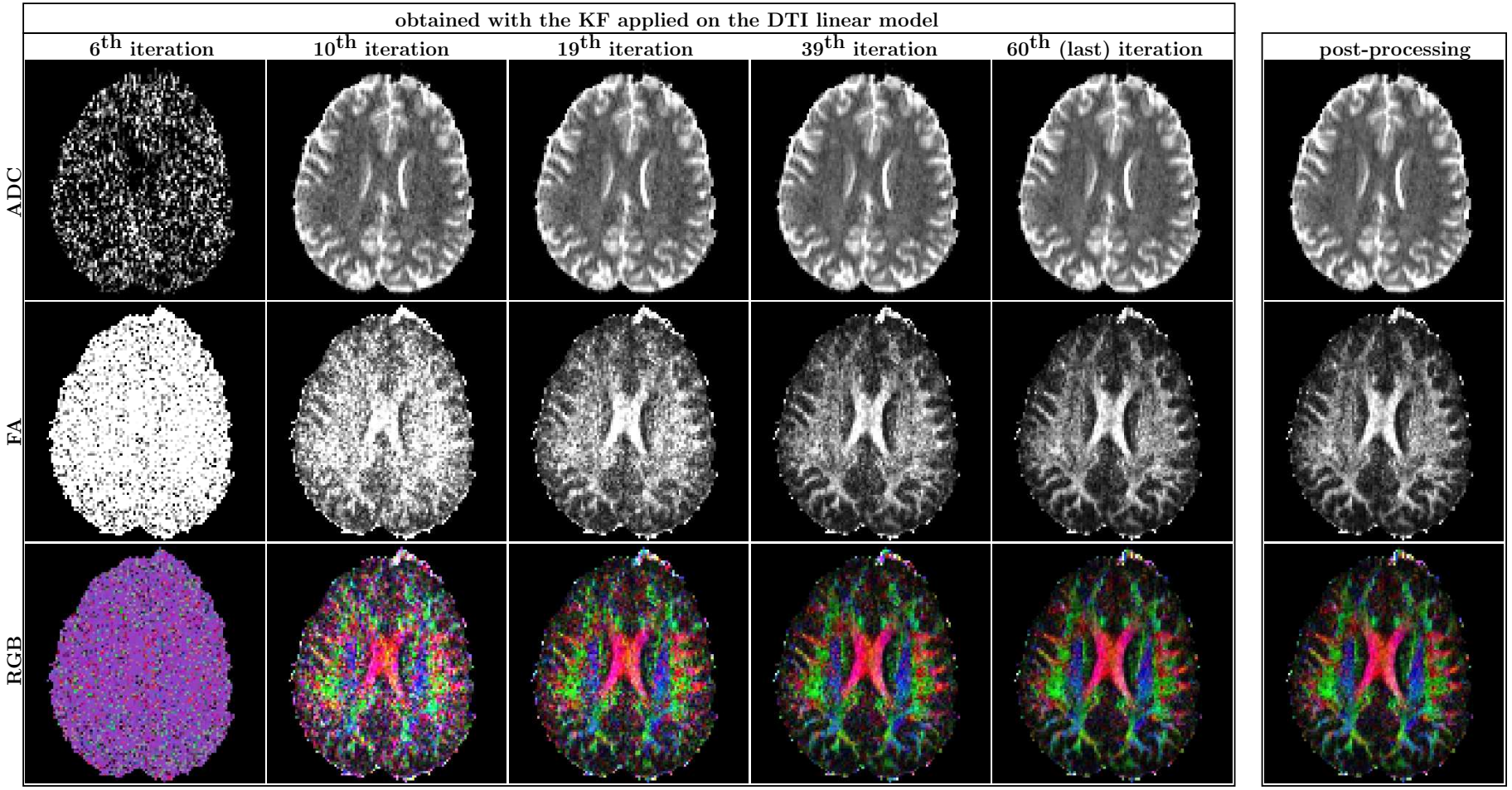


Figure 4.8: KF application on the DTI model. The KF was performed as an OLS estimator.

Fig. 4.9 also depicts the evolution of the mean square error (MSE) between the final offline estimation of the state vector and its RT estimation. Here, $\hat{\mathbf{x}}$ represents the six estimated coefficients of the diffusion tensor. The MSE_i is then calculated on the whole 3D volume at each iteration, such that:

$$\boxed{\text{MSE}_i = \frac{1}{N_{\mathbf{v}}} \sum_{\mathbf{v} \in \mathcal{V}} \sum_{j=1}^{N} \left(\hat{x}_i(j, \mathbf{v}) - x^{offline}(j, \mathbf{v}) \right)^2}, \tag{4.20}$$

with $\mathcal{V}$ the entire 3D image volume containing $N_{\mathbf{v}}$ voxels. $\hat{x}_i(j, \mathbf{v})$ is the RT estimation at iteration i of the j^{th} coefficient of the state vector for the voxel $\mathbf{v}$. $x^{offline}(j, \mathbf{v})$ corresponds to the offline estimated j^{th} coefficient of the state vector for the voxel $\mathbf{v}$. Here, the state vector is the diffusion tensor, therefore $N = 6$. The curve representing the MSE_i in fig 4.9 presents a similar behavior as the curve of the trace of $\mathbf{P}_i$.

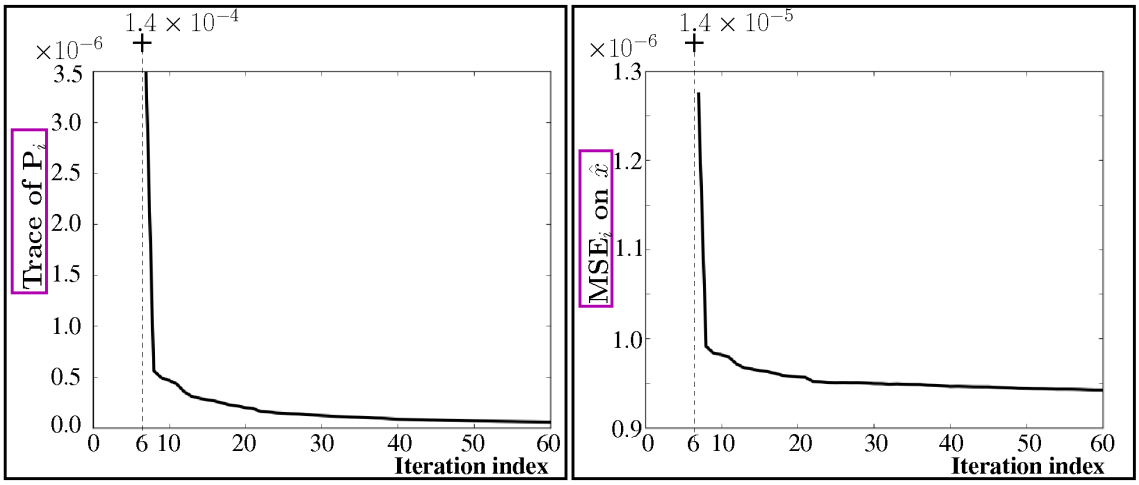


Figure 4.9: Left: evolution of the trace of P_i during the KF process. The trace of P_i, written as $Tr\,(P_i)$, is initialized such that $Tr\,(P_0) = 6.0 \times 10^6$. Right: evolution of MSE_i during the KF process. MSE_i is computed using eq. 4.20.

Finally, we also quantified the performance of this RT process in terms of time duration. We distinguished two times: first, the time t_{KF} required for a KF update, and secondly, the time t_{map} required to process the the FA, ADC and RGB maps. The time to display the maps in RT was negligible compared to these other times, therefore, we will not indicate it. We run the algorithms over the brain inside the 3D volume on a linux 2.8 GHz workstation. The times of the RT process are shown in fig. 4.10 when using only one CPU and when using the cluster of 80 CPUs (each CPU is a linux 2 GHz machine). It is visible that the parallelization and distribution of the C++ code caused a huge decrease of the global computation time $(t_{KF} + t_{map})$ for one iteration of the process. Whereas this time is close to $T_R = 14s$ on one CPU, it then falls far below T_R with the cluster. The computation process is performed in RT, as the update after one iteration is done before the next iteration happens.

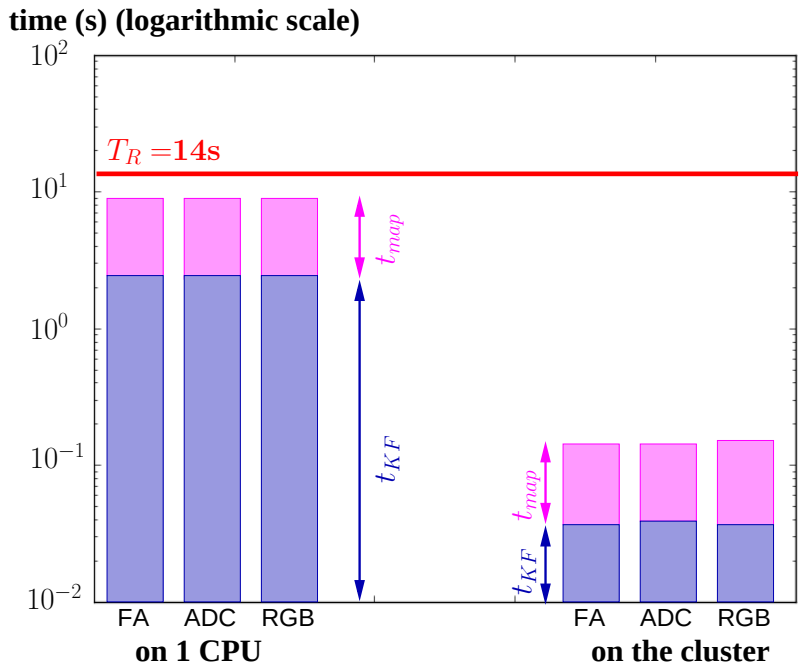


Figure 4.10: Times to get the FA, ADC and RGB maps' update after a new iteration: t_{KF} (blue box) is the time required for one iteration of the KF algorithm and t_{map} (magenta box) is the time for processing the maps. These times are compared to $T_R = 14s$.

4.5.2 RT application for aQBI and sa-aQBI

Here we show the results obtained applying the OLS version of the regularized KF on the aQBI and sa-aQBI linear models with a maximum SH order of 4 —meaning that the size of the state vector is equal to $N = 15$—. Fig. 4.11 depicts the RT aQBI maps (GFA and dODF) and the RT sa-aQBI maps, for which the corrected GFA (cGFA) is calculated using eq. 3.21 with the dODF coefficients replaced by the cdODF coefficients. As for the DTI results, fig. 4.11 highlights how the maps are refined in time. As before, no highly significant change in the GFA maps is noticed after the 39^{th} iteration. Nevertheless, some changes are more visible on the dODF and cdODF maps after the 39^{th} iteration. All maps obtained at the last iteration are the same as the ones obtained with the post-processing algorithms.

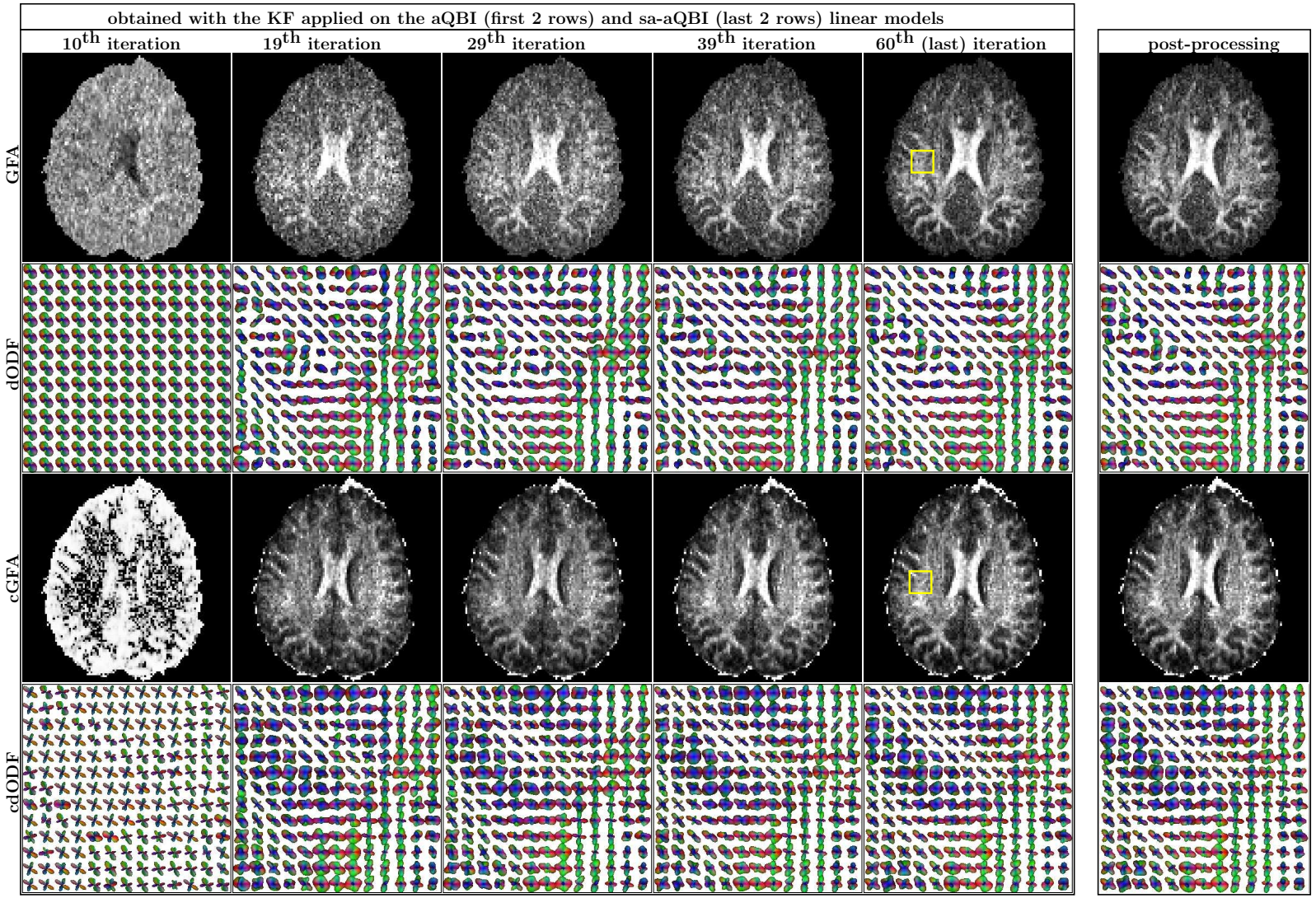


Figure 4.11: KF application on the aQBI and sa-aQBI models. The KF was performed as an OLS estimator, but using the regularized version. The maximum SH order was set to 4. The ROI for which we show the dODF and cdODF results is shown in yellow in the fourth column.

We here again calculated the trace of P_i at each iteration i, as well as the MSE_i defined by eq. 4.20. Here, the state vector was either the vector $\mathbf{C}^{\text{DW}}$ for the aQBI linear model or the vector $\mathbf{C}^{\text{SA}}$ for the sa-aQBI linear model. All curves (fig. 4.12) show a decrease of the error in RT. Note that the orders of magnitude of the MSE_i for the DTI, aQBI and sa-aQBI models are not the same as the corresponding state vectors are of different orders of magnitude (for the DTI, the first (and usually higher) coefficient of $\mathbf{D}$ is around 10^{-10}; for the aQBI, the first (and usually higher) coefficient of $\mathbf{C}^{\text{DW}}$ is around 10^{2}; for the sa-aQBI, the first (and usually higher) coefficient of$\mathbf{C}^{\text{SA}}$ is around 10^{-1}). The decrease for MSE_i does here not reach a *plateau* as it did previously for the linear DTI model. This confirms the different choice for the number of orientations (or here iterations) usually used for DTI and aQBI or sa-aQBI models: whereas for DTI, 40 orientations will be a very sufficient choice, for the aQBI and sa-aQBI models, it is better to take between 50 to 200 orientations [Deriche et al. (2009)], meaning that the improvement gained will still be significant after the 50^{th} orientation.

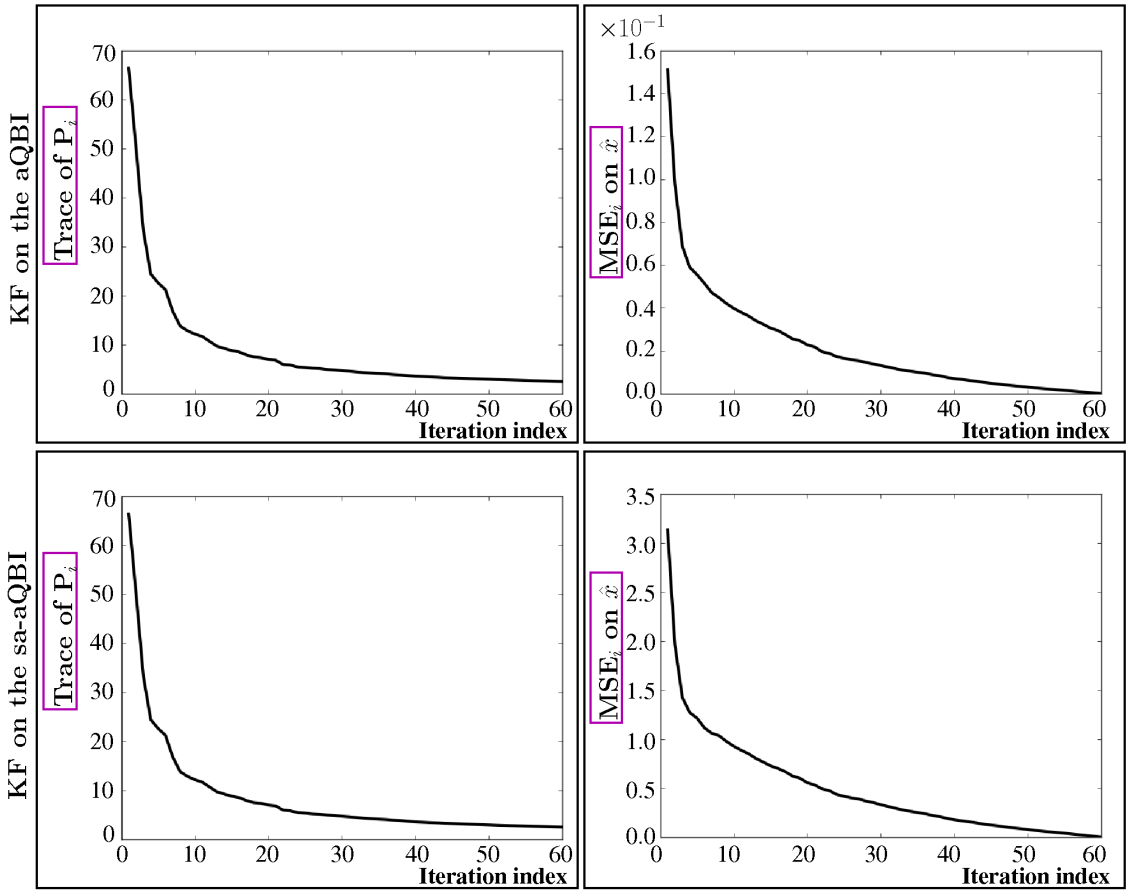


Figure 4.12: MSEs during the KF process on the aQBI and sa-aQBI model. The trace of P_i is initialized such that $Tr\,(P_0) = 10^6$ (regularized KF version). The results are the same whether the maximum SH order was set to 4 or 8.

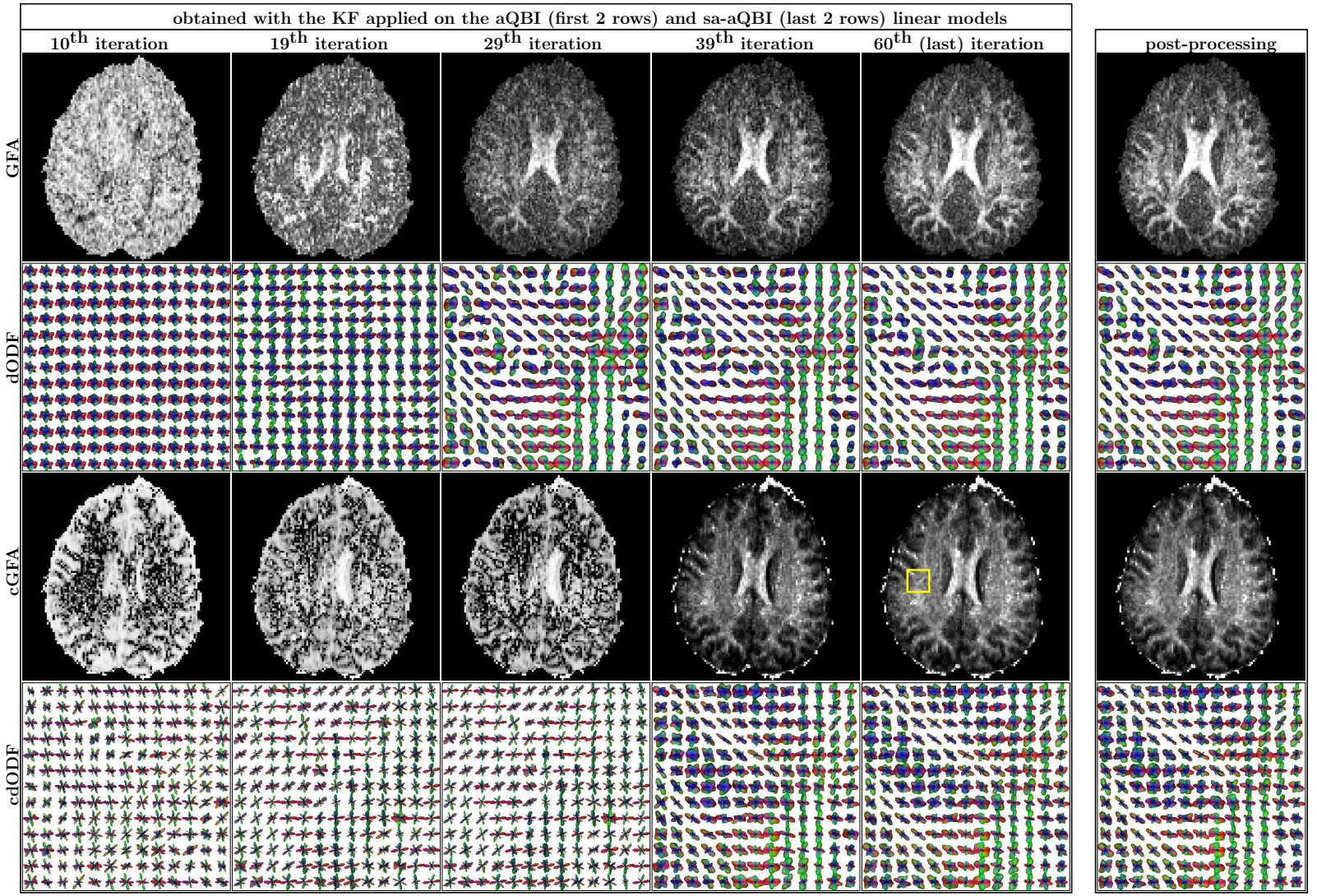


Figure 4.13: KF application on the aQBI and sa-aQBI models. The KF was performed as an OLS estimator, but using the regularized version. The maximum SH order was set to 6. The ROI for which we show the dODF and cdODF results is shown in yellow in the fourth column.

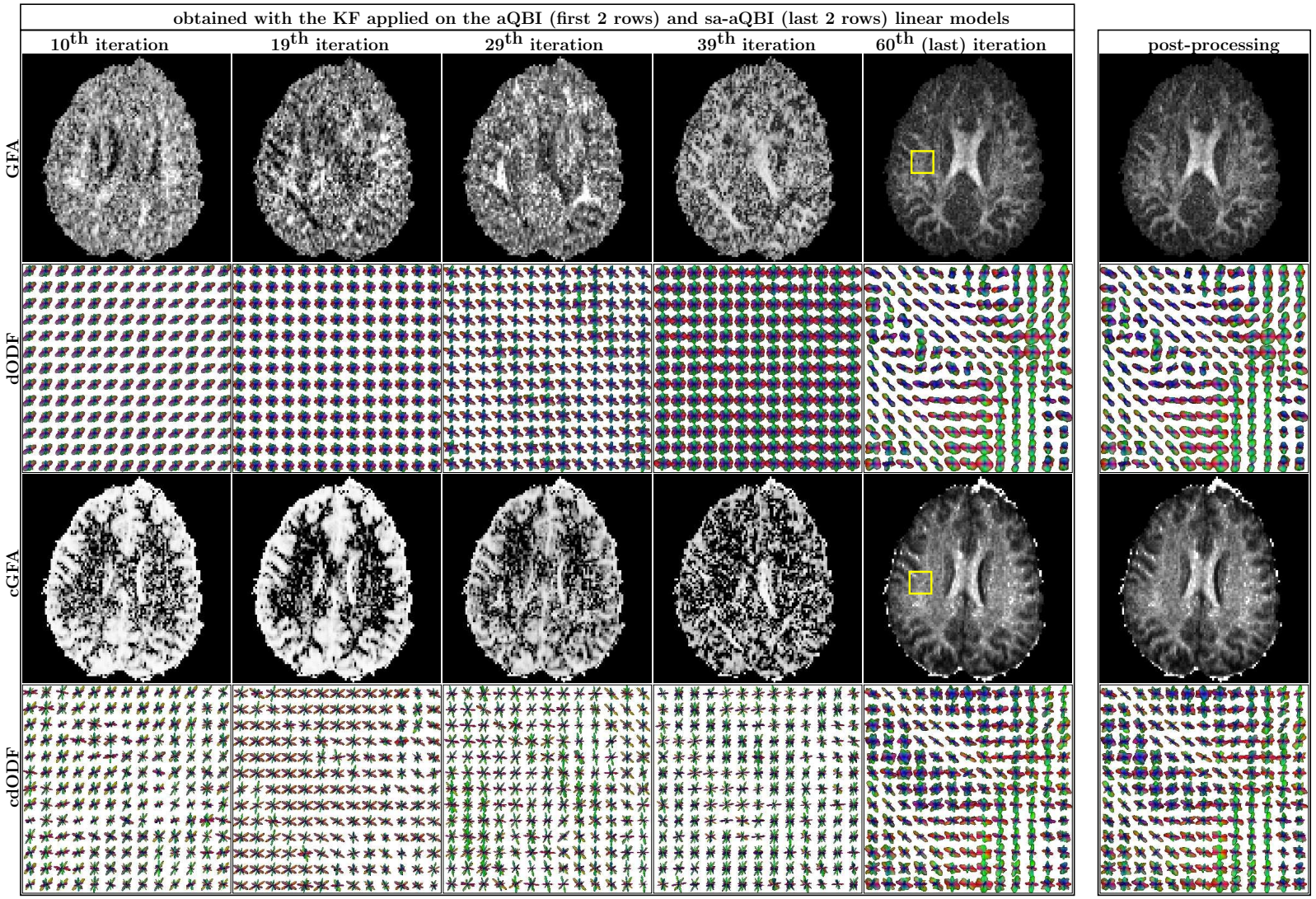


Figure 4.14: KF application on the aQBI and sa-aQBI models. The KF was performed as an OLS estimator, but using the regularized version. The maximum SH order was set to 8. The ROI for which we show the dODF and cdODF results is shown in yellow in the fourth column.

We did the same experiment with a maximum SH order of 6 and 8 —meaning that the size of the state vector is equal to $N = 28$ and $N = 45$, respectively—. This choice enables to get more details in the state vector, but also increases the noise impact on this vector. Indeed the SH decomposition of the DW signal on a SH basis acts a bit as a smoothing filter regarding noise issues. Nevertheless, this is less and less true, as N increases. This observation gives an explanation for the results in fig. 4.13 and fig. 4.14: the maps show no clear structural information for the first iterations (until at least the 39^{th} iteration for fig. 4.14. In this case, an RT process does not bring any extra information for at least the first 39^{th} iterations. We will be interested in showing if a noise correction enables to improve this situation in the chapter 7.

We also computed the evolutions of the trace of P_i at each iteration i, as well as of the MSE_i. We do not show them as they are identical to the ones in fig. 4.12. This shows that these RT error indicators have to be analyzed with respect to the maximum SH order used: if the latter is high (*e.g.* equal to 8), then these RT error indicators do no indicate the same image quality as when the maximum SH order is low.

Finally, we measured the two durations t_{KF} and t_{map} defined in the previous subsection. For these HARDI results, t_{map} represents the time required to process the GFA map for the aQBI model (or the cGFA for the sa-aQBI model). As before, we do not indicate the duration to display the maps in RT, since it is negligible compared to the other durations. The results are shown in fig. 4.15 for the GFA map and cGFA map reconstructions when using only one CPU and when using the cluster of 80 CPUs. For the maximum SH orders set to 6 and 8, the update and map reconstruction time was obtained higher than the repetition time T_R when using the workstation only (1 CPU). With the cluster, this time decreases far below T_R. The results are very similar when using the aQBI or the sa-aQBI model.

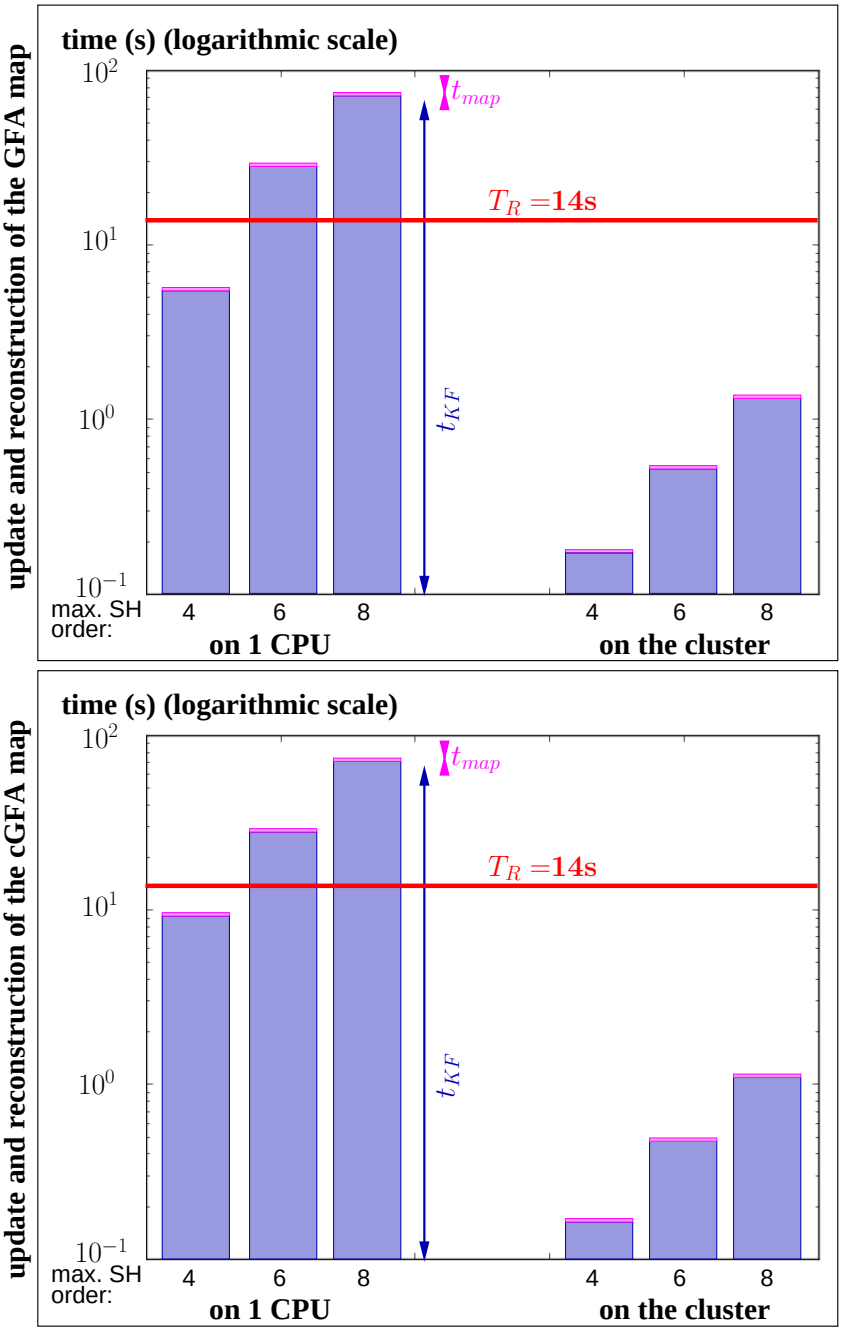


Figure 4.15: Times to get the GFA map and cGFA map update after a new iteration: t_{KF} (blue box) is the time required for one iteration of the KF algorithm and t_{map} (magenta box) is the time for processing the GFA map for the aQBI model (top) and the cGFA for the sa-aQBI model (bottom). These times are compared to $T_R = 14s$.

We did the same measurement for the dODF and cdODF maps calculated for the whole brain. Computing dODF or cdODF maps is actually very time consuming. Indeed, the representation of such maps is done using many little triangles on a CPU. The display is then performed using the Open Graphics Library (OpenGL). The use of triangles leads to

a huge increase in computation time. This was overcome by switching to a computation technique based on GPU (graphical process unit) and more specifically using *shaders* as they are already available in the library we used for the developments.

Concerning the execution of the denoising algorithm on the cluster, additionally to the display task, a transfer task has to be accounted for. Indeed, each CPU of the cluster has to transfer the large-sized dODF and cdODF reconstructed maps on the server that displays the maps. This is taking a supplementary time, compared to the one-CPU process. To measure these times, we run the algorithms as previously. The dODF and cdODF calculation times of the RT process are shown in fig. 4.16 when using only one CPU and when using the cluster of 80 CPUs. We do not indicate the time to display the maps for some results with one CPU, as they were negligible compared to the other times. Pay attention that the display times $t_{display}$ indicated for the cases using the cluster also include the transfer time required when using the cluster. This explains why the green boxes are higher when using the cluster. We also notice in fig. 4.16 that the green box has a tendency to decrease when the maximum SH order increases, when using the cluster. This is surprising as the maps have a bigger size when the maximum SH order increases. This observation can be explained by the fact that, for a high maximum SH order, there might be a certain overlap between the reconstruction time and the transfer time. For example, a CPU may have finished its reconstruction task earlier than another CPU and the transfer may begin for this CPU earlier. This overlapping induced a decrease of the green box when the maximum SH order increases, but this does not mean that the display and the transfer of the maps happen faster with a higher maximum SH order. As before with the other maps, we obtained a global time far below $T_R = 14s$ with the use of the cluster.

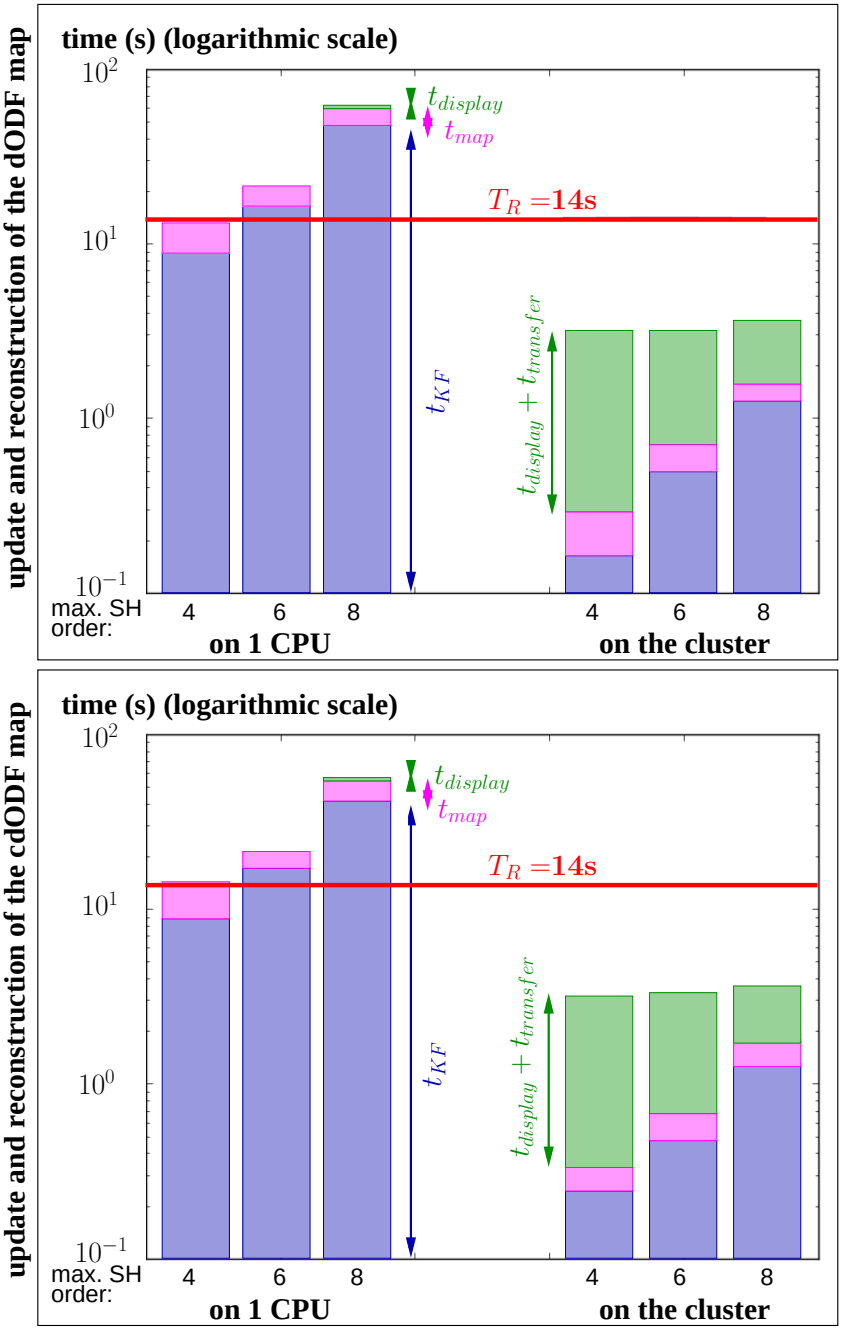


Figure 4.16: Times to get the dODF map and cdODF map update after a new iteration: t_{KF} (blue box) is the time required for one iteration of the KF algorithm and t_{map} (magenta box) is the time for processing the dODF map for the aQBI model (top) and the cdODF for the sa-aQBI model (bottom). Finally, $t_{display}$ (green box) is the time required to display the maps. It additionally contains the time required for the transfer of the data from the CPUs of the cluster to the server, when using the cluster. When $t_{display}$ is not shown, it is because it is very negligible compared to the other times. All times are compared to $T_R = 14s$.

4.6 Conclusion of this chapter

The motivations for RT dMRI are twofold: improving the patient comfort and the diagnosis, opening the way to more interactions with the acquisition parameters online to refine the results. After providing an overview of the various incremental techniques introduced in the literature, the Kalman framework was presented as well as its implementation to compute in RT local models such as DTI, aQBI and sa-aQBI. The hardware architecture was also detailed and used to provide proofs of concept of the feasability of doing RT inference of local modeling. In the next chapter, we will address the issue of non Gaussian noise distribution and show how to deal with it accurately in RT.

4.7 Appendix A: The KF derivation in a static framework (without regularization)

The KF equations in a static framework, as the one described in section 4.2, can be derivated from a WLS estimation detailed in [Chui and Chen (1987)], on which we base this appendix. The weighted error is written as $\mathbf{H}\boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon}$ the noise vector equal to $\mathbf{y} - \mathbf{A}\mathbf{x}$ and with $\mathbf{H}$ the weight matrix. The latter is defined such that $\mathbf{R}^{-1} = \mathbf{H}^{\mathrm{T}}\mathbf{H}$, with $\mathbf{R}$ the noise covariance matrix, presented in section 4.2 which is supposed to be diagonal with diagonal elements named R_i $\forall i \in [\![1; K]\!]$. Minimizing $\mathbf{H}\boldsymbol{\epsilon}$ in the least squares' sense means minimizing:

$$\begin{aligned}(\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon}) &= (\mathbf{H}\mathbf{y} - \mathbf{H}\mathbf{A}\mathbf{x})^{\mathrm{T}}(\mathbf{H}\mathbf{y} - \mathbf{H}\mathbf{A}\mathbf{x}) \\ &= (\mathbf{y} - \mathbf{A}\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{A}\mathbf{x}).\end{aligned} \tag{4.21}$$

To minimize $(\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon})$, we look at its derivative with respect to $\mathbf{x}$. This leads us to the normal equation:

$$\left(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\right)\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y}. \tag{4.22}$$

We obtained this normal equation from eq. 4.21. The demonstration is in appendix B. Let us rewrite eq. 4.22 for iteration i, *i.e.* when only i measurements have been done ($\mathbf{A}_i$ is then an $i\times$N matrix, $\mathbf{R}_i$ an $i\times i$ matrix and $\mathbf{y}_i$ an $i\times 1$ vector):

$$\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)\hat{\mathbf{x}}_i = \mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{y}_i. \tag{4.23}$$

To obtain our sequential estimator, our objective is to link $\hat{\mathbf{x}}_{i-1}$ to $\hat{\mathbf{x}}_i$. For that, let us first detail the term $\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i$, considering $\mathbf{A}_i = \begin{bmatrix} \mathbf{A}_{i-1} \\ \mathbf{a}_i \end{bmatrix}$, with $\mathbf{a}_i$ representing the i^{th} row of the matrix $\mathbf{A}$:

$$\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i = \begin{bmatrix} \mathbf{A}_{i-1}^{\mathrm{T}} & \mathbf{a}_i^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{i-1}^{-1} & 0 \\ 0 & R_i^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{i-1} \\ \mathbf{a}_i \end{bmatrix}, \tag{4.24}$$

From eq. 4.24, we obtain:

$$\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i = \mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1} + \mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i. \tag{4.25}$$

Now, let us detail the term $\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{y}_i$.

$$\begin{aligned}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{y}_i &= \begin{bmatrix} \mathbf{A}_{i-1}^{\mathrm{T}} & \mathbf{a}_i^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{i-1}^{-1} & 0 \\ 0 & R_i^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{i-1} \\ y_i \end{bmatrix} \\ &= \mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{y}_{i-1} + \mathbf{a}_i^{\mathrm{T}}R_i^{-1}y_i.\end{aligned} \tag{4.26}$$

Injecting eq. 4.25 and 4.26 into eq. 4.23 leads us to:

$$\left(\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1} + \mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i\right)\hat{\mathbf{x}}_i = \mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{y}_{i-1} + \mathbf{a}_i^{\mathrm{T}}R_i^{-1}y_i. \tag{4.27}$$

And, when rewriting eq. 4.23 at iteration $i-1$, we obtain:

$$\left(\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1}+\right)\hat{\mathbf{x}}_{i-1}=\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{y}_{i-1},$$

Then:

$$\left(\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1}+\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i\right)\hat{\mathbf{x}}_{i-1}=\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{y}_{i-1}+\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i\hat{\mathbf{x}}_{i-1}. \tag{4.28}$$

We substract eq. 4.27 with eq. 4.28:

$$\left(\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1}+\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i\right)\left(\hat{\mathbf{x}}_i-\hat{\mathbf{x}}_{i-1}\right)=\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\left(y_i-\mathbf{a}_i\hat{\mathbf{x}}_{i-1}\right). \tag{4.29}$$

We set $\mathbf{k}_i=\left(\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1}+\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i\right)^{-1}\mathbf{a}_i^{\mathrm{T}}R_i^{-1}$, which represents the gain, and this gives us: $\boxed{\hat{\mathbf{x}}_i=\hat{\mathbf{x}}_{i-1}+\mathbf{k}_i\left(y_i-\mathbf{a}_i\hat{\mathbf{x}}_{i-1}\right)}$. The expression of $\mathbf{k}_i$ can be simplified into: $\mathbf{k}_i=\mathbf{D}_i\mathbf{a}_i^{\mathrm{T}}R_i^{-1}$, with

$$\begin{aligned}\mathbf{D}_i&=\left(\mathbf{A}_{i-1}^{\mathrm{T}}\mathbf{R}_{i-1}^{-1}\mathbf{A}_{i-1}+\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i\right)^{-1}\\&=\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}.\end{aligned} \tag{4.30}$$

In the appendix C, we prove that the estimation error covariance matrix $\mathbf{P}_i$ is equal to $\mathbf{D}_i$. From the two expressions of $\mathbf{D}_i$ written above in 4.30, we can write, with $\mathbf{D}_i$ replaced by $\mathbf{P}_i$:

$$\mathbf{P}_i^{-1}=\mathbf{P}_{i-1}^{-1}+\mathbf{a}_i^{\mathrm{T}}R_i^{-1}\mathbf{a}_i. \tag{4.31}$$

Here, we use the *matrix inversion lemma* (**Lemma 1.**), whose proof is given in [Deriche et al. (2009)] and which is stated such that:

Lemma 1. *Let* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *be* $n\times n$, $m\times m$ *and* $n\times m$ *matrices respectively such that* $\left(\mathbf{A}^{-1}+\mathbf{C}\mathbf{B}^{-1}\mathbf{C}^{\mathrm{T}}\right)$ *and* $\left(\mathbf{B}+\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}\right)$ *are nonsingular, then:*

$$\left(\mathbf{A}^{-1}+\mathbf{C}\mathbf{B}^{-1}\mathbf{C}^{\mathrm{T}}\right)^{-1}=\mathbf{A}-\mathbf{A}\mathbf{C}\left(\mathbf{B}+\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}\right)^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{A}.$$

This lemma enables us to inverse eq. 4.31:

$$\mathbf{P}_i=\mathbf{P}_{i-1}-\mathbf{P}_{i-1}\mathbf{a}_i^{\mathrm{T}}\left(R_i+\mathbf{a}_i\mathbf{P}_{i-1}\mathbf{a}_i^{\mathrm{T}}\right)^{-1}\mathbf{a}_i\mathbf{P}_{i-1}. \tag{4.32}$$

We can show, and the demonstration of it is performed in appendix D, that:

$$\boxed{\mathbf{k}_i=\mathbf{P}_{i-1}\mathbf{a}_i^{\mathrm{T}}\left(R_i+\mathbf{a}_i\mathbf{P}_{i-1}\mathbf{a}_i^{\mathrm{T}}\right)^{-1}}.$$

Thus, from both latter equations, we obtain: $\boxed{\mathbf{P}_i=\mathbf{P}_{i-1}-\mathbf{k}_i\mathbf{a}_i\mathbf{P}_{i-1}}$. Rewriting the three boxed equations of this appendix A, we obtain:

$$\boxed{\begin{aligned}\hat{\mathbf{x}}_i&=\hat{\mathbf{x}}_{i-1}+\mathbf{k}_i\left(y_i-\mathbf{a}_i\hat{\mathbf{x}}_{i-1}\right),\\\mathbf{k}_i&=\left(R_i+\mathbf{a}_i\mathbf{P}_{i-1}\mathbf{a}_i^{\mathrm{T}}\right)^{-1}\mathbf{P}_{i-1}\mathbf{a}_i^{\mathrm{T}},\\\mathbf{P}_i&=\mathbf{P}_{i-1}-\mathbf{k}_i\mathbf{a}_i\mathbf{P}_{i-1},\end{aligned}}$$

which correspond to the KF equations presented in section 4.2.

4.8 Appendix B: Demonstration of the normal equation

Here, we want to minimize the weighted error $\mathbf{H}\boldsymbol{\epsilon}$. In appendix A, we have seen that: $(\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon}) = (\mathbf{y}-\mathbf{A}\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-\mathbf{A}\mathbf{x})$. This expression can be developed such that:

$$(\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon}) = \mathbf{y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} + \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\mathbf{x}.$$

And yet, we have:

$$\begin{aligned}\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} &= \left(\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y}\right)^{\mathrm{T}} \text{ (because it is scalar)}\\ &= \mathbf{y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\mathbf{x} \text{ (because } \mathbf{R}^{\mathrm{T}} = \mathbf{R}).\end{aligned}$$

Consequently, we obtain: $(\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\mathbf{x} - 2\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} + \mathbf{y}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y}$. Let us notice that $\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}$ is positive. Therefore, to find the optimum vector $\hat{\mathbf{x}}$ minimizing $(\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon})$, we just need to look for $\hat{\mathbf{x}}$ such that:

$$\frac{\partial\left((\mathbf{H}\boldsymbol{\epsilon})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\epsilon})\right)(\hat{\mathbf{x}})}{\partial\mathbf{x}} = 0. \tag{4.33}$$

And yet, there is a **Proposition 1.** about the matrix derivative stating that:

Proposition 1. *For a vector* $\mathbf{v}\in\mathbb{R}^k$ *and a matrix* $\mathbf{M}\in\mathbb{R}^{k\times k}$*:*

$$\frac{\partial}{\partial\mathbf{v}}\left(\mathbf{v}^{\mathrm{T}}\mathbf{M}\mathbf{v}\right) = \left(\mathbf{M}+\mathbf{M}^{\mathrm{T}}\right)\mathbf{v},$$

and in particular, if $\mathbf{M}$ *is symmetric,* $\mathbf{M}^{\mathrm{T}} = \mathbf{M}$ *and:*

$$\frac{\partial}{\partial\mathbf{v}}\left(\mathbf{v}^{\mathrm{T}}\mathbf{M}\mathbf{v}\right) = 2\mathbf{M}\mathbf{v},$$

Knowing the above proposition, eq. 4.33, which is quadratic in $\mathbf{x}$ simplifies into: $2\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\hat{\mathbf{x}} - 2\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} = 0$, which finally leads to: $\left(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A}\right)\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y}$.

4.9 Appendix C: Formula for the estimation error covariance matrix $\mathbf{P}_i$

Here, we want to prove that: $\mathbf{P}_i = \left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}$. For that, we exactly follow the mathematical steps of appendix D in [Deriche et al. (2009)]. The estimation error covariance matrix is defined such that: $\mathbf{P}_i = E\left[\left(\mathbf{x}-\hat{\mathbf{x}}_\mathbf{i}\right)\left(\mathbf{x}-\hat{\mathbf{x}}_\mathbf{i}\right)^{\mathrm{T}}\right]$. From this, we can write that:

$$
\begin{aligned}
&E\left[\left(\mathbf{x}-\hat{\mathbf{x}}_\mathbf{i}\right)\left(\mathbf{x}-\hat{\mathbf{x}}_\mathbf{i}\right)^{\mathrm{T}}\right] \\
&= E\left[\left(\mathbf{x}-\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{y}_i\right)\left(\mathbf{x}-\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{y}_i\right)^{\mathrm{T}}\right] \text{ (using} \\
&\text{eq. 4.23)} \\
&= E\left[\left(\mathbf{x}-\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\left(\mathbf{A}_i\mathbf{x}+\epsilon_i\right)\right)\left(\mathbf{x}-\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\left(\mathbf{A}_i\mathbf{x}+\epsilon_i\right)\right)^{\mathrm{T}}\right] \\
&= E\left[\left(\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\epsilon_i\right)\left(\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\epsilon_i\right)^{\mathrm{T}}\right] \\
&= \left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}E\left[\epsilon_i\epsilon_i^{\mathrm{T}}\right]\mathbf{R}_i^{-1}\mathbf{A}_i\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1} \\
&= \left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{R}_i\mathbf{R}_i^{-1}\mathbf{A}_i\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1} \\
&= \left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)\left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1} \\
&= \left(\mathbf{A}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{A}_i\right)^{-1}.
\end{aligned}
$$

4.10 Appendix D: Formula for the Kalman gain matrix $\mathbf{k}_i$

In appendix A, when we introduced the Kalman gain, we expressed it as $\mathbf{k}_i = \mathbf{P}_i \mathbf{a}_i^{\mathrm{T}} R_i^{-1}$. We want to demonstrate that
$\mathbf{k}_i = \mathbf{P}_i \mathbf{a}_i^{\mathrm{T}} R_i^{-1} = \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(R_i + \mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1}$. For that, we exactly follow the mathematical steps of appendix E in [Deriche et al. (2009)]. Using the *matrix inversion lemma* (**Lemma 1.**) stated in appendix A, we have:

$$
\begin{aligned}
&\mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(R_i + \mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1} \\
&= \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(R_i^{-1} - R_i^{-1} \mathbf{a}_i \left(\mathbf{P}_{i-1}^{-1} + \mathbf{a}_i R_i^{-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1} \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \right) \\
&= \left(\mathbf{P}_{i-1} - \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \mathbf{a}_i \left(\mathbf{P}_{i-1}^{-1} + \mathbf{a}_i R_i^{-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1} \right) \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \\
&= \left(\mathbf{P}_{i-1} - \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(\mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} + R_i \right)^{-1} \left(\mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} + R_i \right) R_i^{-1} \mathbf{a}_i \left(\mathbf{P}_{i-1}^{-1} + \mathbf{a}_i R_i^{-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1} \right) \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \\
&= \left(\mathbf{P}_{i-1} - \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(\mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} + R_i \right)^{-1} \left(\mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \mathbf{a}_i + \mathbf{a}_i \right) \left(\mathbf{P}_{i-1}^{-1} + \mathbf{a}_i R_i^{-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1} \right) \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \\
&= \left(\mathbf{P}_{i-1} - \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(\mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} + R_i \right)^{-1} \mathbf{a}_i \mathbf{P}_{i-1} \left(\mathbf{a}_i^{\mathrm{T}} R_i^{-1} \mathbf{a}_i + \mathbf{P}_{i-1}^{-1} \right) \left(\mathbf{P}_{i-1}^{-1} + \mathbf{a}_i R_i^{-1} \mathbf{a}_i^{\mathrm{T}} \right)^{-1} \right) \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \\
&= \left(\mathbf{P}_{i-1} - \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} \left(\mathbf{a}_i \mathbf{P}_{i-1} \mathbf{a}_i^{\mathrm{T}} + R_i \right)^{-1} \mathbf{a}_i \mathbf{P}_{i-1} \right) \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \\
&= \mathbf{P}_i \mathbf{a}_i^{\mathrm{T}} R_i^{-1} \text{ (using eq. 4.32)} \\
&= \mathbf{k}_i.
\end{aligned}
$$

Part III

Real-time noise correction for real-time dMRI (rtdMRI)

Chapter 5

Modeling of noise in MRI

Among the phenomena which corrupt the quality of MRI images, noise is a considerable one, particularly in dMRI where the sensitization to the diffusion process is characterized by a strong signal decay. Indeed, the dMRI measurement corresponds to the loss of phase consistency in the transverse magnetization signal, a loss, which is due to the water molecules' diffusion (see chapter 3). The signal level in dMRI quite often falls down to the noise level. What are the origins and the nature of this noise? Can we estimate it and if so, how? What actual —measurable ?— effect does it have on images? This chapter attempts to assemble all up-to-date knowledge about MRI noise while answering these questions. The first section gives a detailed description of noise according to different MRI sequence parameters. The second section summarizes the various methods of the literature available to estimate the noise, highlighting the difficulty to accurately estimate it. Finally, in the third section, we describe some classical image quality indices and propose a new one specific to dMRI images.

5.1 Origins of noise in MRI

When we perform an MRI exam on a subject, the collected data are always noisy. How is noise defined and what are its origins?

5.1.1 The origins and the experimental aspects of noise

Noise can be considered as being all the signal, which does not originate from the measured physical phenomenon and corresponds to what covers up and degrades the true signal. This parasite can have several origins: it can be a **thermal** noise (also called **Johnson-Nyquist** noise), which is due to random motion of charge carriers in electrical conductors of the MRI scanner system, as well as in the subject's body, which is also conductive. Noise can also come from the **external environment** (outside temperature changes...). Finally, a last possibility is that noise originates from **quantization error** due to analog-to-digital signal conversion. Among all the sources of corruption described above, only the thermal noise is random [Barbier (2004)]. In our study, we will not deal with noise

coming from the external environment that we consider to have very little effect on brain MRI results. Concerning the quantization error, it is also small and is negligible regarding thermal noise if the latter's standard deviation is not approximately equal or lower than one [Dietrich et al. (2008b)], which was always the case in our studies.

The noise, which we are interested in, is therefore the random thermal noise. It comes from **ohmic losses** in the receiving circuit, as well as from noise due to the pre-amplifier. The latter is used to amplify the measured signal before the analog-to-digital conversion followed by the double inverse Fourier transform of the measured signal (or the parallel reconstruction technique). The circuit between the receiving antenna and the pre-amplifier must have low losses and the pre-amplifier itself must have a low noise figure. This is usually achieved [Vlaardingerbroek and den Boer (2004)]. From the ohmic losses only remain those from the receiving circuit whose sources are divided in two: first, the **RF antenna** has its own losses. Secondly, eddy current losses exist in the **subject** itself produced by the Brownian motion of charge carriers within the body's conductive tissues. These eddy currents generate randomly varying magnetic fields, which then induce a noise electromotive force in the RF receiving antenna [Hoult and Richards (1976); Redpath (1998)]. At mid to high magnetic fields ($\mathbf{B_0} \geq 1T$), the dominant source of the receiving circuit is the one coming from the subject, except if the RF antenna is very small. In any case, the impact of thermal noise is that, even when no magnetic excitation is applied to acquire a signal, the voltage across the receiving coil, which measures the transverse magnetization, is not equal to zero: it contains the noise signal. This thermal noise is **white**, and therefore theoretically contains all frequencies. However, during the MRI signal detection process, a low pass filtering is used to only detect signals in a frequency read bandwidth (RBW) centered around the frequency of the MRI signal. Therefore, only the noise signals with frequencies contained in this RBW have an impact on the final result.

Let us now define the SNR in the MRI hardware context. In the first chapter, we saw with eq. 2.9 that the output of the receiving antenna, in a general case when no field gradient is applied, is proportional to:

$$e_{prop}(t) = \omega_0 \int_{\text{volume}} e^{-t/T_2^*(\mathbf{r})} A_\perp(\mathbf{r}, 0) B_\perp(\mathbf{r}) \sin\left(\omega_0 t + \theta_B(\mathbf{r}) - \phi_0(\mathbf{r})\right) d^3\mathbf{r}, \tag{5.1}$$

with ω_0 the Larmor angular frequency, T_2^* the effective transverse relaxation time and $A_\perp(\mathbf{r}, 0)$ the initial transverse magnetization at the localization given by $\mathbf{r}$. $B_\perp(\mathbf{r})$ is the magnitude of the transverse magnetic field for a unit current in the coil with $\theta_B(\mathbf{r})$ its angle measured in the laboratory frame of reference. Finally, $\phi_0(\mathbf{r})$ is the initial phase of the transverse magnetization. Now, for what concerns noise, we can write the root mean square of the noise voltage. Using the definition of the Johnson-Nyquist noise, the root mean square of the noise voltage is:

$$V_{noise} = \sqrt{4 k_B T \cdot \Delta f \cdot (R_a + R_s)}, \tag{5.2}$$

with k_B the Boltzmann's constant, T the absolute temperature of the resistive object and Δf the RBW. Finally, R_a is the resistance due to the antenna, whereas R_s is the

resistance due to subject. We then define the SNR at the location $\mathbf{r}$ in this hardware context as the ratio between the instantaneous output of the receiver to the root mean square noise voltage, and consequently [Vlaardingerbroek and den Boer (2004)]:

$$\begin{aligned} \mathrm{SNR}(\mathbf{r}) &\propto \frac{|e_{prop}(t)|}{V_{noise}}, \\ &\propto \frac{\omega_0 e^{-t/T_2^*(\mathbf{r})} |A_\perp(\mathbf{r},0)||B_\perp(\mathbf{r})|}{\sqrt{4k_B T \cdot \Delta f \cdot (R_a + R_s)}}. \end{aligned} \tag{5.3}$$

Several teams studied the effects of the hardware configuration on the SNR, and namely the effect of the field strength (neglecting the effect of $\mathbf{B_0}$ on T_2^*). For mid to high field systems ($\mathbf{B_0} \geq 1T$), with a not too small antenna, the SNR increases linearly with $\mathbf{B_0}$. Therefore, a gain in image quality is obtained with the use of high field scanners. Studies about hardware factors impacting the SNR can be found in [Haacke et al. (1999); Redpath (1998); Edelstein et al. (1986)].

From eq. 5.3, an SNR expression can be derived in the specific case of 2D imaging that depends on the scanning parameters. For a single-shot echo-planar DW spin echo pulse sequence in 2D, we have the following proportionality relation, which accounts for the use of partially parallel MRI reconstruction and for which T_2^* is replaced by T_2 because of the refocusing pulse in the sequence:

$$\mathrm{SNR}(\mathbf{r}) \propto \frac{\exp(-T_E/T_2(\mathbf{r}))(\Delta x \cdot \Delta y \cdot \Delta z)\sqrt{N_x N_y N_{acq}} B_0}{\sqrt{\Delta f \cdot R}\mathcal{G}}, \tag{5.4}$$

with Δx, Δy and Δz the dimensions of the voxel which indicate the resolution level. N_x and N_y are the dimensions in number of voxels of the image (*i.e.* the acquisition matrix). N_{acq} is the number of repetitions of the same acquisition. R is, as in chapter 2, the parallel acceleration factor. Finally, $\mathcal{G}$ is the quality factor of the receiving antenna. Some conclusions can be obtained from eq. 5.4: as before, we see that the SNR increases with $\mathbf{B_0}$. However, the time T_2 also decreases with $\mathbf{B_0}$. To make sure that the SNR gain obtained with an increase of $\mathbf{B_0}$ is not too much impacted by the decrease of T_2, a powerful gradient system has to be used: gradients with a high amplitude and high slew rates enable to reduce the echo time T_E and therefore reduce the T_2 attenuation. Eq. 5.4 also informs us that the higher the spatial resolution, the lower the SNR. Moreover, the higher the acquisition matrix size (for a same resolution), the higher the SNR. The SNR can be improved by a factor $\sqrt{p}$ when repeating p times the same acquisition. Finally, a low value of the RBW increases the SNR, but low RBW values increase the echo train duration inducing more geometrical distortions (due to susceptibility effects...). Some trade-off has therefore to be found [Redpath (1998)].

An important choice impacting the SNR is about the receiver. The choice of a phased-array coil increases the SNR, in comparison to the use of a volumic coil exploring the same FOV, as mentioned in the first chapter. Research teams studied more deeply the impact of a phased-array system on the noise. They highlighted the fact that in phased

array systems, **noise correlations** exist [Roemer et al. (1990); Hayes and Roemer (1990); Harpen (1992); Redpath (1992); Brown et al. (2007)]. These correlations reduce the SNR performance of the array. Noise correlations have two sources [Redpath (1992); Constantinides et al. (1997)]: they can come from mutual coupling between the channels. They are then due to the noise components arising from the coils themselves. Such correlation type can be eliminated by an adapted preamplifier hardware configuration. The second source of noise correlations is noise arising from the subject. The resulting correlations are then impossible to eliminate. If they impact the SNR performance of the phased array, they also produce a correlated noise signal in MRI images. This is an important point to keep in mind for next subsection.

Further analysis is possible on the impact of other contrasts and sequence choices on the SNR. But this is behind the scope of our introduction on noise. In the rest of this section, we will detail the characteristics of thermal noise in the case of an acquisition with one channel, and in the case of parallel MRI. We will insist on what makes this noise difficult to correct.

5.1.2 Noise modeling

The MRI signal received on each channel is complex with a real and an imaginary component. The expression of the MRI signal before the double inverse Fourier transform is given by eq. 2.13 in chapter 2. The thermal noise, described in the previous subsection, corrupts this complex signal. It also has a real and an imaginary component. These two noise components are assumed to be zero-mean and uncorrelated Gaussian distributed [Henkelman (1985)] which is consistent with the previous description of a white random noise. The zero-mean uncorrelated Gaussian distribution of each noise component remains true after the inverse Fourier transform is applied to each real and imaginary parts of the channel [Gudbjartsson and Patz (1995)]. Indeed, the inverse Fourier transform does not change the zero-mean Gaussianity characteristic of a function. However, in general the magnitude of the complex MRI signal is exploited. The magnitude of this complex signal leads to the usual MRI images examined by physicians. Indeed, it is most common to use the magnitude of the measured signal, instead of its real and imaginary parts, to avoid artifacts due to the phase, as explained in [Henkelman (1985); Constantinides et al. (1997); Nowak (1999)]. In most clinical MRI scanners, there exist a way to extract the complex **k**-space data, but in general, through proprietary file formats and tools not designed to be used in clinical routine. Furthermore, reconstruction pipelines are generally not opened to the customer and there is consequently no possibility to deal with noise correction at the complex signal level. Therefore, it turns to be necessary to develop ways to directly analyse noise from the magnitude images. Besides, the community of researchers working on estimation and correction of noise on the magnitude is large.

In this section, we will therefore present the statistics of the noise corrupting the magnitude. Let us consider the general case of a multiple-channel acquisition. Assuming an SoS reconstruction, the magnitude signal is written such that [Constantinides et al.

(1997)]:

$$M = \sqrt{\sum_{c=1}^{n} \left[M_{r_c}^2 + M_{i_c}^2\right]}, \tag{5.5}$$

with n the number of channels and M_{r_c} and M_{i_c} the real and imaginary parts, respectively, of the MRI complex signal measured by the channel c. The magnitude can also be expressed by the following relation:

$$\boxed{M = \sqrt{\sum_{c=1}^{n} \left[(S_{r_c} + \epsilon_{r_c})^2 + (S_{i_c} + \epsilon_{i_c})^2\right]}}, \tag{5.6}$$

with S_{r_c} and S_{i_c} the real and imaginary parts, respectively, of the noise-free complex signal S_c received by the channel c. The terms ϵ_{r_c} and ϵ_{i_c} represent the real and imaginary parts, respectively, of the complex noise ϵ_c corrupting the signal received by the channel c. For a single-channel acquisition, this complex noise has always been assumed to be a zero-mean and uncorrelated Gaussian distributed process, as mentioned earlier [Henkelman (1985)]. Concerning a multiple-channel acquisition, the real and imaginary parts of the complex noise signals are also assumed to be zero-mean Gaussian processes. However, we explained in the previous subsection that noise correlations exist. Therefore, the complex noise on a channel is eventually correlated with the complex noise signals on the other channels. Consequently, ϵ_{r_c} and ϵ_{i_c} may be correlated with the other real and imaginary parts of the signals from the rest of the channels. At first, we will neglect these correlations' issues and study noise in the simplified case. If we want to study the noise corrupting the magnitude M, we may write:

$$\boxed{M = S + \epsilon}, \text{ with } \begin{cases} S = \sqrt{\sum_{c=1}^{n} \left[S_{r_c}^2 + S_{i_c}^2\right]} = \sqrt{\sum_{c=1}^{n} S_c^2}, \\ \epsilon = \sqrt{\sum_{c=1}^{n} \left[\underbrace{\epsilon_{r_c}^2 + \epsilon_{i_c}^2}_{\epsilon_c^2} + 2\left(S_{r_c}\epsilon_{r_c} + S_{i_c}\epsilon_{i_c}\right)\right]}. \end{cases} \tag{5.7}$$

The noise term ϵ corrupting the magnitude M is the noise which we are interested in and that many researchers have studied. The non-linear transformation to get ϵ modifies the characterics of the noise: ϵ, in opposition to ϵ_{r_c} and ϵ_{i_c}, is no longer zero-mean Gaussian distributed. It is also a **noise-free signal dependent noise**.

Fig. 5.1 provides a short history of the investigations conducted about noise in MRI from the magnitude of the signal. First, in 1984, the team of Edelstein studied the noise in the MRI image background, in the single-channel case (*i.e.* $n = 1$). They showed that the **background magnitude** M follows a **Rayleigh distribution**, defined for $M \geq 0$ as [Edelstein et al. (1984)]:

$$\underline{\text{RAYLEIGH PDF:}} \quad \boxed{p(M;\sigma) = \frac{M}{\sigma^2}.\exp\left(-\frac{M^2}{2\sigma^2}\right)}, \tag{5.8}$$

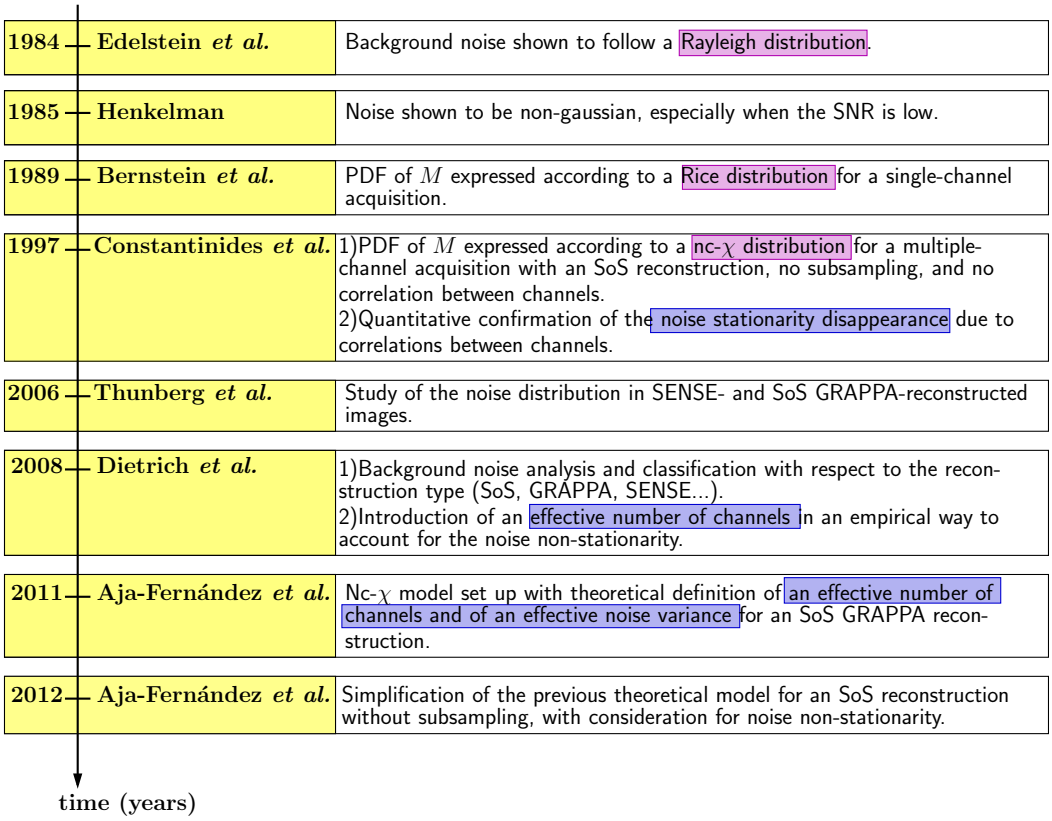


Figure 5.1: A brief history* of MRI noise modeling. *(Inspired by Stephen Hawking...)

with $p(M;\sigma)$ the PDF of M with σ its standard deviation. The latter PDF is drawn in dark blue in fig. 5.2, on page 112. It looks like a Gaussian, with the big difference that it is not symmetric and that it contains a positive skewness. Moreover, its mean is not zero. Then, in 1985, Mark Henkelman emphasized the non zero-mean Gaussian nature of the noise in MRI magnitude images when $n = 1$, especially in regions with a low SNR. It has to be well accounted for a good noise correction or a good SNR measurement. In 1989, Mattew Bernstein *et al.* were the first to express, for the whole image, the PDF of M concluding it follows a Rice distribution in the case of a single-channel acquisition ($n = 1$). This distribution is defined for $M \geq 0$ and $S \geq 0$ and written as [Bernstein et al. (1989); Rice (1952)]:

$$\underline{\text{RICIAN PDF:}} \quad \boxed{p(M;S,\sigma) = \frac{M}{\sigma^2}.\exp\left(-\frac{M^2+S^2}{2\sigma^2}\right)\cdot I_0\left(\frac{S\cdot M}{\sigma^2}\right)}, \tag{5.9}$$

with I_0 the modified Bessel function of the first kind of order 0 (defined at the end of the manuscript, in appendix B) and σ the noise standard deviation on the imaginary and real components of the receiving channel. Chris Constantinides *et al.* studied the case of multiple-channel acquisitions ($n > 1$) with an SoS reconstruction, assuming that there is no noise correlation in the phased array system. Moreover, the impact of parallel MRI accelerated sequence with subsampling on the noise characteristics was not tackled.

Finally all channels were assumed to have the same variance. In this case, they showed that the PDF of the signal received by n channels is a **nc-χ distribution**, defined for $M \geq 0$ and $S > 0$, and written such that [Constantinides et al. (1997)]:

$$\underline{\text{NC-}\chi \text{ PDF:}} \quad \boxed{p(M;S,\sigma,n) = \frac{S}{\sigma^2}\left(\frac{M}{S}\right)^n \exp\left(-\frac{M^2+S^2}{2\sigma^2}\right) \cdot I_{n-1}\left(\frac{S\cdot M}{\sigma^2}\right)}, \tag{5.10}$$

with I_{n-1} the modified Bessel function of the first kind of order $n-1$. But, when correlations between channels exist, the noise —more precisely its distribution— loses its stationarity, as pointed out by [Hayes and Roemer (1990); Redpath (1992); Constantinides et al. (1997)]. This means that the noise standard deviation σ is then **voxel-dependent**, as well as **signal-dependent**. This non-stationarity is indeed caused by the fact that the correlations are dependent on the noisy signals at each voxel. As the noisy signals are voxel-dependent, so are also the correlations. In 2006, Per Thunberg *et al.* showed the spatial variation of noise with two partially parallel MRI techniques, namely SENSE (with and without regularization) and SoS GRAPPA [Thunberg and Zetterberg (2007)]. In 2008, Olaf Dietrich *et al.* studied histograms of the background noise in different reconstruction situations to build a classification of the noise PDF depending on many reconstruction criteria [Dietrich et al. (2008a)]. They also introduced, for the SoS reconstruction, an effective number of channels in an empirical way to account for the correlations' effects on the PDF of M. Finally in 2011 and 2012, the team of Santiago Aja-Fernández and Antonio Tristán-Vega built a theoretical definition of this effective number of channels, as well as of an effective noise variance, in case of an SoS GRAPPA reconstruction [Aja-Fernández et al. (2011)] and in case of an SoS reconstruction without subsampling [Aja-Fernández and Tristán-Vega (2012)]. They showed that the nc-χ distribution defined with these two effective parameters was a good approximation for the PDF of M in such cases.

We will now detail the noise nature for the single-channel and the multiple-channel acquisitions. The SoS GRAPPA case is of particular interest for us, as we will deal with real data obtained using this reconstruction process.

Noise distribution in case of a single-channel acquisition

In case of a single-channel acquisition, the magnitude follows a Rice distribution, which we rewrite here for $M \geq 0$ and $S \geq 0$:

$$\underline{\text{RICIAN PDF:}} \quad \boxed{p(M;S,\sigma) = \frac{M}{\sigma^2}.\exp\left(-\frac{M^2+S^2}{2\sigma^2}\right) \cdot I_0\left(\frac{S\cdot M}{\sigma^2}\right)}. \tag{5.11}$$

Its first and second moments are:

$$\boxed{\begin{cases} E(M) = \sqrt{2}\cdot\sigma\cdot\Gamma\left(\frac{3}{2}\right)\cdot {}_1F_1\left(-\frac{1}{2};1;-\frac{S^2}{2\sigma^2}\right), \\ E(M^2) = S^2 + 2\cdot\sigma^2, \end{cases}} \tag{5.12}$$

with Γ the Gamma function, defined at the end of the manuscript in appendix B. ${}_1F_1$ is the confluent hypergeometric function of the first kind defined in the same appendix.

From eq. 5.11, the PDF followed by the noise ϵ can be derived, using the following **Theorem 1.**:

Theorem 1. *f_X is the PDF of the random variable X. If a random variable Y is written such that $Y = g(X)$, with g a differentiable and strictly monotonic function, then the resulting PDF followed by Y is expressed such that:*

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \left|\frac{1}{g'(g^{-1}(y))}\right|,$$

where g^{-1} is the inverse function of g and g' is the derivative of g.

Then, the PDF of ϵ, with $\epsilon = M - S$, is given by:

$$p(\epsilon; S, \sigma) = \frac{\epsilon + S}{\sigma^2}.\exp\left(-\frac{(\epsilon+S)^2 + S^2}{2\sigma^2}\right) \cdot I_0\left(\frac{S\cdot(\epsilon+S)}{\sigma^2}\right). \tag{5.13}$$

Both PDF curves given by eq. 5.11 and eq. 5.13 are drawn in fig. 5.2 for different values of S/σ. When $S = 0$, M follows a Rayleigh distribution given by eq. 5.8. It corresponds to the dark blue curve in fig. 5.2. We also note that the higher the ratio S/σ, the more the PDF of ϵ looks like a zero-mean Gaussian distribution.

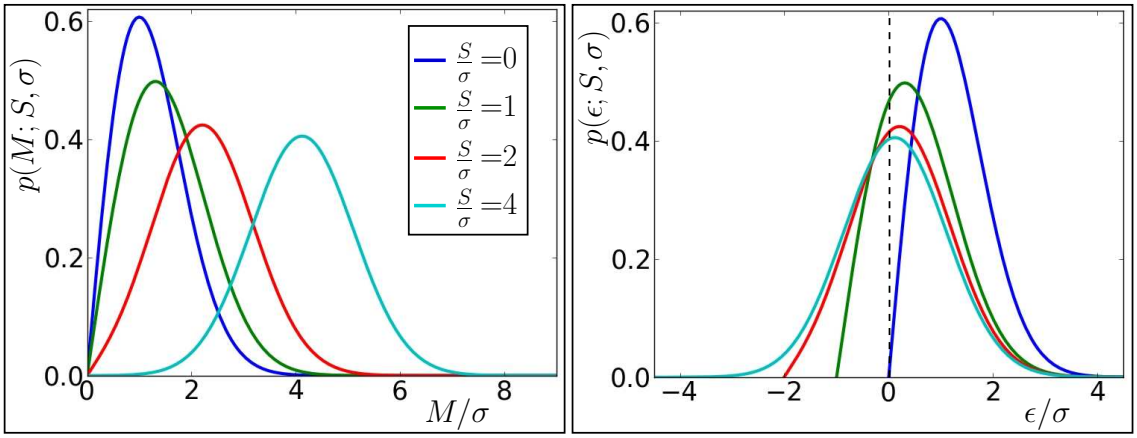


Figure 5.2: PDF of M (left) and ϵ (right) in case of a single-channel acquisition for several values of S/σ. The higher the ratio S/σ, the more the PDF of ϵ looks like a zero-mean Gaussian. The legend for colors indicated in the square on the left also applies on the right.

In the previous chapter, we saw that for ratios $S/\sigma \geq 4$, the PDF followed by ϵ can be approximated by a Gaussian $\mathcal{N}(0, \sigma^2)$ [Henkelman (1985); Sijbers et al. (1998)]. Indeed, fig. 5.3 shows the convergence of $E(M)$ on S, as well as the convergence of the noise standard deviation of M, written as σ_M, on σ for high SNR values. Considering now the noise term $\epsilon = M - S$, its mean is $E(\epsilon) = E(M - S)$ and therefore it converges on 0. Its noise standard deviation σ_ϵ is equal to σ_M and therefore converges on σ. For low SNR values, the approximation of a zero-mean Gaussian noise with a standard deviation equal to σ is no longer valid. We then speak of the **Rician bias** [Gudbjartsson and Patz

(1995); Sijbers et al. (1998)], which corresponds to the value of ϵ at low SNR. Compared to a zero-mean Gaussian noise, the noise ϵ described here is far more difficult to remove, precisely because it is **dependent on the noise-free signal**. Indeed, the PDF given by eq. 5.13 depends on S. Least-squares estimators, which are efficient to remove zero-mean Gaussian noise, will here not be able to remove the Rician bias at low SNR values. And yet, we will see in section 5.1.3 that the SNR can be very low in dMRI and therefore it is required to account for the non zero-mean and non Gaussian nature of the noise.

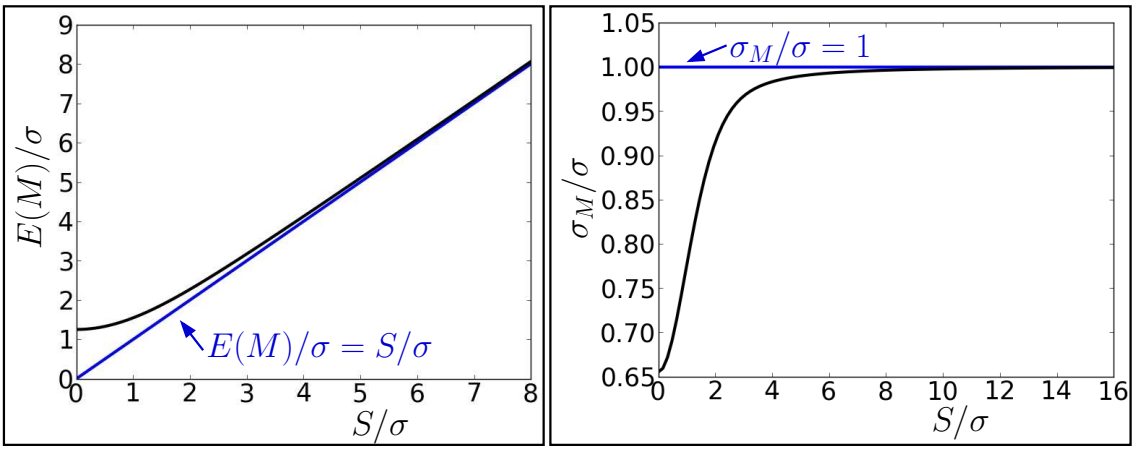


Figure 5.3: Mean and noise standard deviation of the PDF followed by M in case of a single-channel acquisition.

What are the noise characteristics when the acquisition is performed using a phased-array coil system?

Noise distribution in case of a multiple-channel acquisition

In neuroimaging, it is increasingly common to acquire MRI data using receiving antenna arrays containing from 8 to 32 channels. There are several reconstruction techniques in parallel MRI. Well-known examples are GRAPPA and SENSE among others. A simple SoS reconstruction without subsampling can also be performed. In this manuscript, we will only focus on three reconstruction algorithms: the SoS reconstruction without subsampling, the SoS GRAPPA reconstruction and finally the SENSE algorithm.

SoS reconstruction without subsampling. We rewrite here the expression of the magnitude signal obtained with an SoS reconstruction [Constantinides et al. (1997)]:

$$M = \sqrt{\sum_{c=1}^{n} ((S_{r_c} + \epsilon_{r_c})^2 + (S_{i_c} + \epsilon_{i_c})^2)}, \tag{5.14}$$

with S_{r_c} and S_{i_c} the real and imaginary parts, respectively, of the noise-free complex signal S_c received by the channel c. The terms ϵ_{r_k} et ϵ_{i_k} represent the real and imaginary parts, respectively, of the complex noise corrupting the signal received by the channel c. These two noise components for each channel are assumed to be zero-mean and uncorrelated Gaussian distributed [Henkelman (1985); Constantinides et al. (1997)]. We consider the following hypotheses: there is no correlation between the channels and all real and imaginary parts of the different channels have all the same noise standard deviation written as σ. Consequently, the noise corrupted final magnitude M, written as $M = S + \epsilon$, follows a nc-χ distribution, defined for $M \geq 0$ and $S > 0$ by [Constantinides et al. (1997)]:

$$\underline{\text{NC-}\chi \text{ PDF:}} \quad \boxed{p(M; S, \sigma, n) = \frac{S}{\sigma^2} \left(\frac{M}{S}\right)^n \exp\left(-\frac{M^2 + S^2}{2\sigma^2}\right) \cdot I_{n-1}\left(\frac{S \cdot M}{\sigma^2}\right)}. \tag{5.15}$$

Its first and second moments are:

$$\boxed{\begin{cases} E(M) = \sqrt{2} \cdot \sigma \cdot (n)^{(1/2)} \cdot {}_1F_1\left(-\frac{1}{2}; n; -\frac{S^2}{2\sigma^2}\right), \\ E(M^2) = S^2 + 2 \cdot n \cdot \sigma^2, \end{cases}} \tag{5.16}$$

with $(n)^{(1/2)}$ calculated applying the definition of the Pochhammer rising factorial symbol given at the end of the manuscript in appendix B. For a single-channel acquisition (*i.e.* $n = 1$), eq. 5.15 simplifies to the Rice distribution given by eq. 5.11 and eq. 5.16 simplifies to eq. 5.12. When $S = 0$, M follows a **central χ distribution**, which equation is the following for $M \geq 0$:

$$\underline{\text{C-}\chi \text{ PDF:}} \quad \boxed{p(M; 0, \sigma, n) = \frac{1}{\Gamma(n)\sigma^2} \left(\frac{M}{2\sigma^2}\right)^{n-1} M^n \exp\left(-\frac{M^2}{2\sigma^2}\right)}. \tag{5.17}$$

As for the Rician distribution, it is possible to derive the PDF followed by the noise ϵ, using the **Theorem 1**, page 112:

$$p(\epsilon; S, \sigma, n) = \frac{S}{\sigma^2} \left(\frac{(\epsilon + S)}{S}\right)^n \exp\left(-\frac{(\epsilon + S)^2 + S^2}{2\sigma^2}\right) \cdot I_{n-1}\left(\frac{S \cdot (\epsilon + S)}{\sigma^2}\right). \tag{5.18}$$

Both PDF curves given by eq. 5.15 and eq. 5.18 are drawn in fig. 5.4 for different values of S/σ, as well as for different values of n. The central χ distributions (when $S = 0$) corresponds to the dark blue curves in the left column in fig. 5.4.

We also remark that the higher the n, the smaller the skewness of the curves, especially at low SNR values: the curves are then more symmetric. Also, as in fig. 5.2, the higher the ratio S/σ, the more the PDF of ϵ looks like a zero-mean Gaussian.

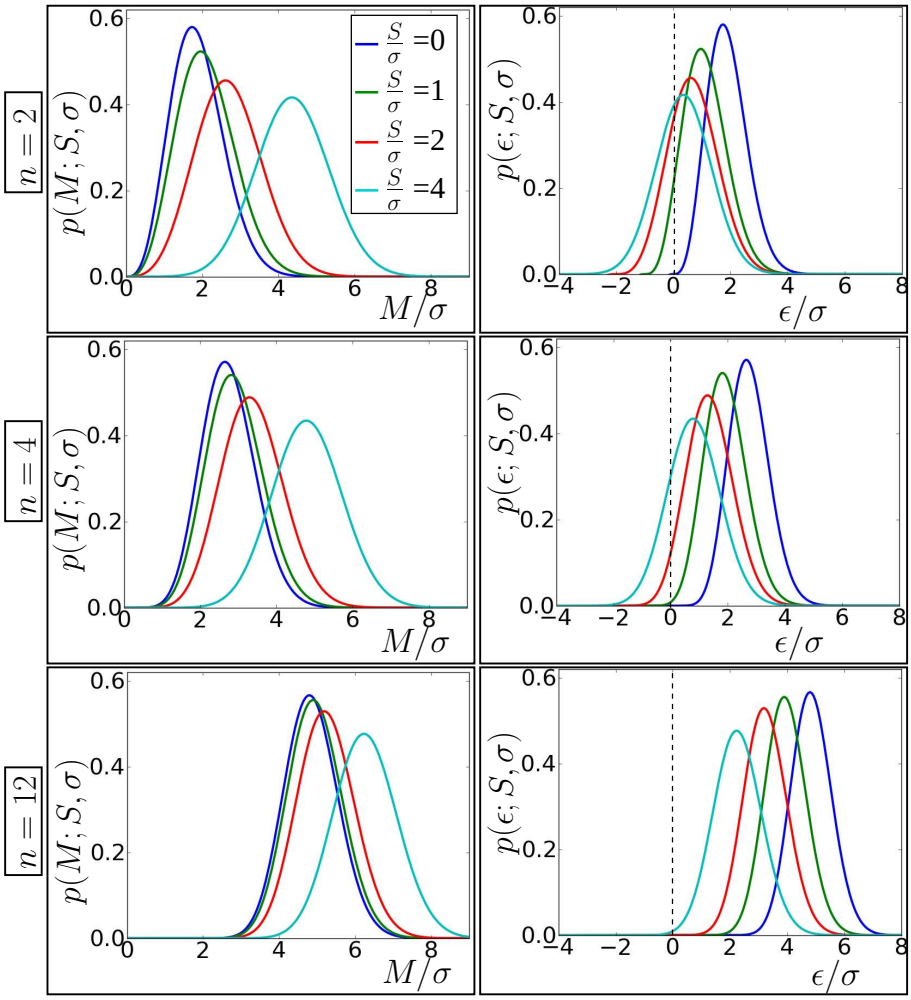


Figure 5.4: PDF of M (left) and ϵ (right) in case of a multiple-channel acquisition under the three following hypotheses: **no subsampling, no correlation between the channels, and same variance for each channel**. The PDF curves are represented for $n = 2/4/12$ and for several values of S/σ. The higher the n, the smaller the skewness of the curves, especially at low SNR values: the curves are then more symmetric. Also, as in fig. 5.2, the higher the ratio S/σ, the more the PDF of ϵ looks like a zero-mean Gaussian. The convergence on this similarity is slower as n increases. The legend for colors indicated in the top left square applies to all plots.

The latter observation is confirmed by fig. 5.5, which shows the convergences of $E(M)$ and σ_M on S and σ, respectively. This convergence is however slower as n increases. Considering now the noise term $\epsilon = M - S$, its mean is $E(\epsilon) = E(M - S)$ and therefore

it converges on 0. Its noise standard deviation σ_ϵ is equal to σ_M and therefore converges on σ. As for the previously described Rician noise, this nc-χ noise is very difficult to remove especially for low SNR values, as it can no longer be approximated by a zero-mean Gaussian noise. A **nc-χ bias** appears, which **depends on the noise-free signal**. Moreover, compared to the Rician PDF, the approximation by a zero-mean Gaussian noise is here valid for larger SNR values than in the single-channel case. This effect is increased with higher values of n.

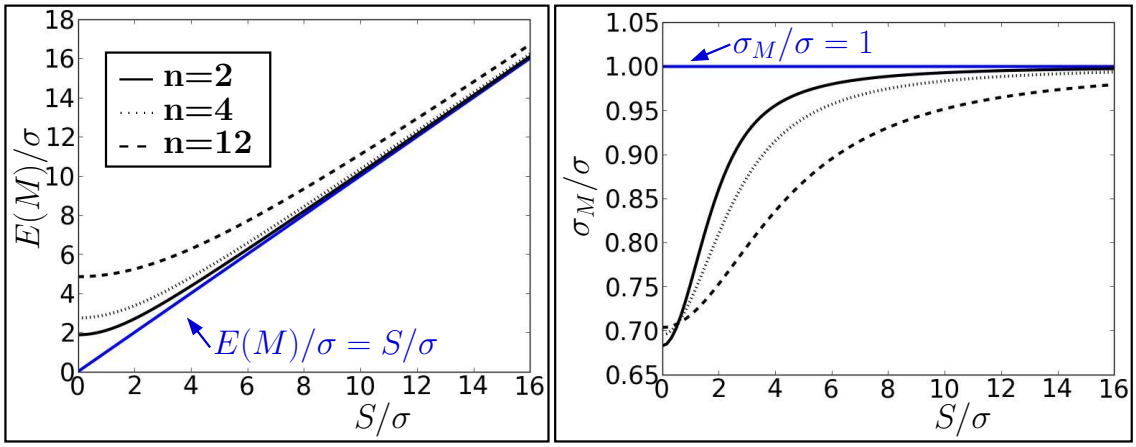


Figure 5.5: Mean and noise standard deviation of the PDF followed by M in case of a multiple-channel acquisition, under the three same hypotheses as in fig. 5.4. The higher the number of channels n, the slower the convergences of $E(M)$ and σ_M on S and σ, respectively. The legend indicated in the left square also applies for the right square.

To obtain the above described statistical characteristics of the noise with an SoS reconstruction without subsampling, we needed to assume that there is no correlation between the noise signals and that all channels have the same variance σ^2. However, in pratice, these conditions, particularly the first one, are often not satisfied. In phased array systems, noise correlations exist and they induce a spatial variation of noise —more precisely of its distribution—, as mentioned previously.

Concerning SoS reconstructions without subsampling, the authors of [Dietrich et al. (2008a)] proposed to calculate an effective number of channels n_{eff} to account for these correlations. Their study only concerned the background. It was based on producing the histogram of noise in an assumed stationary region of the background. They created a lookup table (LUT) linking integer values of the number of channels n with the corresponding theoretical ratio of the mean value over the standard deviation of the background signal. They obtained the ratio using the characteristics of the nc-χ PDF. Then, they measured the experimental ratio obtained with the measurements inside the large background region and designated the corresponding number of channels in the LUT as n_{eff}. The latter was always obtained lower than the true n used in the experiment. Indeed, this lower n_{eff} reflects the fact that correlations reduce the number of independent channels' outputs.

The authors of [Aja-Fernández and Tristán-Vega (2012)] went recently further in the analysis: they considered the whole image, and not only the background anymore. They showed that the nc-χ model does not hold to describe the distribution of M for SoS reconstructions without subsampling. But, they proved that a nc-χ distribution defined with an effective number of channels, as well as an effective noise variance is a good approximation for the true noisy distribution in the whole volume. It is highlighted in their paper that these both effective parameters are non-stationary, and therefore have to be calculated at each voxel of the volume. This calculation requires to compute the following **voxel-wise covariance matrix**:

$$\mathbf{\Sigma}^2(\mathbf{v}) = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2(\mathbf{v}) & \dots & \sigma_{1n}^2(\mathbf{v}) \\ \sigma_{21}^2(\mathbf{v}) & \sigma_2^2 & \dots & \sigma_{2n}^2(\mathbf{v}) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1}^2(\mathbf{v}) & \sigma_{n2}^2(\mathbf{v}) & \dots & \sigma_n^2 \end{pmatrix}, \tag{5.19}$$

with σ_i^2 the stationary variance of the noise on the channel i and $\sigma_{ij}^2(\mathbf{v})$ the covariance of the noise signals from channels i and j defined at the voxel $\mathbf{v}$. The **effective number of channels** $n_{eff}(\mathbf{v})$ and the **effective noise variance** $\sigma_{eff}^2(\mathbf{v})$ are given by:

$$n_{eff}(\mathbf{v}) = \frac{S^2(\mathbf{v}) \cdot Tr(\mathbf{\Sigma}^2(\mathbf{v})) + Tr(\mathbf{\Sigma}^2)^2(\mathbf{v})}{\mathbf{S}_\mathbf{c}^*(\mathbf{v}) \cdot \mathbf{\Sigma}^2(\mathbf{v}) \cdot \mathbf{S}_\mathbf{c}(\mathbf{v}) + \|\mathbf{\Sigma}^2(\mathbf{v})\|_F^2}, \tag{5.20}$$

$$\sigma_{eff}^2(\mathbf{v}) = \frac{Tr(\mathbf{\Sigma}^2(\mathbf{v}))}{n_{eff}(\mathbf{v})}, \tag{5.21}$$

where $Tr(\cdot)$ is the trace operator. The n×1 vector $\mathbf{S_c}$ contains all complex noise-free signals received on each channel: $\mathbf{S_c}(\mathbf{v}) = [S_1(\mathbf{v}), \dots, S_n(\mathbf{v})]^{\mathrm{T}}$. Its conjugate transpose is written as $\mathbf{S}_\mathbf{c}^*(\mathbf{v})$. Finally, $\|\cdot\|_F$ represents the Frobenius norm, defined, for a matrix $\mathbf{X}$, such that: $\|\mathbf{X}\|_F = \sqrt{Tr\left(\mathbf{X}\mathbf{X}^*\right)}$. These voxel-wise expressions of both effective parameters seem difficult to use in practice: indeed, with this information, calculating $n_{eff}(\mathbf{v})$ and $\sigma_{eff}^2(\mathbf{v})$ requires a simultaneous estimation at each voxel of $n_{eff}(\mathbf{v})$, $\sigma_{eff}^2(\mathbf{v})$ and the noise-free signals' vector $\mathbf{S_c}(\mathbf{v})$. This is very difficult to achieve in pratice. But a simplified scenario with some approximations was developed to account for this theory and we will present it in the paragraph page 124.

How does the distribution model change with an SoS GRAPPA reconstruction?

SoS GRAPPA reconstruction. An SoS GRAPPA reconstruction has been shown to increase the non-stationarity of the noise distribution in the data volume [Dietrich et al. (2008a); Aja-Fernández et al. (2011)]. Besides, the authors of [Dietrich et al. (2008a)] obtained a lower empirical n_{eff} in the background with an SoS GRAPPA reconstruction compared to an SoS reconstruction without subsampling. The team of Santiago Aja-Fernández and Antonio Tristán-Vega performed the same analysis of the PDF of M as previously, with an adaptation for the GRAPPA reconstruction. For that, they used the

voxel-wise GRAPPA interpolation matrix, defined by:

$$\mathbf{W}(\mathbf{v}) = \begin{pmatrix} W_1(1,\mathbf{v}) & \dots & W_n(n,\mathbf{v}) \\ \vdots & \dots & \vdots \\ W_n(1,\mathbf{v}) & \dots & W_n(n,\mathbf{v}) \end{pmatrix}, \tag{5.22}$$

with $\mathbf{W}_i(j,\mathbf{v})$ the GRAPPA reconstruction coefficient in the spatial domain. This coefficient is obtained by the double inverse Fourier transform of the GRAPPA convolution kernel defined in the **k**-space [Aja-Fernández et al. (2011); Breuer et al. (2009)]. Using the GRAPPA interpolation matrix, the covariance matrix of the interpolated data at each voxel $\mathbf{v}$ is then expressed such that:

$$\mathbf{C}^2(\mathbf{v}) = \mathbf{W}(\mathbf{v}) \cdot \mathbf{\Sigma}^2(\mathbf{v}) \cdot \mathbf{W}^*(\mathbf{v}), \tag{5.23}$$

with $\mathbf{\Sigma}^2(\mathbf{v})$ the voxel-wise covariance matrix given by eq. 5.19. The authors of [Aja-Fernández et al. (2011)] then adapted the previous expressions of the effective parameters, replacing the matrix $\mathbf{\Sigma}^2$ by $\mathbf{C}^2$ in eq. 5.20 and 5.21:

$$n_{eff}(\mathbf{v}) = \frac{S^2(\mathbf{v}) \cdot Tr(\mathbf{C}^2(\mathbf{v})) + Tr(\mathbf{C}^2)^2(\mathbf{v})}{\mathbf{S}_\mathbf{c}^*(\mathbf{v}) \cdot \mathbf{C}^2(\mathbf{v}) \cdot \mathbf{S}_\mathbf{c}(\mathbf{v}) + \|\mathbf{C}^2(\mathbf{v})\|_F^2}, \tag{5.24}$$

$$\sigma_{eff}^2(\mathbf{v}) = \frac{Tr(\mathbf{C}^2(\mathbf{v}))}{n_{eff}(\mathbf{v})}. \tag{5.25}$$

Similarly as previously, these expressions are difficult to use in pratice.

Finally, how does the distribution model change with a SENSE reconstruction?

SENSE reconstruction. With a SENSE method, the noise distribution is also non-stationary. It was shown to vary spatially according to the voxel-wise geometry factor, which is an indicator of the geometry arrangement between the coils [Weiger et al. (2001)]. The authors of [Thunberg and Zetterberg (2007)] also studied the mSENSE algorithm [Wang et al. (2001)] that includes a regularization. For both SENSE algorithms, noise varies spatially and quite less uniformly than with GRAPPA. We invite the reader to read [Thunberg and Zetterberg (2007)] for more details on the spatial noise variations with SENSE and mSENSE. Concerning a theoretical PDF that could be defined as for the GRAPPA case on the whole MRI volume, nothing has been deeply experimented yet. However, with SENSE, as the final reconstructed image is generated from the modulus of only one complex image, the number of channels should be intuitively taken as 1: a Rician PDF with a voxel-wise noise variance could then be intuitively used to describe the noisy measured MRI signal statistics, as experimentally demonstrated for a background area in [Dietrich et al. (2008a)] and as assumed for noise correction in [Rajan et al. (2012b)].

The approximative statistical noise models in case of a parallel MRI reconstruction are summed up in fig. 5.6. What is now the impact of this noise on DW images and more particularly on HARDI data?

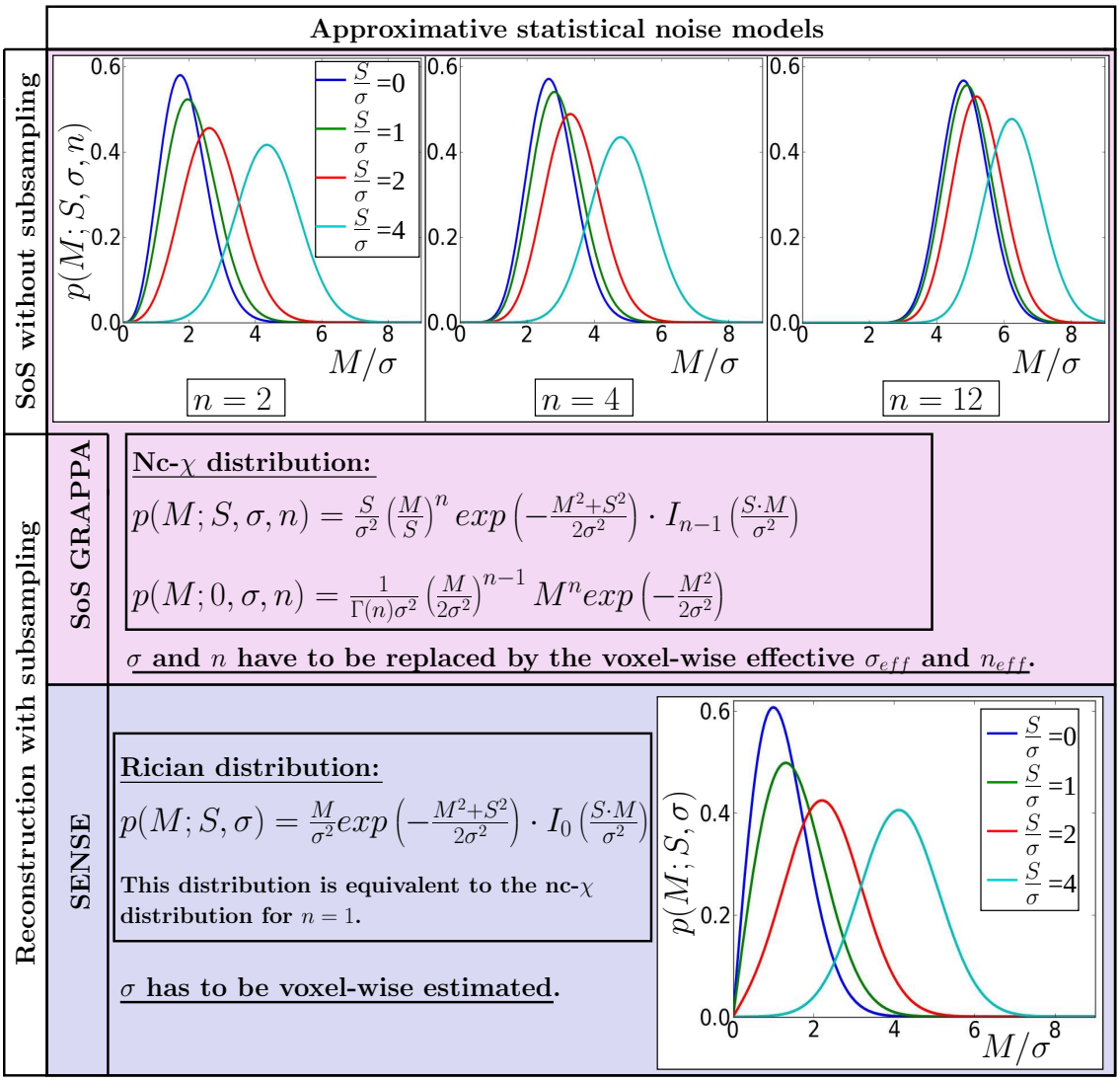


Figure 5.6: Approximative statistical noise models for three parallel MRI reconstructions: SoS, SoS GRAPPA and SENSE (from top to bottom). For both SoS and SoS GRAPPA algorithms, the nc-χ distribution was shown to be a good approximation for the actual noise distribution in MRI images at one condition: to account for the noise non-stationarity, σ and n have to be replaced by voxel-wise effective parameters n_{eff} and σ_{eff}, using eq. 5.20, 5.21, 5.24 and 5.25. For the SENSE algorithm, a Rician distribution is commonly assumed. To account for the noise non-stationarity, the variance has to be estimated at each voxel.

5.1.3 Noise impact on DW images and HARDI models

In DW MRI, the diffusion process in tissues induces a signal loss in the images, as seen in chapter 3. This signal loss increases with b, and for high *b-values* ($b > 3000 s/mm^2$), the signal can easily be swallowed up in noise. Indeed, the higher the diffusion, the lower the DW signal and the higher the noise impact on this signal. Consequently, DW images often have a low SNR, and this is more pronounced on HARDI data, often involving

higher *b-values*. This can be verified in fig. 5.7. A non-zero component appears in the image, whereas the signal is either zero or very low: this is the **bias effect** mentioned in the previous subsection, as well as in [Jones and Basser (2004)] for ADC profiles and in [Clarke et al. (2008)] for fODF profiles. Fig. 5.8 highlights this noise bias effect on the DW signal shape for the case of a pancake shape in a vertical fiber bundle. This bias makes the signal contrast between directions of strong and low diffusion significantly attenuated, and leads, for instance, to some decrease of the FA. Fig. 5.9 shows the noise impact on RGB maps, which are reconstructed from the DW images. Again, the higher the b, the higher the noise level in the image.

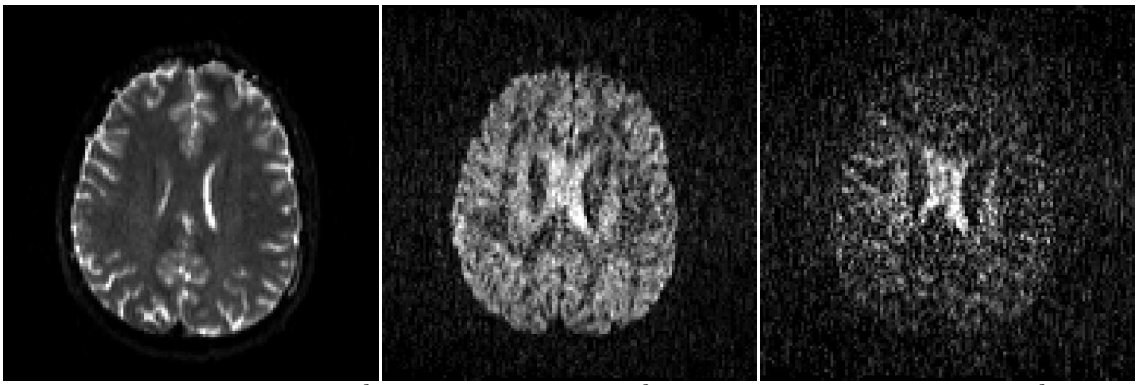


Figure 5.7: DW images for four different *b-values* and a T_2-weighted image. These data are taken from the "Archi database" described in appendix A at the end of the manuscript. Pay attention that the color scale is not the same between the images.

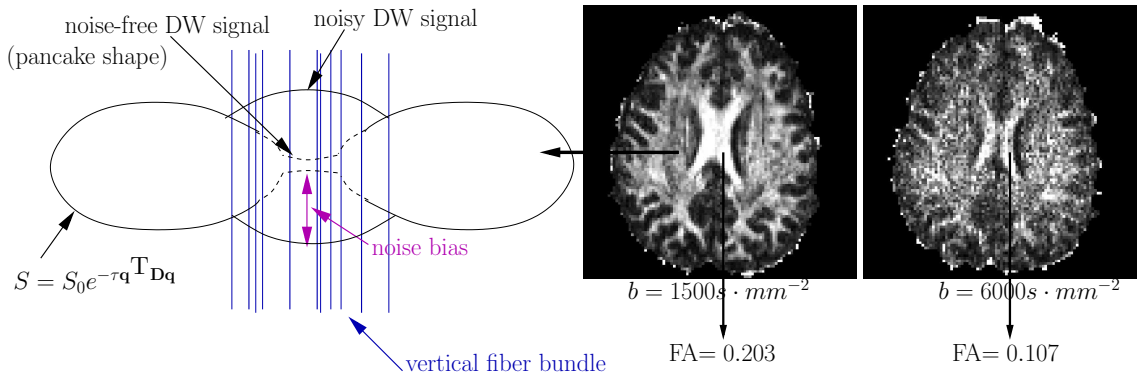


Figure 5.8: Noise bias on the DW signal found in a vertical bundle (*e.g.* the superior longitudinal fasciculus) as shown on the FA map. Two FA maps are presented for two different *b-values*. These data are taken from the "Archi database" described in appendix A at the end of the manuscript. Pay attention that the color scale is not the same between the images.

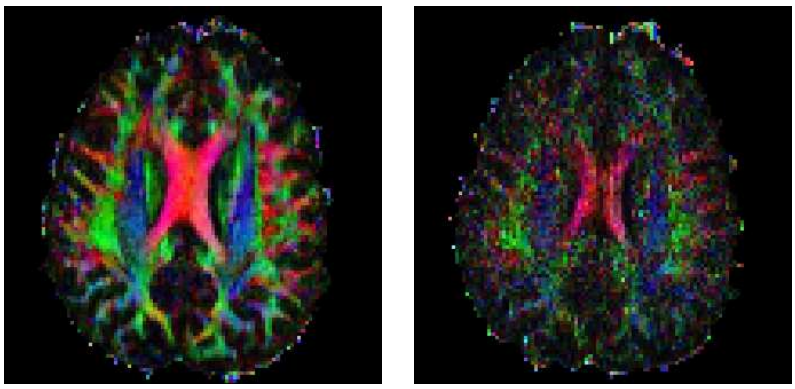


Figure 5.9: RGB maps for two different *b-values*. These data are taken from the "Archi database" described in appendix A at the end of the manuscript. Pay attention that the color scale is not the same between the images.

Finally, we performed the simulation of a noise-free DW signal corresponding to the crossing of two fiber bundles. We simulated the same DW signal with the addition of a nc-χ noise. We performed the simulations using $b = 6000s \cdot mm^{-2}$, $n = 4$ channels and σ in a range of values between 0.23 (SNR= 30) to 14 (SNR= 0.5) (see more details on the other simulation parameters in appendix C). We computed the corresponding noise-free and noisy dODFs. The comparison between both DW signals and the corresponding dODFs is shown in fig. 5.10. A low DW signal along an orientation corresponds to a high level of diffusion and therefore to a dODF peak. Consequently, we see in fig. 5.10 that the yellow and magenta dotted lines representing two lobes in the 3D DW signal generate dODF lobes in the perpendicular directions (where the diffusion process is at its highest). Fig. 5.10 highlights the effects of nc-χ noise on the DW 3D signal and on the dODF. First, noise corrupts the noise-free object and degrades its angular resolution which is shown by the

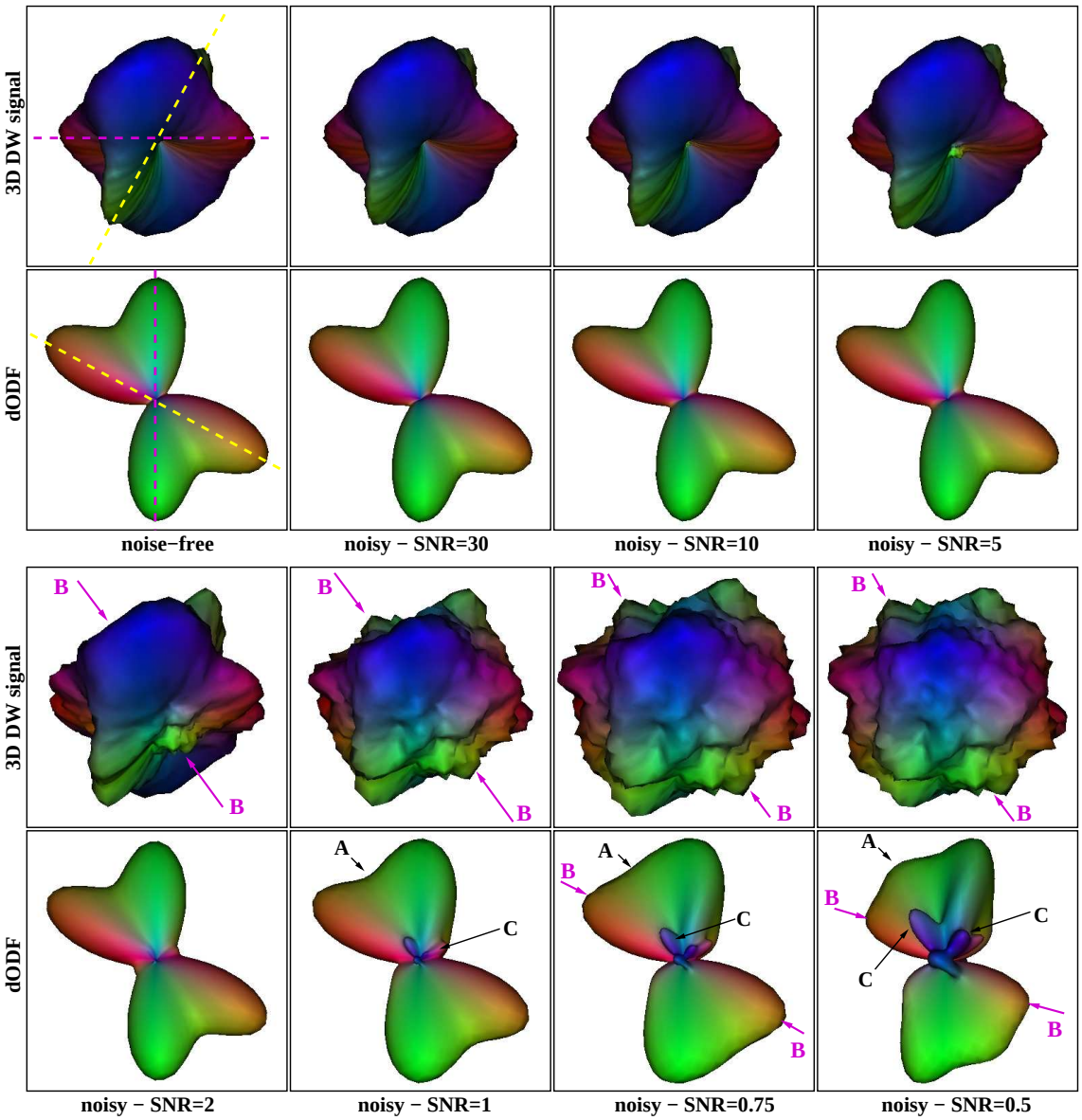


Figure 5.10: Simulations of noise-free and noisy 3D DW signals with their corresponding dODFs for several SNR values. Pay attention that the SNR is calculated from the noise-free DW signal ($S = 7$), and not from the noise-free T_2-weighted signal ($S_0 = 200$). The yellow and magenta dotted lines representing two lobes in the 3D DW signal generate dODF lobes in the perpendicular directions (where the diffusion process is at its highest). On the noisy dODFs, the letter A shows the loss of angular resolution. The letter B indicates the noise bias effect, making the lobes of the dODFs being shrinked. Finally, the letter C shows the wrong peaks created by noise on the dODFs. The simulations were performed using $b = 6000s \cdot mm^{-2}$, $n = 4$ and σ from 0.23 to 14. More details on the simulation parameters are given in appendix C at the end of the manuscript.

letter A in the noisy dODF: it is more difficult to differentiate the two peaks representing the two most probable diffusion direction. Moreover, we can notice that the vertical lobe in the noise-free dODF is deviated after the noise addition (this is particularly visible at

SNR= 0.5). Secondly, the bias due to noise appears in regions of very low noise-free DW signal and is pointed out by the letter B. On the dODF, this bias squashes the dODF lobes, similarly as for the ADC profiles as explained in [Jones and Basser (2004)]. Finally, noise is also responsible for creating spurious peaks in the dODF. This is shown by the letter C (and also by the letter A at SNR= 0.5). This may induce the belief that there is another orientation to account for in the dODF. Again, we see that noise can really bring errors in the data analysis, and consequently lead to putative errors in the direction full yielding false positives during the tractography steps.

That is why an adapted noise correction becomes essential to get reliable and detailed information about the directionality of tissue structures. Before correcting the noise, it is necessary to estimate its level given by the variance parameter.

5.2 MRI noise estimation methods

In this section, we detail two kinds of noise estimation methods: the background-dependent and background-free noise variance estimation techniques. Moreover, to account for the noise non-stationarity, the estimation should be done at each voxel.

5.2.1 Background-dependent noise estimation methods

Most of the noise estimation methods are background-dependent. And most of them define a noise variance globally for the whole data volume. This global variance corresponds to the variance of the noise present on each channel of the receiving coil in the hypotheses of no subsampling, no correlation and the same variance for each channel. The unknown parameter is therefore the σ parameter of the nc-χ PDF eq. 5.10.

Global noise variance estimation

A review of such global noise variance estimation techniques can be found in [Aja-Fernández et al. (2009)]. Let us here point out that another interesting method, called PIESNO, was developed recently by [Koay et al. (2009b)], which proposes a way to identify the noise-only voxels to improve the noise variance estimation.

In this study, we chose to test three commonly used techniques introduced by the review of [Aja-Fernández et al. (2009)]. They correspond to the ones called "mode M1-χ", "mode M2-χ" and "mode V1-χ" in the review. These methods can be applied either in the entire image or in the segmented background. These methods therefore depend on the presence of a background in the image, but do not require any background segmentation. However, in our procedure with real DW data, the background was automatically segmented, as explained in appendix A at the end of the manuscript. Thus, we could easily estimate σ using only the background voxels, leading to a more robust estimation.

The "mode M1-χ" method relies on the computation of the first moment of the measured magnitude in the background. It is derived from the first moment expression given

by eq. 5.16 for a nc-χ distribution in a background area, with ${}_1F_1\left(-\frac{1}{2}; n; 0\right) = 1$. This first moment can be calculated locally in the background with the use of a neighborhood. Its histogram is drawn and the mode of the histogram is then considered. The estimated global noise standard deviation $\hat{\sigma}$ using the "mode M1-χ" method is:

$$\underline{\text{ModeM1-}\chi\text{:}} \quad \boxed{\hat{\sigma} = \left(\sqrt{2}(n)^{(1/2)}\right)^{-1} \text{mode}\left(\langle M_{bg}(\mathbf{v})\rangle\right)}, \tag{5.26}$$

where $M_{bg}(\mathbf{v})$ is the measured magnitude at the voxel $\mathbf{v}$ in the background region and $\text{mode}\left(\langle M_{bg}(\mathbf{v})\rangle\right)$ is the distribution mode of the local mean of $M_{bg}(\mathbf{v})$; n is the number of channels and $(n)^{(1/2)}$ is calculated applying the definition of the Pochhammer symbol, which can be found in appendix B at the end of the manuscript.

The so called "mode M2-χ" method is based on the computation of the second moment of the measured magnitude in the background. It is derived from the second eq. in the eq. system 5.16 and it gives the following estimated noise standard deviation:

$$\underline{\text{ModeM2-}\chi\text{:}} \quad \boxed{\hat{\sigma} = \sqrt{\frac{1}{2n}\text{mode}\left(\langle M_{bg}^2(\mathbf{v})\rangle\right)}}, \tag{5.27}$$

with the same parameters as previously.

Finally, the so called "mode V1-χ" method is based on the computation of the local variance of the measured magnitude in the background. It is derived from eq. 5.26 and eq. 5.27. It gives the following estimated noise standard deviation:

$$\underline{\text{ModeV1-}\chi\text{:}} \quad \boxed{\hat{\sigma} = \sqrt{\left[2n - 2\frac{\Gamma^2(n+\frac{1}{2})}{\Gamma^2(n)}\right]^{-1} \text{mode}\left[\text{Var}(M_{bg}(\mathbf{v}))\right]}}, \tag{5.28}$$

with Γ the Gamma function defined in appendix B at the end of the manuscript. $\text{Var}(M_{bg}(\mathbf{v}))$ is the locally computed variance of $M_{bg}(\mathbf{v})$.

We tested these methods on simulated data. The results are given in subsection 5.2.3. However, these methods estimate a stationary noise variance, whereas we showed page 117 the necessity to estimate a voxel-wise variance. Would it be possible to have such a voxel-wise estimation?

Voxel-wise noise variance estimation

We saw in the previous paragraphs that noise is non-stationary in most cases of parallel MRI. Concerning SENSE reconstructions, a spatially variable noise variance estimation was proposed by [Landman et al. (2009)]. In this manuscript however, we deal with data obtained using SoS and SoS GRAPPA reconstructions, therefore we prefered to focus on the corresponding noise variance estimation techniques.

For the SoS reconstruction without subsampling and for the SoS GRAPPA reconstruction, expressions of a voxel-wise effective noise variance were proposed (see eq. 5.21 and 5.25, respectively). However, these expressions require a simultaneous estimation at each

voxel of $n_{eff}(\mathbf{v})$, $\sigma^2_{eff}(\mathbf{v})$ and the noise-free signals' vector $\mathbf{S_c}(\mathbf{v})$. This is actually very difficult to achieve. However, for the case of the SoS reconstruction without subsampling, the authors of [Aja-Fernández and Tristán-Vega (2012)] and [Aja-Fernández et al. (2013)] proposed to study a simplified scenario with the following hypotheses: the variance of noise is the same for every channel and the noise-free signals received by each channel for the voxel $\mathbf{v}$ are the same. With this simplified scenario, it is possible to obtain the **lower** and **upper bounds** of the effective variance across the whole volume: indeed, the authors of [Aja-Fernández and Tristán-Vega (2012)] showed that σ_{eff} can be estimated in the background (lower bound) and in the high-SNR areas (upper bound). This scenario therefore assumes that the effective variance is stationary through the background, as well as through high-SNR areas. Following this simplified scenario, the authors of [Aja-Fernández et al. (2013)] went further and proposed an estimation of the effective variance, as well as of the effective number of channels, at each voxel $\mathbf{v}$ of the image volume. However, their estimation is only valid if the previously mentioned hypotheses of the scenario are respected. Moreover, it was developed for SoS reconstruction without subsampling only. But, it is the only existing estimation of both voxel-wise effective parameters to date. This estimation requires to follow the computation steps enumerated in table 5.1. For more details about the mathematical demonstration, we invite the reader to look in [Aja-Fernández et al. (2013)].

The authors of [Aja-Fernández et al. (2013)] emphasized the fact that the product $n_{eff}(\mathbf{v}) \cdot \sigma^2_{eff}(\mathbf{v})$ does not require such a simplified scenario to be computed. Indeed, for an SoS reconstruction without subsampling, this product is equal to $Tr\left(\mathbf{\Sigma}^2(\mathbf{v})\right) = \sum_{i=1}^{n} \sigma_i^2$, with σ_i the standard deviation of the i^{th} channel. Thus, $n_{eff}(\mathbf{v}) \cdot \sigma^2_{eff}(\mathbf{v})$ is then stationary. For a GRAPPA reconstruction, this product is equal to $Tr\left(\boldsymbol{C}^2(\mathbf{v})\right)$, which is not stationary, but does not depend on any noise-free signal value and can be estimated using the GRAPPA reconstruction coefficients and the covariance matrix. This detail about the product $n_{eff}(\mathbf{v}) \cdot \sigma^2_{eff}(\mathbf{v})$ can be helpful in case of noise correction methods that rely on this product rather than on each effective parameter separately [Aja-Fernández et al. (2013)]. Nevertheless, we will see in the next chapters that it was not our case. Therefore, we could not use this trick.

The noise variance estimation methods described so far required the presence of a background area in the image volume. Is there any method, which is free from a background presence in the image?

Table 5.1: Summary of the voxel-wise variance estimation method proposed by [Aja-Fernández et al. (2013)] and valid for a simplified scenario in case of an SoS reconstruction without any partially parallel MRI reconstruction.

The estimation works contains the following steps:

1. A **noise level estimator**, written as $\hat{\sigma}_n^2$, is defined such that:

$$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}\left(\langle M_{bg}(\mathbf{v})^2\rangle\right), \tag{5.29}$$

where $M_{bg}(\mathbf{v})$ is the measured magnitude at the voxel $\mathbf{v}$ in the background region. Using eq. 5.27, $\hat{\sigma}_n^2$ is also equal to the product $n_{eff}(\mathbf{v}) \cdot \sigma_{eff}^2(\mathbf{v})$. In the case of an SoS reconstruction without partially parallel MRI reconstruction, this product is equal to $Tr\left(\mathbf{\Sigma}^2(\mathbf{v})\right) = \sum_{i=1}^{n} \sigma_i^2$, which is stationary.

2. Then, the variance of M is calculated in the background region such that:

$$\widehat{\sigma_{M_{bg}}^2} = \text{mode}\left(\text{Var}\left(M_{bg}(\mathbf{v})\right)\right), \tag{5.30}$$

3. From the two latter parameters, the effective noise variance and the effective number of channels are estimated in the background with a stationarity assumption across the background region. Their estimation requires the following iterative process:

$$\begin{cases} \left(\sigma_{eff,B}^2\right)_{t+1} = \dfrac{\widehat{\sigma_{M_{bg}}^2}}{2\left(n_{eff,B}\right)_t - 2\left[\dfrac{\Gamma^2\left(\left(n_{eff,B}\right)_t + \frac{1}{2}\right)}{\Gamma^2\left(\left(n_{eff,B}\right)_t\right)}\right]}, \\ \left(n_{eff,B}\right)_{t+1} = \dfrac{\widehat{\sigma_n^2}}{\left(\sigma_{eff,B}^2\right)_{t+1}}, \end{cases} \tag{5.31}$$

with $\left(n_{eff,B}\right)_0$ set to the real number of channels n. Eq. 5.31 is based on eq. 5.28.

4. Concerning high-signal regions, the effective noise variance and the effective number of channels are estimated such that:

$$\begin{cases} \widehat{\sigma_{eff,S}^2} = \text{mode}\left(\text{Var}\left(M_S(\mathbf{v})\right)\right), \\ n_{eff,S} = \dfrac{\widehat{\sigma_n^2}}{\widehat{\sigma_{eff,S}^2}}, \end{cases} \tag{5.32}$$

5. Finally, the **effective voxel-wise parameters** are obtained such that:

$$\underline{\text{Eff-v:}} \quad \boxed{\begin{cases} \widehat{\sigma_{eff}^2}(\mathbf{v}) = \left(1 - \widehat{\Psi_n}(\mathbf{v})\right) \cdot \widehat{\sigma_{eff,S}^2} + \widehat{\Psi_n}(\mathbf{v}) \cdot \widehat{\sigma_{eff,B}^2}, \\ \widehat{n_{eff}}(\mathbf{v}) = \dfrac{\widehat{\sigma_n^2}}{\widehat{\sigma_{eff}^2}(\mathbf{v})}, \end{cases}} \tag{5.33}$$

with $\Psi_n(\mathbf{v})$ a function estimated such that: $\widehat{\Psi_n}(\mathbf{v}) = \dfrac{\widehat{\sigma_n^2}}{\langle M_{bg}^2(\mathbf{v})\rangle - \widehat{\sigma_n^2}}$.

5.2.2 Background-free noise estimation method

Some methods were developed to estimate a global variance σ^2 without knowledge of the background. These methods assumed a Rician noise [Coupé et al. (2010); Rajan et al. (2010)]. We have extended the technique by [Rajan et al. (2010)] to nc-χ noise in [Brion et al. (2011b)]. It constitutes one contribution of this thesis work. The technique requires to estimate the variance σ_M^2 and the skewness γ of the magnitude MRI signal, at each voxel of the image, using:

$$\sigma_M^2 = E(M^2) - E(M)^2 \tag{5.34}$$

$$\gamma = \left(2E(M)^3 - 3E(M)E(M^2) + E(M^3)\right)/\sigma_M^3, \tag{5.35}$$

where the expectations $E(\cdot)$ can also be replaced by a local spatial mean $\langle\cdot\rangle$. The method relies on the computation of a local correction factor φ which tunes the proximity of the nc-χ distribution toward the central χ distribution for very low SNR, or toward the Gaussian distribution for very high SNR. For a central χ distribution, the variance can be estimated using: $\hat{\sigma} = \sqrt{\left[2n - 2\frac{\Gamma^2(n+\frac{1}{2})}{\Gamma^2(n)}\right]^{-1} \text{mode}\left[\sigma_M^2\right]}$, as in eq. 5.28. For a Gaussian distribution, the variance can be estimated using: $\hat{\sigma} = \sqrt{\text{mode}\left[\sigma_M^2\right]}$, similarly as in eq. 5.32. The proximity of the nc-χ distribution toward the central χ or the Gaussian distribution is measured by the skewness: the smaller the skewness, the less the distribution looks like a central χ distribution and the more it is close to a Gaussian. We created as in [Rajan et al. (2010)] a lookup table $\varphi(\gamma)$ between the local correction factor φ and the local skewness γ for a nc-χ distribution with a given n, by varying the value of S and keeping σ constant. In order to build this lookup table (fig. 5.11), the three first nc-χ moments have been calculated from the range of values of S and σ:

$$\begin{aligned}
E(M) &= \sqrt{2}\sigma(n)^{(\frac{1}{2})} \times {}_1F_1\left(-\tfrac{1}{2}, n, -\tfrac{S^2}{2\sigma^2}\right) \\
E(M^2) &= 2n\sigma^2 + S^2 \\
E(M^3) &= 2\sqrt{2}\sigma^3(n)^{(\frac{3}{2})} \times {}_1F_1\left(-\tfrac{3}{2}, n, -\tfrac{S^2}{2\sigma^2}\right),
\end{aligned}$$

with $(x)^{(a)}$ being the Pochhammer rising factorial symbol and ${}_1F_1\left(a, b, c\right)$ being the confluent hypergeometric function, both detailed in appendix B at the end of the manuscript. Then, γ is obtained by injecting the three expressions above in eq. 5.35. In the same manner, the local variance σ_M is computed using eq. 5.34. Finally, φ is calculated from $\varphi = \sigma^2/\sigma_M^2$. Once the $\varphi(\gamma)$ lookup table is computed, the estimation of the variance can be performed on an image. From the local estimates σ_M^2 (eq. 5.34) and γ (eq. 5.35) computed at each voxel using the neighbors, the final noise standard deviation can be estimated from:

$$\underline{\text{Bg-free:}}\ \boxed{\hat{\sigma} = \sqrt{\text{mode}\left(\varphi \times \sigma_M^2\right)}}, \tag{5.36}$$

where $\text{mode}\left(\varphi \times \sigma_M^2\right)$ is the distribution mode of $\varphi \times \sigma_M^2$ calculated on the image.

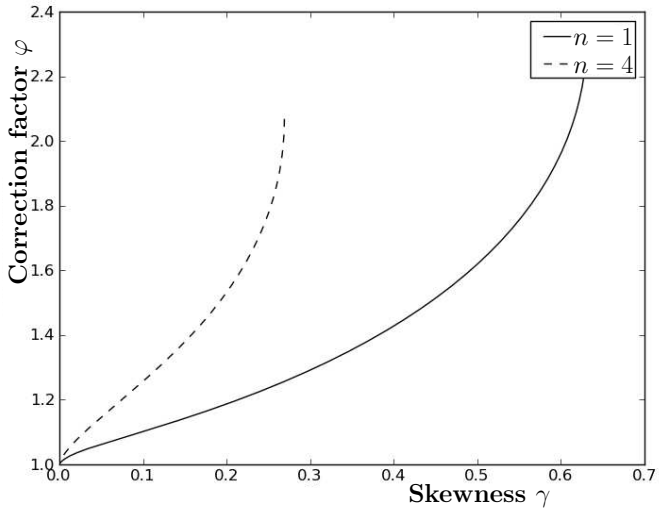


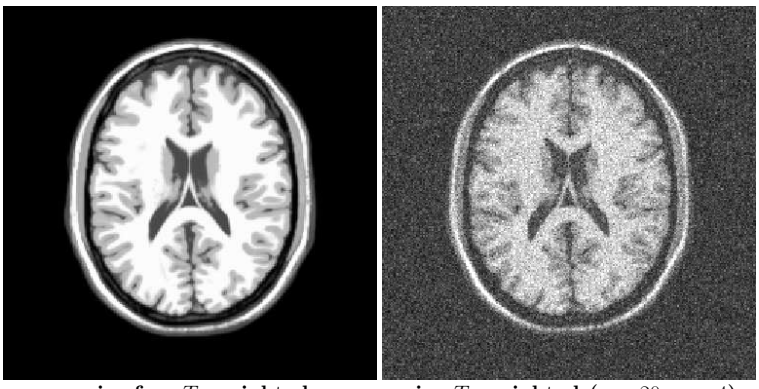
Figure 5.11: Look up tables $\varphi(\gamma)$ used in the background-free noise estimation method for $n = 1$ (Rician case) and $n = 4$.

5.2.3 Comparison between the global noise estimation methods in a stationary simulation case

We compared the previous background-free noise standard deviation method, called "Bg-free", with the methods called "mode M1-χ", "mode M2-χ" and "mode V1-χ" on a simulated T_1-weighted image, on which we added stationary nc-χ noise. The noise-free T_1-weighted image was taken from the BrainWeb database [Collins et al. (1998)]. Its intensity values are comprised between 0 and 255. This T_1-weighted image, shown in fig. 5.12, contains a background. First, we added to this image a Rician noise ($n = 1$) to test the original "Bg-free" technique of [Rajan et al. (2010)]. Second, we added a nc-χ noise with $n = 4$ to the noise-free image to test our extension of the "Bg-free" technique. The noise addition was performed with different global standard deviations. The noise addition process on the BrainWeb simulated image is explained in the appendix C at the end of the manuscript. (It is the same process as the one used on our simulated DW volume).

Concerning our experiment, our purpose was to evaluate the proposed nc-χ extended background-free noise variance estimation technique. The latter was therefore compared to the "mode M1-χ", "mode M2-χ" and "mode V1-χ" methods, known to offer very good results as demonstrated in [Aja-Fernández et al. (2009)]. We thus tested all methods below:

- the "Bg-free" technique on the image without its background,
- the "mode M1-χ", "mode M2-χ" and "mode V1-χ" techniques on the segmented background.



noise-free T_1-weighted **noisy T_1-weighted ($\sigma = 20$, $n = 4$)**

Figure 5.12: Noise-free and noisy T_1-weighted images.

All techniques were performed using a 7×7 neighborhood, and the histograms were calculated using 500 bins "Bg-free" technique and using 1000 bins for the other methods. Concerning the parameters' choice for the "mode M1-χ", "mode M2-χ" and "mode V1-χ" techniques, we followed the advice in [Aja-Fernández et al. (2009)]. Concerning the "Bg-free" technique, it was shown to be quite sensitive to the choice of these parameters and we made our decision after several trials.

We present our results in fig. 5.13 in a similar manner as in [Rajan et al. (2010)]. Concerning the comparison between the "Bg-free" technique and the "mode M1-χ", "mode M2-χ" and "mode V1-χ" techniques, it is clear that the latter methods perform better than the former one. The methods which use the background information are therefore more robust than the "Bg-free" technique which estimates the noise level without the background knowledge. It seems that the "Bg-free" technique, in both Rician and nc-χ noise cases, overestimate the standard deviation for low noise levels and then underestimate it for high noise levels. This latter method does not necessitate any background in the data, which can be a determinant point when working with some data that do not have enough background (for instance when using a small FOV). However, it has to be well kept in mind that this method is sensitive to the choice of the neighborhood size and that it can yield an error on $\hat{\sigma}$ of at least 20% (calculated here for $\sigma = 20$ with $n = 1$). Moreover, this method should be tested on other data to really define its performances, as it was shown to work better on the cardiac image used in [Rajan et al. (2010)].

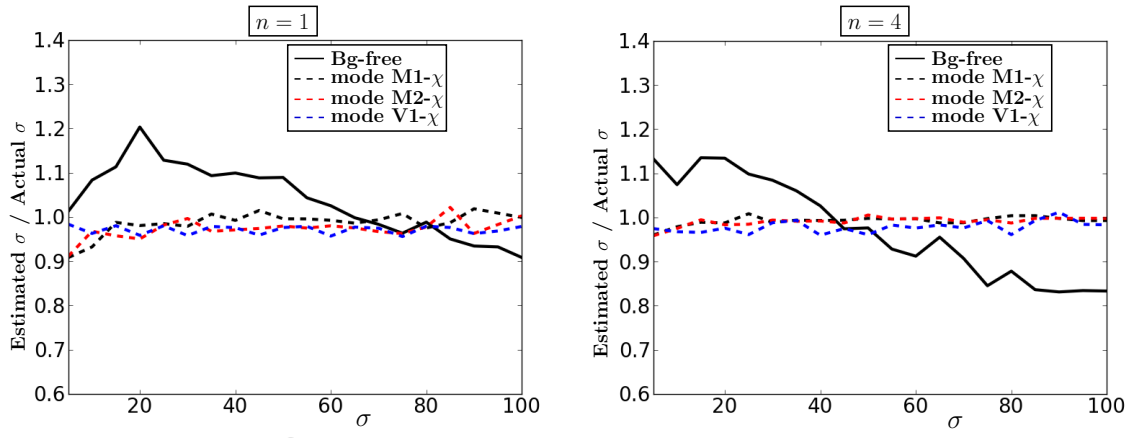


Figure 5.13: The ratio of the estimated σ over the actual σ for various values of σ ranging from 10 to 100. The estimations were performed on the BrainWeb simulated T_1-weighted image with an addition of stationary nc-χ noise with $n = 4$. We used the "Bg-free" technique on the image without its background, as well as the "mode M1-χ", "mode M2-χ" and "mode V1-χ" techniques on the segmented background.

Nevertheless, we have to keep in mind that the "Bg-free" method, as well as the "mode M1-χ", "mode M2-χ" and "mode V1-χ" techniques estimate a global noise variance for the whole image. Therefore, they do not account for the noise non-stationarity, whereas, in practice, with a multiple-channel acquisition, noise is non-stationary in most cases as explained previously. Then, the most accurate method to use is the one presented previously in the paragraph page 124, which was introduced and evaluated in [Aja-Fernández et al. (2013)]. It was shown that, in case of noise non-stationarity, this estimated voxel-wise effective noise variance was better than a global estimated variance and improved a lot some noise correction algorithms which rely on the variance value.

We will now present some indicators to assess the quality of an image.

5.3 Indices of image quality

In this section, we present some image quality indices. Our purpose was not to make a complete review of all indices, therefore we only introduce some classical ones. We first focus on some popular classical indices, and then we present two dMRI-specific indices, which we precisely developed for this thesis work.

5.3.1 Conventional indices

Among the classical indices that we present here, some require to have the knowledge of the noise-free volume. Others can be evaluated with the measured (therefore noisy) volume only.

Mean square error (MSE)

The mean square error (MSE) is based on the calculation of differences between the noisy volume and the noise-free volume. This index gives the average of the squared errors between an image volume $\mathcal{V}$ and its noise-free reference. Both volumes have the same size. The MSE is computed such that:

$$\text{MSE} = \frac{1}{N_v} \sum_{\mathbf{v} \in \mathcal{V}} \left(M(\mathbf{v}) - S(\mathbf{v}) \right)^2 , \tag{5.37}$$

with N_v the number of voxels in the volume $\mathcal{V}$. $M(\mathbf{v})$ and $S(\mathbf{v})$ are the intensity values at the voxel $\mathbf{v}$ of the volume, whose quality is assessed, and of the noise-free volume reference, respectively. The MSE therefore indicates how two image volumes are voxel-wise similar.

We want to underline here that there are other possible indices using a noise-free reference which we did not use in this thesis work because of a lack of time: the mean structural similarity index matrix [Wang et al. (2004a)] and the quality index based on local variance [Aja-Fernández et al. (2006)] among others.

Signal-to-noise ratio (SNR)

The SNR, which compares the noise-free signal level with the noise level, is defined at each voxel of the image volume such that [Constantinides et al. (1997)]:

$$\text{SNR}(\mathbf{v}) = \frac{S(\mathbf{v})}{\sigma(\mathbf{v})}, \tag{5.38}$$

with $S(\mathbf{v})$ the noise-free signal and $\sigma(\mathbf{v})$ the noise standard deviation, both defined at the voxel $\mathbf{v}$. In SoS reconstructions, $\sigma(\mathbf{v})$ has to be replaced by $\sigma_{eff}(\mathbf{v})$ to account for eventual correlations between the noise signals from the n channels.

This SNR definition can be applied in a local noise stationary region of interest (ROI), with $S(\mathbf{v})$ and $\sigma(\mathbf{v})$ replaced by their means. It can also be extended to an estimated SNR, written as $\widehat{\text{SNR}}$ and defined on the measured data only, such that:

$$\widehat{\text{SNR}}_{\text{ROI}} = \frac{\langle M(\mathbf{v}) \rangle_{\text{ROI}}}{\langle \sigma(\mathbf{v}) \rangle_{\text{ROI}}}, \tag{5.39}$$

with $M(\mathbf{v})$ the measured magnitude signal and $\langle \cdot \rangle_{\text{ROI}}$ the expectation operator defined in a unique local noise stationary ROI of the volume.

Many researcher teams proposed different ways to estimate the SNR on the measured data. A review of SNR measurement methods can be found in [Dietrich et al. (2007); Goerner and Clarke (2011)] for the case of multiple-channel acquisition. At the time of these reviews, the authors were not aware of the voxel-wise variance estimation method proposed in [Aja-Fernández et al. (2013)]. Therefore, they did not propose to estimate the SNR following eq. 5.39. These reviews condamn the SNR estimation methods that use separate regions to estimate the signal and noise levels. Indeed, such methods do not account for the fact that the noise standard deviation is voxel-dependent, as well as signal-dependent. Because of these spatial noise variations in the image, such methods are not

accurate. Among the other more accurate methods proposed, some (called "SNR_{mult}" in [Dietrich et al. (2007)]) require many repetitions of images acquired with the same imaging parameters. This is not easily achieved in practice, in particular for dMRI sequences, which last an already long time. Other methods propose to estimate the SNR only in high-signal regions, assuming a Gaussian noise ("SNR_{diff}" in [Dietrich et al. (2007)]). However, we think that, if it is interesting to have a measurement of one of the highest SNR in the image, it is also interesting to measure the SNR in medium to low signal regions. Indeed, in dMRI, these latter regions are more likely to be found, especially at high *b-values*. But, it must also be highlighted that it is very difficult to define an ROI in dMRI images obtained at high *b-values*: the signals are so noisy that the structures disappear on a DW image, as can be seen in fig. 5.7, page 120.

Contrast-to-noise ratio (CNR)

The contrast-to-noise ratio (CNR) brings a complementary information in addition to the SNR. Indeed it indicates how easy or not it is to differentiate structural regions in the volume. It gives a measurement of the contrast between two regions A and B. It is defined such that [Haacke et al. (1999)]:

$$\text{CNR} = |\text{SNR}_\text{A} - \text{SNR}_\text{B}|, \tag{5.40}$$

with SNR_A and SNR_B the SNR calculated for the regions A and B, respectively.

As previously for the SNR, the CNR can be estimated on the measured data, using:

$$\widehat{\text{CNR}} = |\widehat{\text{SNR}}_\text{A} - \widehat{\text{SNR}}_\text{B}|, \tag{5.41}$$

with $\widehat{\text{SNR}}_\text{A}$ and $\widehat{\text{SNR}}_\text{B}$ the estimated SNR for the regions A and B, respectively.

As for the SNR, the CNR index is difficult to get on DW images, for which the structures are not easily distinguishable. That is why we developed some dMRI-specific indices.

5.3.2 dMRI-specific indices

We defined dMRI-specific indices to evaluate the noise impact on DW images.

DTI, aQBI and sa-aQBI MSE

We adapted the definition of the MSE to the DTI, aQBI and sa-aQBI data. This definition relies on the vector $\mathbf{x}$, which is either equal to the diffusion tensor $\mathbf{D}$ for DTI data, the coefficients' vector $\mathbf{C}^{\text{DW}}$ for aQBI data or the coefficients' vector $\mathbf{C}^{\text{SA}}$ for sa-aQBI data:

$$\text{MSE} = \frac{1}{N_v} \sum_{\mathbf{v}\in\mathcal{V}} \sum_{j=1}^{N} \left(\tilde{x}(j,\mathbf{v}) - x(j,\mathbf{v})\right)^2, \tag{5.42}$$

with $\tilde{x}(j,\mathbf{v})$ the j^{th} coefficient of the noisy vector $\tilde{\mathbf{x}}$ calculated on the noisy DW data. $x(j,\mathbf{v})$ is the j^{th} coefficient of the noise-free vector $\mathbf{x}$. This diffusion-specific MSE is faster

to compute compared to the classical MSE given by eq. 5.37. Indeed, there are always less coefficients than DW measurements: $N < K$, K being the number of orientations leading to K different values of $M(i, \mathbf{v})$ with $i \in [\![1; K]\!]$.

FA, GFA and cGFA ratios

The previous image quality index requires the knowledge of the noise-free reference image. It is ideal when working with simulated data, but it cannot be used on real data. That is why we chose to develop an new index which performs a measurement that only depends on the raw (and therefore noisy) data. This new index is an FA, GFA or cGFA ratio which is intuitively similar to a CNR index. However, in opposition to the SNR or CNR index, it is performed on a diffusion derived map (such as the GFA map), on which structures can be distinguished, which is not the case on the simple DW images. It is defined for the aQBI data such that:

$$\text{GFA ratio} = \frac{\text{GFA}_\text{A}}{\text{GFA}_\text{B}}, \tag{5.43}$$

with GFA_A and GFA_B the GFA means calculated on the GFA map for the regions A and B, respectively. Similarly, the FA and cGFA ratios can be computed with the corresponding means.

In contrary to the SNR or CNR indices, this ratio does not contain any noise level. Indeed, the noise variance estimated on the DW image cannot be used on a diffusion derived map such as the GFA. The absence of any noise variance in the index gives the following advantage that this ratio can be easily used on filtered images, for which the noise statistics is usually unknown. Else, the presence of the noise variance in the index would have required to correctly estimate the noise statistics in the filtered image to rightly obtain its noise variance. There is however one requirement for this index: it is to know two ROI whose GFA (or FA, or cGFA depending on the diffusion model) are different. On real data, it is for example advised to use for ROI A a region of medium anisotropy (like in a fiber bundle or a fiber bundle crossing) and for ROI B a region of low anisotropy (like in the CSF). Then, the higher the noise level on the image, the lower the index.

Later in the manuscript, we will use this index to compare the quality of raw and noise corrected diffusion maps. For now, we first need to evaluate this new index in comparing it to the previous ones.

5.3.3 Comparison between the indices on a simulated DW field

We evaluated the new index on a simulated DW volume as the one at $b = 4500s \cdot mm^{-2}$ described in the appendix C at the end of the manuscript. We added stationary nc-χ noise with $n = 4$ and with different values of σ on this DW volume. To perform our evaluation, we measured:

- the MSE developed for the aQBI using eq. 5.42,

- the SNR in ROI A on the *noise-free* DW image shown in fig. 5.14,
- the CNR between the ROI A and B on the *noise-free* DW image shown in fig. 5.14,
- the GFA ratio between the ROI A and B on the *noisy* GFA image.

The results are shown in fig. 5.14 for several values of noise standard deviation. As expected, the aQBI MSE increases with the noise level: the noisier the image, the less it looks like the noise-free reference. The SNR decreases with the noise variance, also as expected. Finally, the higher the noise level, the less distinguishable are ROI A and ROI B: the CNR decreases with σ, as expected. Concerning now our proposed index, the GFA ratio measured directly on the noisy GFA, it also decreases with σ. This new index expresses the loss of contrast on the GFA map between the two ROI. Fig. 5.14 demonstrates that it is a good image quality indicator on the simulated data.

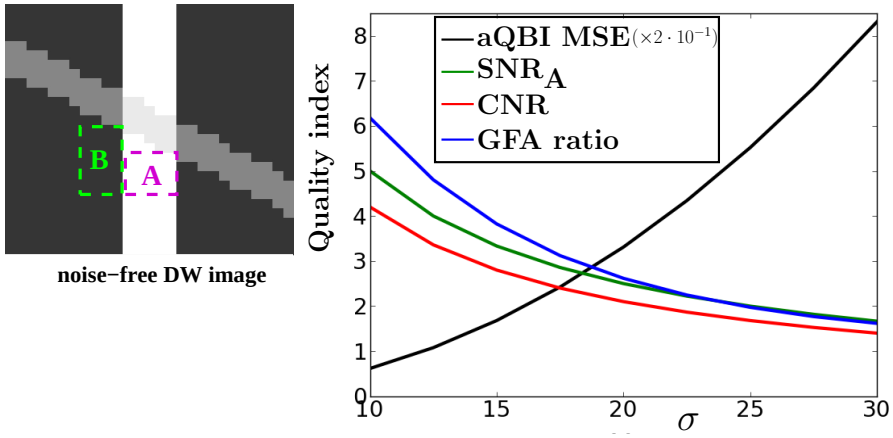


Figure 5.14: Comparison between the GFA ratio and other image quality indices, such as the MSE adapted to the aQBI model, the SNR and the CNR. The comparison was performed on a simulated DW volume with an addition of stationary nc-χ noise with $n = 4$ and with different values of σ.

The results here were shown using the aQBI model. Similar results were obtained using the DTI and sa-aQBI models.

5.4 Conclusion of this chapter

We presented in this chapter the knowledge to date on noise nature and noise models in MRI. We detailed different methods to estimate noise. Among them, one arises from this thesis work and enables a global variance estimation on images without any background, in case of stationary noise. However, noise is in pratice non-stationary and therefore requires a voxel-wise estimation, as the one we explained from the literature. Finally, we presented several indicators to assess the image quality. Among them, we adapted the MSE for dMRI data. We also proposed the GFA (or FA, or cGFA) ratio, which can be calculated on real data, without a noise-free reference. The contributions of our work are listed below.

5.4.1 Contributions of this chapter:

- A detailed state-of-the-art review on the statistical distribution of the measured MRI signal.
- A new noise estimation method adapted to the nc-χ PDF of the measured MRI signal. This method does not necessitate any background in the data, which can be a determinant point when working with some data that do not have enough background. It does however not produce as good results as methods relying on the background knowledge. The technique is valid for stationary noise only. This contribution appears in [Brion et al. (2011b)].
- A new image quality index, the GFA (or FA, or cGFA) ratio, which is specific for dMRI images. This index can be measured without any noise-free reference. Moreover, it can be used on images, for which the noise statistics are unknown (like in filtered images).

In this chapter, we emphasized the impact of noise in dMRI, especially at high *b-values*. We showed the nc-χ bias, which appears at low SNR values and high n values. To be removed, this bias requires a correction adapted to nc-χ noise. That is why a correction accounting for the noise statistics explained in this chapter is necessary to obtain an accurate and reliable information on the anatomical structures presented in the data. Also, it is important to keep in mind our RT objective. The next chapter establishes the issue regarding noise correction for rtdMRI.

Chapter 6

Correcting noise in RT

As we introduced it in chapter 4 and 5, dMRI is generally significantly corrupted by noise unfortunately characterized by a non-Gaussian distribution at high *b-values*. There is a clear interest for developing a framework to compute DTI, HARDI and HYDI in RT and we have shown that it is possible using incremental solvers such as Kalman filters. However, such filters are designed to deal with a zero-mean Gaussian noise distribution and are therefore biased in dMRI. In this chapter, we will first present a detailed state-of-the-art about noise correction methods and we will propose a method to correct the non-Gaussian noise of dMRI that is compatible with RT objectives.

6.1 Constraints stemming from RT

In this section, we will first attempt to summarize most of the existing denoising methods which have been introduced in the field of MRI. Then, we will discuss them in light of an RT requirement.

6.1.1 State-of-the-art of the noise correction methods

Today, there is a plethora of noise correction techniques for application to MR datasets. These numerous techniques usually rely on well-known denoising concepts. We attempted to classify them in groups indicated in fig. 6.1, but some of them combine methods stemming from different groups. We want to highlight here that we cannot exhaustively detail all the denoising techniques, as there are too many, but we will focus on the most important methods quite often referenced in the literature. First methods were introduced in 1985 and there is still active research today to deal with the complex noise model characterizing the magnitude signal of the most recent parallel imaging techniques.

MRI denoising techniques can be separated into two categories: a first category containing methods which do not account for the specific Rician or nc-χ noise model. Such methods usually assume a Gaussian noise. The second category incorporates the noise distribution model in the denoising method. For this last group of techniques, and concerning the dMRI dedicated methods, they can be distinguished whether they incorporate

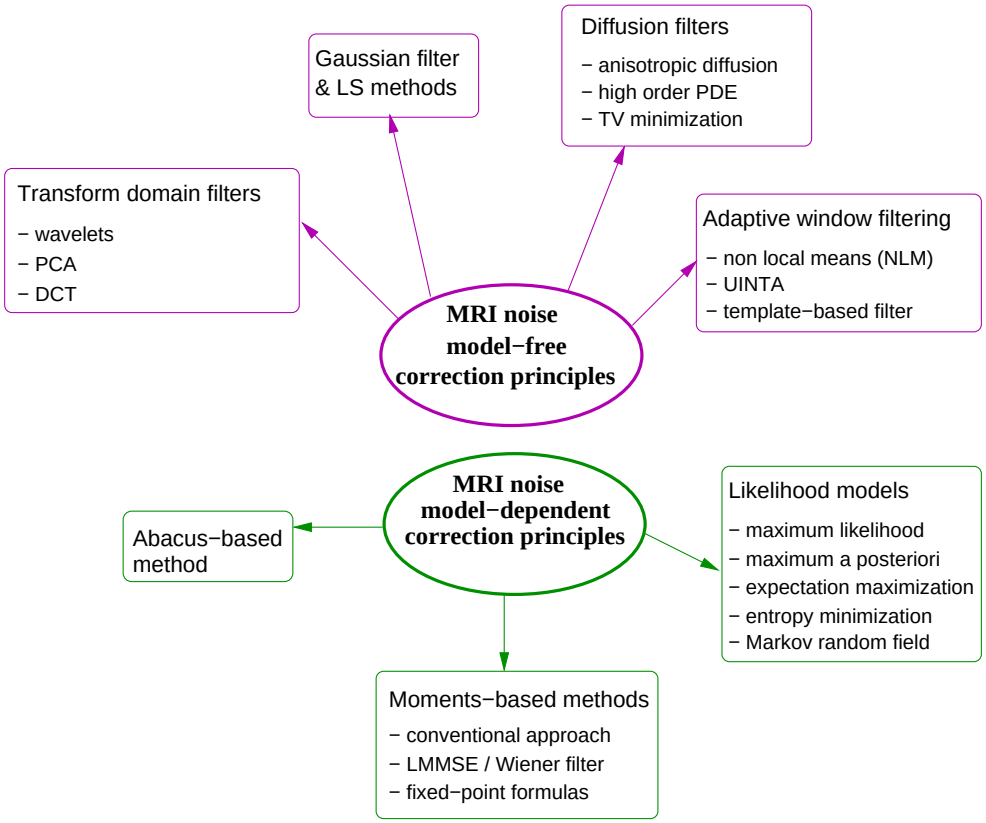


Figure 6.1: Main groups of noise correction methods with an application to MRI images. These methods are either based on a correction principle relying on the MRI Rician or nc-χ noise model (green), or not (magenta).

the noise model directly in the diffusion propagator estimate or in the propagator-derived functions like the dODF, or whether they only correct the DW signal. We will now make a review of families of methods relying on a specific denoising concept. We will first present the techniques, for which the original denoising concept does not account for the MRI noise model.

Principles of noise correction in MRI using model-free approaches.

Transform domain filters. The group of transform domain filters gathers techniques processing images in a transformed domain. Such a domain provides a multiscale representation of an image, for which the lower scales contain the basic components of the image and the higher scales contain the details and boundaries between different structural regions of the image. The idea of transform domain filters is to keep the scales, which are lower than a certain threshold, so that the noise, which is contained in the higher scales, is removed. Well-known examples of such methods are wavelet-based approaches [Healy and Weaver (1992); Donoho and Johnstone (1994); Hilton et al. (1996); Wood and Johnson (1999); Nowak (1999); Pižurica et al. (2003); Delakis et al. (2007)]. The wavelet coeffi-

cients with high SNR are kept, whereas the coefficients with low SNR containing mostly noise are discarded. The wavelet-based method is an edge-preserving technique [Nowak (1999)], however it can produce artifacts due to the structure of the underlying wavelets [Awate and Whitaker (2007); Manjón et al. (2008)]. For this technique, which does not initially account for the MRI noise model, some methods were developed to incorporate the Rician MRI noise nature [Nowak (1999); Wood and Johnson (1999)]. Others included a local noise variance estimation, which accounts for the non-stationarity of this parameter in the case of parallel MRI [Delakis et al. (2007)]. Another transform domain filter was proposed in [Muresan and Parks (2003)], based on the use of a principal component analysis (PCA), assuming a white noise model. The principal components provide a local adapted basis used to decompose the signal, followed by a thresholding to select only the principal wavelets. Finally, the use of the discrete cosine transform (DCT) was also proposed. [Yaroslavsky et al. (2001); Guleryuz (2003, 2007); Manjón et al. (2012)].

Gaussian filter & least squares (LS) methods. This group contains techniques relying on a Gaussian filter [Lindenbaum et al. (1994); Ashburner and Friston (2000)] which present some risk of a final blurring result. The other filters of this group are based on many kinds of least squares (LS) methods, which can be linear or not, constrained or not and also weighted or not: [Chang et al. (2005); Salvador et al. (2005); Koay et al. (2006); Tristán-Vega et al. (2012)]. These methods directly work on the DTI model and are not adapted to other local diffusion models. Because of the simplicity behind the LS algorithm, such techniques are still commonly used. However, it is highlighted in [Tristán-Vega et al. (2012)] that they are not efficient in medium-to-low SNR situations, where the noise bias due to the Rician or to the nc-χ noise nature is not correctly removed.

Diffusion filters. Diffusion filters to remove noise are based on anisotropic diffusion equations [Perona and Malik (1990); Catté et al. (1992); Gerig et al. (1992); Yang et al. (1995); Alvarez (1996); Sapiro and Ringach (1996); Black et al. (1998); Parker et al. (2000); Ling and Bovik (2002); Samsonov and Johnson (2004); Chen and Hsu (2005); Basu et al. (2006); Tschumperlé and Brun (2009); Krissian and Aja-Fernández (2009)]. The diffusion equation, originally proposed by [Perona and Malik (1990)], relies on the idea to apply the diffusion process on an image with a spatially varying diffusion coefficient. The latter is chosen to encourage intra-region smoothing and to limit inter-region smoothing. A good tuning of this coefficient enables to avoid the blurring of the edges contained in the image. To control the smoothing in boundaries' regions, the diffusion coefficient is expressed as a function of the image gradient. A threshold-parameter is required to control the sensitivity to the egdes of the image. This parameter is either determined experimentally or calculated as a function of the noise in the image. The use of the anisotropic diffusion does not incorporate the knowledge of the noise distribution. It was shown in [Sijbers et al. (1999a)], that there is a noticeable advantage in incorporating this distribution knowledge. A comparison between the anisotropic diffusion filter of [Yang et al.

(1995)] and the same method with an additional maximum likelihood estimation of the filter parameters accounting for the Rician noise nature showed that the second method outperformed the first one. Among the previous techniques, the method by [Samsonov and Johnson (2004)] proposes a local estimation of the noise variance to cope with the non-stationarity of this parameter due to parallel MRI. Other diffusion filters were also developed aiming at minimizing the total variation (TV) norm of the image [Rudin et al. (1992); Strong and Chan (1996); McGraw et al. (2004)]. The underlying idea is that the TV norm, which corresponds to the integral of the absolute image gradient, contains the details of the image. As noise corresponds to the spurious details, minimizing the TV norm of the image will help reducing the noise. However, such methods may generate staircase effects yielding gradual contrast variations in homogeneous regions of the image. The TV minimization technique leads to the resolution of a partial differential equation (PDE). Other filters based on higher order PDEs were developed to outperform the performances of the TV minimization technique and the anisotropic diffusion equation method [Lysaker et al. (2003); Greer and Bertozzi (2004); Kim and Lim (2009); Liu et al. (2011)]. We can also cite here some variational methods using such PDEs for directly regularizing the DTI or HARDI calculations [Chefd'Hotel et al. (2002); Assemlal et al. (2007); Wang et al. (2004b)].

Adaptive window filtering. The last group of denoising methods which relies on a correction concept that does not account for the noise distribution is the group of adaptive window filters. Such estimators are based on the choice of an adapted window containg neighborhing information to improve noise removal. Among these filters, the template-based filter of [Ahn et al. (1999); Guo et al. (2006)] proposes to test different templates (or windows) of different sizes and geometries around the voxel to correct and to choose the biggest template containing the voxels of interest representing a locally constant region. The method therefore excludes any template containing edges. The underlying idea of such a concept is similar to the diffusion filters that stop the diffusion process at edges to keep the structural boundaries intact in the image. Other adaptive window filters search the optimal neighbors accross the entire image (or in a very large part of it). To find similar structural neighbors to the voxel of interest, the method compares the local patch of voxels around the neighbor with the patch of the voxel of interest. The choice of the optimal neighbors is done using a similarity criterion between the patches. Such a concept taking advantage of the recurrences inside an image appeared in the non local means (NLM) filter of [Buades et al. (2005)] and in the unsupervised, information-theoretic, adaptive (UINTA) image filtering for image restoration of [Awate and Whitaker (2006)]. The NLM filter was extended in many other works, which incorporated the Rician nature of the MRI noise [Manjón et al. (2008); Wiest-Daesslé et al. (2008); Descoteaux et al. (2008)]. A paper additionally proposed to account for the spatially varying noise variance with the use of parallel MRI [Manjón et al. (2010)]. Other approaches attempted to reduce the heavy computational cost of the filter [Coupé et al. (2008b); Tristán-Vega et al. (2012)].

Finally, this NLM filter was also combined with other denoising or optimizing methods [Coupé et al. (2008a); Tristán-Vega and Aja-Fernández (2010); Hong et al. (2012); Manjón et al. (2012); He and Greenshields (2009); Rajan et al. (2011, 2012a,b)]. A unique NLM method by [Aja-Fernández et al. (2013)] was developed to account for the nc-χ noise model. Additionally, it relies on the use of the effective parameters n_{eff} and σ_{eff} detailed in the previous chapter, page 117), and therefore accounts more precisely for the non-stationarity of the MRI noise, in case of an SoS reconstruction without subsampling of the **k**-space.

MRI noise model-dependent correction principles.

Abacus-based method. Among the noise removal techniques relying on the MRI noise distribution, the abacus-based method introduced by [Henkelman (1985)] is actually the first historical technique proposed to correct for noise in MRI images. This method consists of using a lookup table (LUT) correction scheme accounting for the Rician bias. This technique can be seen as the first work which attempts to evaluate and extract the noise bias from the MRI measurements.

Moments-based methods. The second historical MRI denoising method came out with the conventional approach based on the expression of the Rician 2^{nd} order moment [McGibney and Smith (1993); Miller and Joseph (1993)]. A technique very close to the two latter, with an additional absolute value operator, was developed by [Gudbjartsson and Patz (1995)]. All these techniques do not perform so well, but are often referred as historical MRI denoising methods. More elaborated techniques also rely on the knowledge of the moments of the measured magnitude M. The linear minimum mean square error (LMMSE) estimator, which was adapted for MRI Rician noise removal by [Aja-Fernández et al. (2008a,b)], is based on the 2^{nd} and 4^{th} order moments of M. It presents a closed-form solution, which does not require any iterative optimization procedure. Therefore, the LMMSE is a computationally straightforward and efficient technique. A DTI dedicated anisotropic extension of the LMMSE was combined with a maximum likelihood technique permitting a spatially varying noise level estimation to account for the variance non-stationarity in case of parallel imaging [Caan et al. (2010)]. Another improvement of the original LMMSE was proposed in [Tristán-Vega and Aja-Fernández (2010)] which consisted in using simultaneously the joint information contained in the DW volumes along the different diffusion orientations. Such a technique is not consistent with an RT denoising. We will come back to it later. Finally, an extension of the LMMSE to nc-χ noise was also proposed in [Brion et al. (2011b)] with further improvements yielding anisotropic schemes [Brion et al. (2011f); Vegas-Sánchez-Ferrero et al. (2012); Casaseca-de-la-Higuera et al. (2012)]. Later, the LMMSE was proposed to be used with the effective parameters n_{eff} and σ_{eff} to account for the noise correlations in parallel MRI, as explained in the previous chapter. This last technique is adapted to the case of an SoS reconstruction without subsampling in the **k**-space [Aja-Fernández et al. (2013)]. As the Wiener filter is a sequential version of the LMMSE, we can also cite in this group the method by [Martin-

Fernández et al. (2009)], which relies on a Wiener filter incorporating the Rician noise characteristics. A third group of moments-based methods concerns the techniques using fixed point formulas of either the SNR or the noise-free magnitude S [Koay and Basser (2006); Koay et al. (2009a)]. A fixed point formula is a relation of equality between a parameter and a function of this parameter. Finding the solution of the fixed-point formula of the SNR of an MRI image enables to obtain the noise-free MRI image [Koay and Basser (2006)]. Based on the fixed-point formula of the noise-free magnitude S, a framework was derived to transform the magnitude signals so that they follow Gaussian distributions, allowing the use of least squares approaches [Koay et al. (2009a)]. These two methods were adapted for a Rician noise, as well as for a nc-χ noise. In practice, they require either many repetitions of the measurements or a very large number of measurements performed at different *b-values*, which are conditions that are unaffordable in clinical routine, as they would make the scan far too long.

Likelihood models. The methods relying on a likelihood model also contain this powerful advantage of incorporating the noise model, making them more adapted to MRI denoising than other methods. Several techniques making use of the likelihood models have been investigated. The maximum likelihood (ML) technique provides a noise-free estimate, which maximizes the likelihood function of a sample of independent observations. It was adapted to the Rician properties of the MRI noise in [Sijbers et al. (1998)] and was shown to yield accurate results at low SNRs [Sijbers et al. (1999b)], compared to other methods, which do not account for the specific Rician noise nature. The ML method was also used in [Sijbers and den Dekker (2004); Haldar et al. (2012)], directly in the Fourier domain on the complex data before the SoS reconstruction. A simple Gaussian noise could then be assumed for these two methods, as noise is Gaussian before the SoS recombination. A comparison was performed in [Sijbers and den Dekker (2004)] between the ML technique applied either on the complex data containing Gaussian noise, or on the magnitude data containing Rician noise: in the common case of phase artifacts on the real and imaginary channels, the technique performed on the magnitude data was shown to yield better results. Many other works were proposed to inject the ML method in some local diffusion models and to denoise simultaneously the set of DW data along different diffusion orientations [Fillard et al. (2007); Clarke et al. (2008); Brion et al. (2009)]. Such methods take advantage of the joint information contained in the DW data set, but are dedicated to a specific local diffusion model. Finally, other methods combine the ML technique with a NLM filter to get the best of both methods [He and Greenshields (2009); Rajan et al. (2011, 2012a,b)]. Bayesian methods like the maximum a posteriori (MAP) techniques, which maximize a posterior probability describing the data, can also be used for denoising [Basu et al. (2006); Andersson (2008); Lam et al. (2012)]. To iteratively find the ML or the MAP estimates of parameters, an expectation-maximization can be used [Marzetta (1995)]. Other denoising methods based on likelihood models rely on the Markov random field properties of the data [Liévin et al. (2002); Awate and Whitaker

(2007)]. Sometimes, a regularization using Markov random fields is directly performed on the noisy diffusion tensors [Martin-Fernández et al. (2004); Poupon et al. (2000)]. Finally, some Markov random field based techniques were combined with an entropy minimization algorithm for improved performances [Awate and Whitaker (2005, 2006)].

Characteristics of the denoising methods.

Among this huge list of denoising methods, it is important to choose an adequate technique with respect to a particular situation. To this aim, we propose here to summarize the key features of these techniques, as shown in fig. 6.2.

Different characteristics to look at concerning MRI noise correction techniques	**Examples**	**Counterexamples**
applied before the SoS reconstruction using the zero-mean Gaussian characteristics of MRI noise	[Sijbers & den Dekker, 2004] [Haldar et al., 2012]	all other cited methods
Rician noise dedicated	[Nowak, 1999] [Brion et al., 2009]...	[Perona & Malik, 1990] [Healy & Weaver, 1992]...
nc-χ noise dedicated	[Koay & Basser, 2006] [Rajan et al., 2012]...	most of other cited methods
with a spatially varying noise variance	[Samsonov & Johnson, 2004] [Delakis et al., 2007]...	most of other cited methods
accounting for the nc-χ statistical model proposed by [Aja-Fernández et al.,2011,2012]	[Aja-Fernández et al., 2012]	all other cited methods
dMRI specific method	[Martin-Fernández et al., 2004] [Koay et al., 2009]...	[Sijbers et al., 1998] [Buades et al., 2005]...
processing the DW data along all orientations together	[Fillard et al., 2007] [Tristán-Vega et al., 2010]...	[Aja-Fernández et al., 2008a,b] [Manjón et al., 2008]...
local diffusion model dependent (DTI, aQBI, SDT...)	[Chen & Hsu, 2005] [Clarke et al., 2007]...	[Parker et al., 2000] [Basu et al., 2006]...
requiring other than the typical dMRI settings	[Koay & Basser, 2006] [Koay et al., 2009]...	[Koay et al., 2006]...
without critical parameter tuning	[Liévin et al., 2002] [Awate & Whitaker,2005,2007]...	[Perona & Malik, 1990] [Lysaker et al., 2003]...
compared to the state-of-the-art methods	[Buades et al., 2005] [Aja-Fernández et al., 2008a,b]...	[Gudbjartsson & Patz, 1995]...
processing slices (at one orientation) independently (for a possible distribution on a cluster of CPU)	[Aja-Fernández et al., 2012]...	[Manjón et al., 2010]...

Figure 6.2: Some characteristics to look at concerning MRI noise correction methods. The characteristics in the yellow frame concern how much the method accounts for the true noise statistics. The characteristics in the pink frame are specific to dMRI data denoising.

The characteristics in yellow in fig. 6.2 concern how much the method accounts for the true noise statistics. In our previous review, most of the MRI denoising methods are adapted to Rician noise, but to our knowledge, none of them are adapted to nc-χ noise distributions that characterize the noise present in the SoS GRAPPA technique described

in the previous chapter. We are particularly interested in this acquisition case, because our real data were acquired with a multiple-channel acquisition and reconstructed with the SoS GRAPPA algorithm. The techniques approaching the most the nc-χ noise model proposed for the SoS GRAPPA technique in [Aja-Fernández et al. (2011)] (see in the previous chapter, page 117) are the ones, which incorporate the nc-χ distribution knowledge, and additionally account for a spatially varying noise variance across the image, and more precisely, account for the non-stationnarities of both the effective noise variance and the effective number of channels. The methods in [Aja-Fernández et al. (2013)] propose such a framework, but the latter is adapted to an SoS reconstruction without subsampling in the **k**-space.

Among the review of noise removal methods, two techniques [Sijbers and den Dekker (2004); Haldar et al. (2012)] directly work in the Fourier domain on the complex data and therefore use the zero-mean Gaussian characteristics of the noise. Indeed, this choice to denoise the data before the SoS reconstruction on each real and imaginary data is very clever, as the noise is then simply zero-mean Gaussian. However, this choice was not usually performed in the denoising community. This can be explained by the fact that the real and imaginary images may not be systematically available [Nowak (1999); Lam et al. (2012)]. Another reason is that it is better to process the magnitude because of artifacts in the phase images [Henkelman (1985); Constantinides et al. (1997); Nowak (1999); Sijbers and den Dekker (2004)]. Therefore, the majority of the correction techniques were developed for an application after the SoS reconstruction on the magnitude images. Consequently, they have to account for a nc-χ noise, as explained in the previous chapter. We also deliberately made this choice to develop an RT denoising technique. But we shall keep in mind that it is also possible to denoise the real and imaginary data from zero-mean Gaussian noise.

Other characteristics of MRI denoising methods, shown in pink in fig. 6.2, are specific to dMRI data denoising. The method can for example be specific towards the dMRI modality. Such dMRI specific method may process all DW data along all diffusion orientations together, taking advantage of the joint information. Some are dependent on the local diffusion model and apply directly on the model. In that case, the denoising method is not generic towards the different diffusion models. Finally, some techniques require unusual dMRI settings, like for example many different *b-values* as in [Koay et al. (2009a)]. Denoising methods can also be distinguished with respect to the necessity or not of a critical parameter tuning by the user. The ideal method should adapt itself to any kind of images without the requirement of a parameter tuning. Some methods tend to this ideal: [Liévin et al. (2002); Awate and Whitaker (2005, 2007)]. A good technique has usually been compared to the state-of-the-art methods to prove its relevance. Finally, when wanting to distribute the computation on a cluster of CPU, it is necessary that the technique processes slices (or small groups of slices) independently. Indeed, the more the data volume is split into small volumes to process independently, the more time will be saved during the denoising task.

To conclude with this state-of-the-art of the noise correction methods, we want to highlight that no gold standard technique has really been brought to light in the past years, although some techniques like the NLM are quite often cited in the community. Moreover, no RT noise removal had ever been investigated until recently, with some works emanating from this thesis [Brion et al. (2010, 2011f,d, 2012a)], followed by others [Casaseca-de-la-Higuera et al. (2012)], as we will see in the next chapter. Before, it is essential to explain that, for choosing the most appropriate method for our RT denoising purpose, we looked at the contraints imposed by our willing to perform the correction in RT.

6.1.2 Constraints imposed by RT

Our goal was to perform a noise correction, which would be easily integrated in the rtdMRI framework environment described in chapter 4. The desired method has to give noise corrected results before each new measurement during the ongoing scan. To obtain this RT correction, four **conditions** are mandatory:

1. the method should not require the knowledge of all the DW dataset, but only a subset of it,
2. the method should address as many local diffusion models as possible,
3. the method should not depend on other measurements than the DW measurements and be usable with clinical protocols,
4. the method should compute any intermediate result in less than the repetition time T_R (usually around 10s), and should therefore be very fast.

With these requirements, it is straightforward to see that some previously mentioned methods cannot be used for our purpose. Indeed, all techniques processing the orientations together, like [Fillard et al. (2007); Clarke et al. (2008); Brion et al. (2009); Martin-Fernández et al. (2009); Tristán-Vega and Aja-Fernández (2010)], do not satisfy the first condition. To respect the second condition, we excluded all methods, which apply to only one diffusion local model, like [Chefd'Hotel et al. (2002); Wang et al. (2004b); Chang et al. (2005); Salvador et al. (2005); Koay et al. (2006); Assemlal et al. (2007); Caan et al. (2010); Tristán-Vega et al. (2012)]. Concerning the third condition, we also excluded [Koay and Basser (2006); Koay et al. (2009a)], which require extra measurements for the denoising process. Finally, the fourth condition invited us to look at the computational efficiency of the method, as well as its compatibility with a distribution over the nodes of a cluster. It is essential to understand here that the time parameter in the method is not only a criterion to look at after having been assured that the technique performs well. It is a point to consider in the same time as when evaluating the denoising performances of the technique. To respect this last condition, we decided to work with the LMMSE estimator —referred simply as LMMSE— of [Aja-Fernández et al. (2008a,b)], originally developed for a Rician noise and shown to outperform, in terms of time and quality performances,

other methods. An also very good —maybe better?— technique, the NLM filter relies on a weighted mean of voxels inside a large and therefore non local search window. The weight of a voxel is computed using the similarity of the patch surrounding this voxel and the patch surrounding the voxel of interest. This technique had not been considered in our choice because of its too long computational time. Indeed, this last method is quite long as it requires to compare neighborhoods across the entire data volume. Its computation time is of several hours for a 181×217×181 volume on a 3 GHz CPU [Coupé et al. (2008b)]. But, the optimized version developed by [Coupé et al. (2008b)] uses a selection of the most relevant voxels, and a blockwise implementation with some parallelization on a cluster. It reduces the computation time to around one minute for the same volume with 8 CPU of 3 GHz each. More recently, [Tristán-Vega et al. (2012)] proposed a further acceleration with the measurement of the similarity between patches using a truncated local Taylor series expansion. Compared to the method by [Coupé et al. (2008b)], it was shown to be at least 1.7 times faster.

We wanted to address the nc-χ nature of the noise in GRAPPA DW images and consequently, we investigated the possibility to extend the LMMSE approach originally developed under the assumption of a Rician noise to a nc-χ noise. A similar study was also conducted by [Aja-Fernández et al. (2013)] to SoS configurations. We explain in the next section the theoretical basis of the LMMSE and we show afterwards how this method combines the RT compatibility with good performances on our dMRI data.

6.2 Correction with a linear minimum mean square error (LMMSE) estimator

Because RT imposes some strong requirements on the computational aspects of the denoising method, we have decided to focus on the LMMSE method introduced in [Aja-Fernández et al. (2008a)] and [Aja-Fernández et al. (2008b)]. We will present this technique, first for a single-channel acquisition and then we will show the extension for a multiple-channel acquisition with a nc-χ noise. We will detail three particular cases of the LMMSE. Finally, we will propose a trade-off with the empirical estimation of global effective parameters to account for putative noise correlations.

6.2.1 The original LMMSE adapted to Rician distributions

The generic expression for the LMMSE estimator $\hat{\theta}$ of a parameter θ from a vector of data $\mathbf{X}$ is [Kay (1993)]:

$$\hat{\theta} = E(\theta) + \mathrm{Cov}(\theta, \mathbf{X}) \cdot \mathrm{Var}(\mathbf{X})^{-1} \cdot (\mathbf{X} - E(\mathbf{X})) \,, \tag{6.1}$$

where $E(\theta)$ is the expectation of θ, $\mathrm{Cov}(\theta, \mathbf{X})$ is the covariance of θ and $\mathbf{X}$ and $\mathrm{Var}(\mathbf{X})$ is the variance of $\mathbf{X}$. Replacing θ by the squared noise-free MR signal magnitude S^2 and $\mathbf{X}$

by the squared measured magnitude M^2 results in:

$$\hat{S^2} = E(S^2) + \mathrm{Cov}(S^2, M^2)\mathrm{Var}(M^2)^{-1} \times \left(M^2 - E(M^2)\right). \tag{6.2}$$

This expression is true for any voxel and any diffusion orientation; thus, to simplify the notation, we do not add the corresponding indices. Eq. 6.2 allows an estimation of the noise-free magnitude and relies on the knowledge of $E(S^2)$, $\mathrm{Cov}(S^2, M^2)$ and $\mathrm{Var}(M^2)$. The authors of [Aja-Fernández et al. (2008a)] expressed the three latter terms in function of measurable quantities. We explain their approach here. The expression of the variance $\mathrm{Var}(M^2)$ is straightforward:

$$\begin{aligned}\mathrm{Var}(M^2) &= E\left([M^2 - E(M^2)]^2\right), \\ &= E\left(M^4 - 2ME(M^2) + E(M^2)^2\right).\end{aligned}$$

Therefore, using the expectation linearity, we obtain:

$$\boxed{\mathrm{Var}(M^2) = E(M^4) - E(M^2)^2}. \tag{6.3}$$

Now, we detail the approach of [Aja-Fernández et al. (2008a)] to simplify the covariance term $\mathrm{Cov}(S^2, M^2)$. In the case of a single-channel acquisition, the magnitude is written such as: $M = \sqrt{(S_r + \epsilon_r)^2 + (S_i + \epsilon_i)^2}$, with S_r and S_i the real and imaginary parts, respectively, of the noise-free signal $S = \sqrt{S_r^2 + S_i^2}$, and ϵ_r and ϵ_i the real and imaginary parts, respectively, of the noise signal. ϵ_r and ϵ_i are supposed to be uncorrelated, zero-mean Gaussian noises of variance σ^2 [Henkelman (1985)]. Therefore, we can write:

$$\begin{cases} E(\epsilon^k) = 0, \text{if k is odd}, \\ E(\epsilon^2) = \sigma^2, \\ E(\epsilon^4) = 3\sigma^4. \end{cases} \tag{6.4}$$

The imaginary and real parts of the noise signal are supposed to be independent. The same hypothesis is applied between the noise ϵ and the noise-free signal S, thus:

$$\begin{cases} E(\epsilon_r \epsilon_i) = E(\epsilon_r)E(\epsilon_i) = 0, \\ E(\epsilon S) = E(\epsilon)E(S) = 0. \end{cases} \tag{6.5}$$

We can then express the covariance term $\mathrm{Cov}(S^2, M^2)$ of eq. 6.2 such that:

$$\begin{aligned}\mathrm{Cov}(S^2, M^2) &= E\left([S^2 - E(S^2)][M^2 - E(M^2)]\right), \\ &= E\left(S^2M^2 - S^2E(M^2) - E(S^2)M^2 + E(S^2)E(M^2)\right), \\ &= E\left(S^2[(S_r + \epsilon_r)^2 + (S_i + \epsilon_i)^2]\right) - E(S^2)E(M^2), \\ &= E\left(S^2[S_r^2 + S_i^2 + 2S_r\epsilon_r + \epsilon_r^2 + 2S_i\epsilon_i + \epsilon_i^2]\right) - E(S^2)E(M^2).\end{aligned}$$

When simplifying this last line using the equation systems 6.4 and 6.5, we obtain:

$$\boxed{\mathrm{Cov}(S^2, M^2) = E(S^4) + 2E(S^2)\sigma^2 - E(S^2)E(M^2)}. \tag{6.6}$$

We can then rewrite eq. 6.2 using the expressions of the variance and covariance terms given by eq. 6.3 and eq. 6.6, respectively:

$$\hat{S^2} = E(S^2) + \frac{\left(E(S^4) + 2E(S^2)\sigma^2 - E(S^2)E(M^2)\right)}{E(M^4) - E(M^2)^2} \times \left(M^2 - E(M^2)\right). \tag{6.7}$$

Now, it is necessary to express the 2^{nd} and 4^{th} order moments of S in measurable quantities, so that S^2 can be estimated with the LMMSE. In the case of a single-channel acquisition, we saw in chapter 5 that M follows a Rician distribution. We therefore know that:

$$\boxed{E(M^2) = E(S^2) + 2\sigma^2}, \tag{6.8}$$

and:

$$\boxed{E(M^4) = E(S^4) + 8\sigma^2 E(S^2) + 8\sigma^4}. \tag{6.9}$$

These two latter equations can be found, without assuming any Rician distribution, as done in [Tristán-Vega and Aja-Fernández (2010)] for a real noise-free signal. The extension to a complex noise-free signal S can be found in the appendix A (section 6.5), at the end of this chapter. Using eq. 6.8 and 6.9, the 2^{nd} and 4^{th} order moments of S are determined by:

$$\begin{cases} E(S^2) = E(M^2) - 2\sigma^2, \\ E(S^4) = E(M^4) - 8\sigma^2 E(M^2) - 8\sigma^4. \end{cases} \tag{6.10}$$

Finally, using the latter equations' system, the LMMSE eq. 6.7 becomes:

$$\hat{S^2} = E(M^2) - 2\sigma^2 + \left(1 - \frac{4\sigma^2[E(M^2) - \sigma^2]}{E(M^4) - E(M^2)^2}\right) \times \left(M^2 - E(M^2)\right). \tag{6.11}$$

The authors of [Aja-Fernández et al. (2008a)] use the assumption of local ergodicity to replace the expectation $E(\cdot)$ by $\langle\cdot\rangle$ corresponding to a local spatial mean calculated on a neighborhood. Then, the LMMSE in the case of a single-channel acquisition is written such that:

$$\boxed{\hat{S^2} = \underbrace{\langle M^2\rangle - 2\sigma^2}_{J_R} + \underbrace{\left(1 - \frac{4\sigma^2[\langle M^2\rangle - \sigma^2]}{\langle M^4\rangle - \langle M^2\rangle^2}\right)}_{K_R} \times \left(M^2 - \langle M^2\rangle\right)}. \tag{6.12}$$

Eq. 6.12 is the equation of the original LMMSE introduced in [Aja-Fernández et al. (2008a,b)] and is valid for Rician noise only. The J_R term is equal to the conventional estimator introduced in [McGibney and Smith (1993); Miller and Joseph (1993)] which consists of applying a **mean** and a **Rician bias removal** to obtain the noise-free squared magnitude. The parameter K_R is a **data attachment term** that modulates the effect of the J_R term in heterogeneous regions. To resume, the LMMSE acts as a filter allowing some smoothing in homogeneous regions without losing details in the features of heterogeneous regions. Particular cases can appear when the denominator $\langle M^4\rangle - \langle M^2\rangle^2$ in the K_R term

approaches zero or also when the final squared LMMSE estimation $\hat{S}^2$ is negative. These cases do not often appear in practice. They will be deeply discussed in section 6.2.3.

Let us now detail how we extended this LMMSE to nc-χ distributions for the multiple-channel acquisition case (SoS and SoS GRAPPA).

6.2.2 The extended LMMSE adapted to nc-χ distributions

The extension of the LMMSE to nc-χ noise constitutes one contribution of this thesis work [Brion et al. (2011b,c,a)]. To extend the LMMSE estimator to nc-χ distributions, we followed the process documented in [Aja-Fernández et al. (2008b)] with the assumption of a multiple-channel acquisition yielding a nc-χ distribution given by eq. 5.15, page 114, in chapter 5. We demonstrated the LMMSE extension to nc-χ distributions in the assumption of an SoS reconstruction without subsampling, without correlation between the channels and with each channel having the same variance. This simplified case is the one described in chapter 5, page 114. We consider the SoS reconstruction method to obtain the magnitude: $M = \sqrt{\sum_{c=1}^{n}\left[(S_{r_c}+\epsilon_{r_c})^2+(S_{i_c}+\epsilon_{i_c})^2\right]}$, where $S = \sqrt{\sum_{c=1}^{n}\left[S_{r_c}^2+S_{i_c}^2\right]} = \sqrt{\sum_{c=1}^{n}S_c^2}$. The terms S_{r_c} and S_{i_c} are the real and imaginary parts, respectively, of the noise-free complex signal S_c received by the channel c. ϵ_{r_c} and ϵ_{i_c} are the real and imaginary parts, respectively, of the complex noise ϵ_c corrupting the signal received by the channel c. They are assumed to be zero-mean, uncorrelated and independent Gaussian noises. Because of these assumptions, the variance and covariance terms can be written:

$$\begin{cases} \mathrm{Var}(M^2,M^2) = E(M^4) - E(M^2)^2, \\ \mathrm{Cov}(S^2,M^2) = E(S^4) + 2n\sigma^2E(S^2) - E(S^2)E(M^2). \end{cases} \tag{6.13}$$

The derivation of these latter values is detailed in appendix B (section 6.6), at the end of this chapter. Injecting both variance and covariance expressions into eq. 6.2, we obtain:

$$\hat{S}^2 = E(S^2) + \frac{E(S^4) + 2n\sigma^2E(S^2) - E(S^2)E(M^2)}{E(M^4) - E(M^2)^2} \times \left(M^2 - E(M^2)^2\right). \tag{6.14}$$

Now, if we use the 2^{nd} and 4^{th} order moments of the nc-χ distribution —whose demonstration is in appendix C (section 6.7) at the end of this chapter—, we have:

$$\begin{cases} E(S^2) = E(M^2) - 2n\sigma^2, \\ E(S^4) = E(M^4) - 4(n+1)\sigma^2E(M^2) + 4n(n+1)\sigma^4, \end{cases} \tag{6.15}$$

and the LMMSE eq. 6.14 for a nc-χ noise can finally be expressed as:

$$\boxed{\hat{S}^2 = \langle M^2\rangle - 2n\sigma^2 + \left(1 - \frac{4\sigma^2\left[\langle M^2\rangle - n\sigma^2\right]}{\langle M^4\rangle - \langle M^2\rangle^2}\right) \times \left(M^2 - \langle M^2\rangle\right)}. \tag{6.16}$$

As performed in [Aja-Fernández et al. (2008a)], and under the assumption of local ergodicity, we replaced the expectation $E(\cdot)$ by $\langle\cdot\rangle$ corresponding to a local spatial mean calculated on a neighborhood. For a single-channel acquisition (*i.e.* $n = 1$), eq. 6.16 simplifies to its Rician form given by eq. 6.12, as expected.

This demonstration was performed for the SoS reconstruction without subsampling, without correlation between the channels and with each channel having the same variance. However, the demonstration is also valid for the more general SoS and SoS GRAPPA reconstructions with σ and n replaced by $\sigma_{eff}(\mathbf{v})$ and $n_{eff}(\mathbf{v})$. In [Aja-Fernández et al. (2013)], the LMMSE was performed using the expressions of the effective parameters for the simplified scenario described in table 5.1, in the previous chapter.

6.2.3 Singularities in the LMMSE

After having described the LMMSE adapted to Rician and nc-χ noises, we have to mention three particular cases of the LMMSE which require a specific treatment:

1. It can happen that the LMMSE result $\hat{S}^2$ is negative. This is an impossible solution. In a fruitful discussion with Dr. Aja-Fernández and with Dr. Tristán-Vega, these two authors of [Aja-Fernández et al. (2008b)] and [Tristán-Vega and Aja-Fernández (2010)] recommended us to set $\hat{S} = 0$, since this artifact generally occurs in CSF regions and in the background, but very seldom in the grey matter. In dMRI, the main region of interest is the white matter, therefore this case does not have an impact on the post-analysis of the corrected DW data.

2. In LMMSE eq. 6.12, we set $K_R = \left(1 - \frac{4\sigma^2[\langle M^2\rangle - \sigma^2]}{\langle M^4\rangle - \langle M^2\rangle^2}\right)$. And in eq. 6.16, we can set similarly: $K_C = \left(1 - \frac{4\sigma^2[\langle M^2\rangle - n\sigma^2]}{\langle M^4\rangle - \langle M^2\rangle^2}\right)$. These two expressions K_R and K_C exist only if the denominators of the fractions are not equal to zero. These denominators can also be written as $\mathrm{Var}(M^2)$. If the term $\mathrm{Var}(M^2)$, calculated using the neighborhood, is equal to zero, then it means that the voxel studied is located in an homogeneous region. In the appendix A of [Aja-Fernández et al. (2008b)], the authors introduced a simple spatial mean to compute $\hat{S}^2$ in the Rician noise situation. We chose a similar approach in the nc-χ noise situation. The estimation is then done simply using: $\hat{S}^2 = E(M^2) - 2\sigma^2$ for the Rician case and $\hat{S}^2 = E(M^2) - 2n\sigma^2$ for the nc-χ case. This solution is equivalent to setting K_R and K_C equal to zero. If this estimation yields negative squared results, we are in the same case as previously and follow the same treatment.

3. Finally, Aja-Fernández et al. (2008b) explained that $K_R < 0$ corresponds to numerical artifacts. After empirical tests, they decided to set $K_R = \max(K_R, 0)$. We have followed this choice in the nc-χ noise situation, which suited well to our data.

6.2.4 Empirical tuning of global effective parameters in the nc-χ model

Practically, using eq. 6.16 requires a good estimate of the noise standard deviation σ. Ideally, to account for the spatially varying nature of the noise due to the noise correlations, it is recommended in [Aja-Fernández et al. (2011); Aja-Fernández and Tristán-Vega (2012)] to calculate an effective noise standard deviation, as well as an effective number of channels at each voxel. As this is very time-consuming, which goes against our speed constraint, we proposed a **trade-off**: the latter consists of limiting the estimation to a **global effective variance** and a **global effective number of channels** used for all the voxels in the volume. We determined both global parameters σ_{eff} and n_{eff} empirically in the following manner: we tested different values for n_{eff}, each of which impacted the computation of σ_{eff}, which was performed using the "mode M1-χ" method resumed by eq. 5.26 (page 124) with n replaced by n_{eff}. We rewrite the estimation equation in our context here:

$$\underline{\text{ModeM1-}\chi\text{:}}\ \boxed{\hat{\sigma}_{eff} = \left(\sqrt{2}(n_{eff})^{(1/2)}\right)^{-1} \text{mode}\left(\langle M_{bg}(\mathbf{v})\rangle\right)}, \tag{6.17}$$

where $M_{bg}(\mathbf{v})$ is the measured magnitude at the voxel $\mathbf{v}$ in the background region and mode $(\langle M_{bg}(\mathbf{v})\rangle)$ is the distribution mode of the local mean of $M_{bg}(\mathbf{v})$; $(n_{eff})^{(1/2)}$ is calculated applying the definition of the Pochhammer symbol, which can be found in appendix B at the end of the manuscript. Then we chose the values of $(n_{eff}; \sigma_{eff})$ that produced the highest FA, GFA or cGFA ratio — as defined in chapter 5, subsection 5.3.2, depending on the local diffusion model— between a white matter (WM) region and a cerebrospinal fluid (CSF) region. This concerned the correction performed on our real data and some experimental details are given in subsection 6.3.4.

6.3 Results & discussion

6.3.1 Generation of our simulated data

To validate our correction methods, we generated simulated DW data. Here we briefly present these simulated data. More details can be found in the appendix C at the end of the manuscript.

For ground truth, we created 3D volumes (volume size=27×31×27) of noise-free T_2-weighted and DW data (with 500 diffusion orientations) depicting a fiber crossing. The DW volume was artificially corrupted with nc-χ noise to perform validations. The noise-free diffusion data were modeled using a Gaussian mixture to create two fiber bundles crossing at 60°:

$$S(b, \mathbf{o}_i) = S_0 \sum_{k=1}^{2} f_k e^{-b\mathbf{o}_i^T \mathbf{D}_k \mathbf{o}_i}, \tag{6.18}$$

with $\mathbf{o}_i$ corresponding to the orientation at the i^{th} iteration, f_k being the volumic fraction $(\sum_{k=1}^{2} f_k = 1)$ and $\mathbf{D}_k$ being the diffusion tensor, both associated to the k^{th} fiber bundle.

All other notations are the same as before. This generation of data performed with eq. 6.18 assumes that there is no water molecule exchange between the two bundle compartments. This assumption is acceptable, as exchange time between compartments is much longer than the diffusion time used in dMRI experiments. To simulate two bundles crossing at 60°, we set the same $[\lambda_1, \lambda_2, \lambda_3] = [1.7 \times 10^{-9}, 0.3 \times 10^{-9}, 0.3 \times 10^{-9}]$ eigenvalues (in $m^2 \cdot s^{-1}$) [Tuch (2002); Descoteaux (2008)] to both bundles and computed the eigenvectors to obtain the desired angle. The *b-value* was set to $b = 4500s \cdot mm^{-2}$. To add the nc-χ noise with a number of channels $n = 4$, we first calculated the noise-free signals $S(c)$ that each channel would receive, assuming that all channels receive the same signal: $\forall c \in [\![1; n]\!], S(c) = S/\sqrt{n}$, with S the noise-free signal given by eq. 6.18. Thus, the noisy signal is given by the nonlinear transform $M = \left(\sum_{c=1}^{n} \left(S(c) + \epsilon_r(c)\right)^2 + \epsilon_i(c)^2 \right)^{1/2}$, where $\epsilon_r(c)$ and $\epsilon_i(c)$ are the real and imaginary noises, respectively, on the channel c. They are generated using a Gaussian noise distribution of standard deviation $\sigma = 20$ for a T_2-weighted signal $S_0 = 200$. The parameter values for these simulations were chosen close to the real values measured on our real DW data acquired at $b = 4500s \cdot mm^{-2}$. Moreover, we decided to work with $b = 4500s \cdot mm^{-2}$, a typical *b-value* for HARDI which also represents a challenging case for noise correction, as the DW signal is very low at this *b-value*.

6.3.2 Results on simulated data & discussion

We compared the original LMMSE adapted to Rician distribution developed by Aja-Fernández et al. (2008b) with the LMMSE adapted to nc-χ distributions on these simulated DW data at $b = 4500s \cdot mm^{-2}$ corrupted by a nc-χ noise with $\sigma = 20$. For simplicity, we designate these methods as the Rice LMMSE and the nc-χ LMMSE, respectively. No noise correlation was involved in the simulated data, so we used the real number of receiver channels n in the nc-χ LMMSE analyses of those data. We computed both LMMSE using a 5×5×5 neighborhood. We will discuss this neighborhood choice at the end of the subsection. We show the results for the aQBI model, for which the maximum SH order N was set to 8. This high value yields 45 coefficients for the modified SH basis. With this high number, noise has a higher impact on the results. This case, which is a little higher than the common setting (usually $N = 4$ or $N = 6$) permits to assess the performances of the denoising tools at a higher noise level. In the aQBI model, the Laplace-Beltrami regularization λ was first set to 0 and then to 0.006 for the generation of the aQBI maps. This permits a comparison between the correction alone and the correction followed by the regularization. Before the computation of the LMMSE, the noise variance σ was estimated using the "mode M1-χ" method resumed by eq. 5.26 (page 124) with $n = 1$ for the Rice LMMSE and with $n = 4$ for the nc-χ LMMSE.

Fig. 6.3 shows the results on the dODF fields. It allows to compare the noisy dODF and the corrected dODF, with either the Rice or the nc-χ LMMSE, with the noise-free dODF. Moreover, the impact of the regularization can also be analyzed. Because of a normalization of each dODF by its maximal amplitude, the region outside the bundles, where

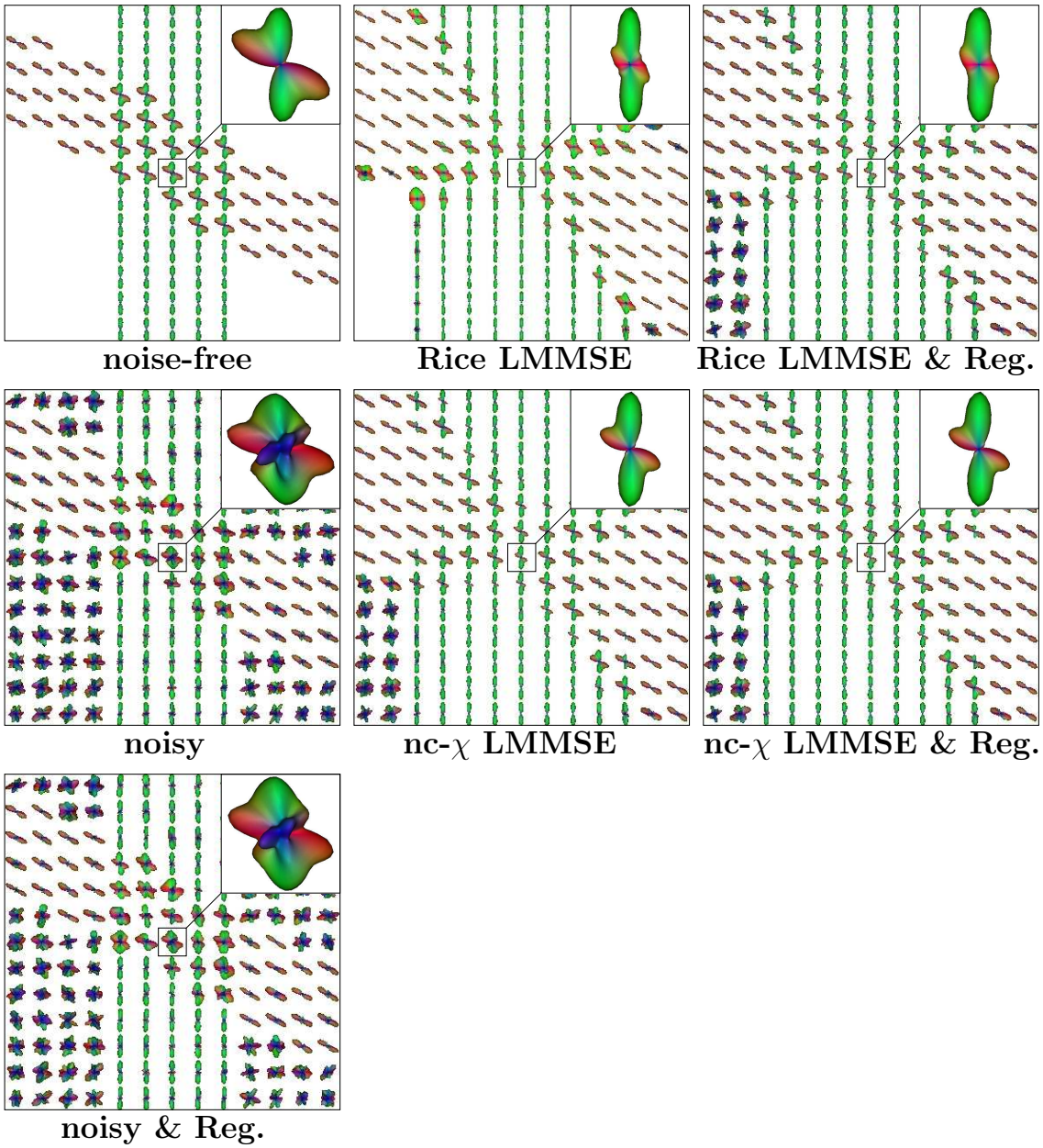


Figure 6.3: Comparison between the Rice LMMSE and the nc-χ LMMSE on simulated DW data corrupted by nc-χ noise with $\sigma = 20$. The dODF fields were generated from the noise-free, the noisy and the results of both LMMSE without and with regularization (when "& Reg." is written). On each field a zoom of a dODF from the crossing region is displayed.

the configuration is defined by isotropic tensors, appears blank in some cases. The results presented in fig. 6.3 show that both correction techniques, without any regularization, improve the noisy dODF field yielding a higher similarity to the noise-free field. However, there is an oversmoothing effect on the dODF field generated by both LMMSE. Indeed, the vertical bundle of dODF after these LMMSE is almost 80% wider than the original one. This effect of the LMMSE is due to the isotropic averages performed using neigh-

borhoods. We will come back to it later. The zoomed dODF in fig. 6.3 reveals that only the nc-χ LMMSE retrieves the noise-free angular information composed of two distinct peaks. Finally, the impact of the regularization is a very small smoothing on each dODF which does not significantly change the results. These qualitative results are confirmed by the graph in fig. 6.4 which shows the MSE calculated as explained in section 5.3 of the previous chapter, such that:

$$\text{MSE} = \frac{1}{N_v} \sum_{\mathbf{v} \in \mathcal{V}} \sum_{j=1}^{N} \left(\tilde{x}(j, \mathbf{v}) - x(j, \mathbf{v}) \right)^2 , \tag{6.19}$$

with N_v the number of voxels in the region shown in fig. 6.3, $\tilde{x}(j, \mathbf{v})$ the j^{th} coefficient of the noisy vector $\tilde{\mathbf{x}}$ calculated on the noisy DW data. $x(j, \mathbf{v})$ is the j^{th} coefficient of the noise-free vector $\mathbf{x}$. The MSE was calculated on all configurations. The MSE calculated on the results obtained with regularization was found practically the same as the MSE of the corresponding result without regularization. For very small values of σ ($\sigma \leq 3.5$), the regularization produces a slightly higher MSE, whereas for values of σ higher than 3.5, it produces a slightly lower MSE than the MSE of the results without regularization. Consequently, for these values of noise levels, the regularization does not damage the LMMSE correction and even improves it. The graph shows that the lowest MSE was achieved with the nc-χ LMMSE (with regularization) for the panel of noise levels.

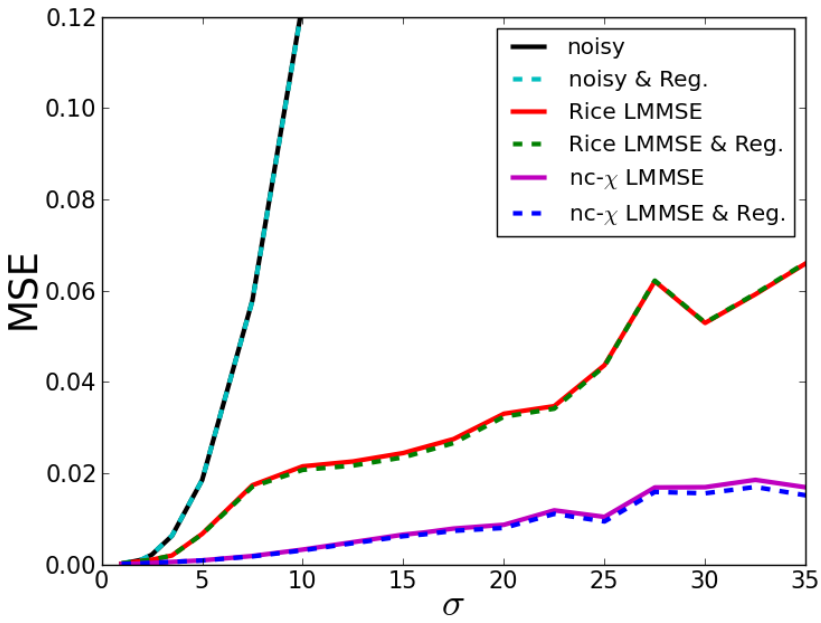


Figure 6.4: Comparison between the MSE of the noisy dODF field without and with regularization and the MSE of the dODF fields after the Rice LMMSE and after the nc-χ LMMSE also without and with regularization. The noisy MSE is of 2.3 at $\sigma = 35$. The black and cyan dotted curves were obtained superimposed, as well as the red and green dotted curves, and finally the magenta and blue dotted curves.

Choice of the neighborhood used in the nc-χ LMMSE.

Fig. 6.3 shows LMMSE results performed with a 3D neighborhood of 5×5×5 voxels. We chose a 3D neighborhood instead of a 2D neighborhood, as the resolution of the generated 3D volume is isotropic (This is also true for our real data). This allowed us to consider neighbors in the three spatial dimensions. Then, it seemed more accurate to use the same number of neighbors for all directions around the central voxel being considered for correction. The last choice to make concerned the number of neighbors. For that, we compared the results obtained with the nc-χ LMMSE with 3×3×3, 5×5×5 and finally 7×7×7 neighborhoods. Fig. 6.5 shows the comparison between the three results. It is obvious that when the neighborhood size increases, the smoothing effect also increases and this makes the MSE greater. However, it also permits to better retrieve the crossing configuration of the zoomed dODF, with a more accurate removal of the noise (see the black arrows). To choose the optimal neighborhood size, we considered a trade-off to obtain an accurate noise removal and a good crossing definition with the most limited smoothing effect. The 5×5×5 size offered us this trade-off. To better control the smoothing effect, a solution will be presented in the next chapter allowing to cancel it. To get an efficient noise correction, it requires enough amount of neighbor information. Typically, the 3×3×3

neighborhood will not be large enough. Therefore we made our choice of the neighborhood size not only looking at the smoothing effect and MSE, but also looking at the crossing precision retrieved by the method.

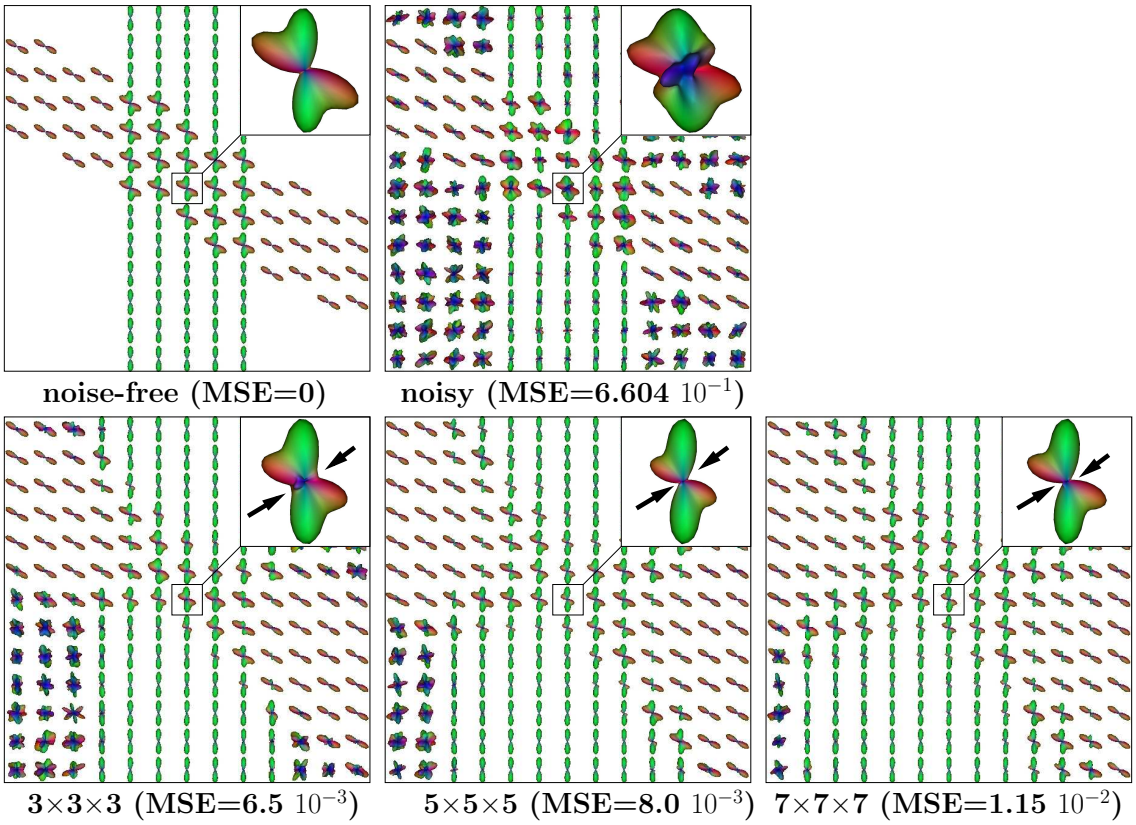


Figure 6.5: Comparison between the 3×3×3, 5×5×5 and 7×7×7 neighborhood sizes used in the nc-χ LMMSE. The nc-χ LMMSE was performed on a noisy DW volume with $\sigma = 20$ (similarly as before). The dODF fields were generated from the noise-free configuration without regularization and from the noisy and the results of the LMMSE with regularization. On each field a zoom of a dODF from the crossing region is displayed. The MSE is also indicated.

The size of the neighborhood also has an impact on the computation time of the method and it has to be accounted when performing RT denoising. We will come back to it later when discussing the nc-χ LMMSE results on real data.

6.3.3 Presentation of our real data

We applied the correction methods on real data, which acquisition protocol is detailed in the appendix C at the end of the manuscript. Here, we briefly present the DW data acquired at four *b-values*.

These data were collected on a human brain using a Magnetom Tim Trio 3T MRI system (Siemens Medical Solutions, Erlangen, Germany), employing a spherical direction sampling of 60 orientations uniformly distributed over each shell at $b = 1500/3000/4500/6000s\cdot$

mm^{-2}. Three T_2-weighted volumes were acquired at $b = 0s \cdot mm^{-2}$. The acquisition parameters were as follows: $T_E/T_R = 116ms/14s$, field of view FOV= $220mm$, matrix 128×128, 70 slices, resolution $1.7{\times}1.7{\times}1.7mm^3$, GRAPPA factor of 2, read bandwidth RBW=$1628Hz/pixel$. Fig. 6.6 shows three DW images obtained for three different diffusion gradients. The acquisition was performed with a 12-element head coil available on the Tim Trio, for which the 12 coil elements are combined into 4 groups of 3 coil elements. These groups are received through 4 distinct receiver channels, yielding $n = 4$.

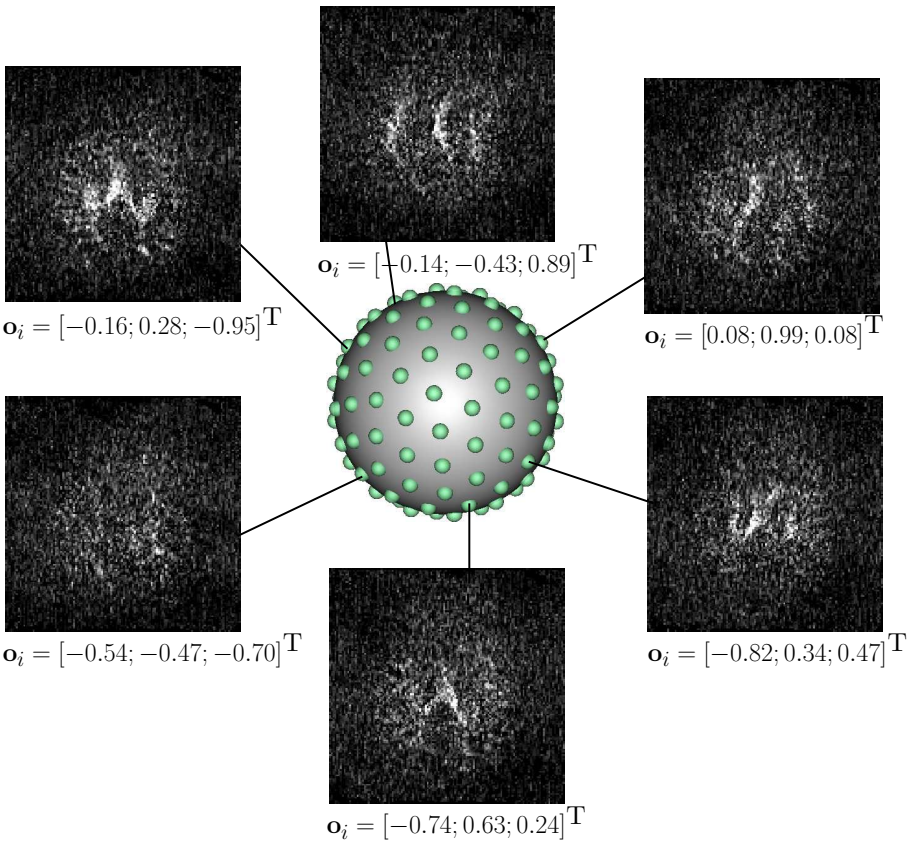


Figure 6.6: DW data for three diffusion gradients. In the center, the sphere of the **q** space is represented, with green dots corresponding to the projections of the orientations chosen for the measurement of the diffusion. Three DW images are shown for three diffusion orientations $\mathbf{o}_i$.

6.3.4 Results on real data & discussion

We performed a similar comparison on the real DW data acquired at $b = 4500s \cdot mm^{-2}$. Again, we computed the results for the aQBI model, for which the maximum SH order N was set to 8 and the Laplace-Beltrami regularization λ was first set to 0 and then to 0.006 for the generation of the aQBI maps. This permits a comparison between the correction alone and the correction followed by the regularization. In the case of the Rice LMMSE, the noise variance σ was estimated using the "mode M1-χ" method resumed by eq. 5.26 (page 124) with $n = 1$ for the Rice LMMSE Concerning the nc-χ LMMSE, we accounted

for the possible noise correlations in the real data in two different ways. The first way calculates global effective parameters for the whole DW volume. Therefore, this solution is not rigorous regarding the analysis in [Aja-Fernández et al. (2011); Aja-Fernández and Tristán-Vega (2012)], as the effective parameters should be voxel-wise. However, this first solution allows to already account for eventual noise correlations and is above all far more computationnaly efficient than a voxel-wise technique. This trade-off is explained in subsection 6.2.4. Fig. 6.7 reveals how in practice we obtained the optimum global effective parameters for the nc-χ LMMSE applied with regularization at $b = 4500s{\cdot}mm^{-2}$, following the steps detailed in subsection 6.2.4. The second way to handle these correlations is the more rigorous calculation of the voxel-wise parameters as described in table 5.1, page 126 [Aja-Fernández et al. (2013)]. This technique can be applied in RT, when performed separately on each DW volume acquired for a given diffusion orientation: in the RT process, the method is executed on the new DW volume acquired before the LMMSE algorithm. This solution was developed for a simplified scenario in case of an SoS reconstruction without any partially parallel MRI reconstruction. Therefore, this solution is not entirely adapted to our real data acquired with GRAPPA. Nevertheless, to date, it is the only solution proposed to obtain both n_{eff} and σ_{eff} separately.

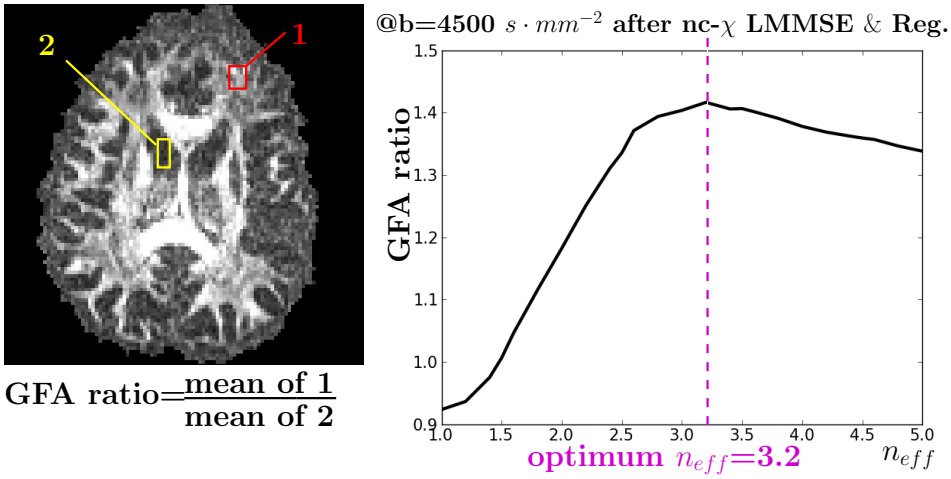


Figure 6.7: The method to determine the global σ_{eff} and the global n_{eff} for the nc-χ LMMSE applied with regularization at $b = 4500s \cdot mm^{-2}$. Two ROI are chosen on the GFA map (the map shown here (left) is the raw GFA at $b = 1500s \cdot mm^{-2}$). The GFA ratio is computed as indicated for a panel of values of n_{eff}. The optimum n_{eff} is chosen when the GFA ratio is at its maximum.

Fig. 6.8 depicts the results obtained with the Rice and the nc-χ LMMSE performed on the DW data at $b = 4500s \cdot mm^{-2}$. Both methods, the global one and the voxel-wise one, were used to estimate the effective parameters. These results on real data show that the oversmoothing effect is high with the Rice LMMSE, whereas it is far more acceptable with the nc-χ LMMSE. The comparison between the global and the voxel-wise estimations of the effective parameters show a slightly higher smoothing on the DW image obtained using the global parameters. In both cases, the smoothing happens in approximately

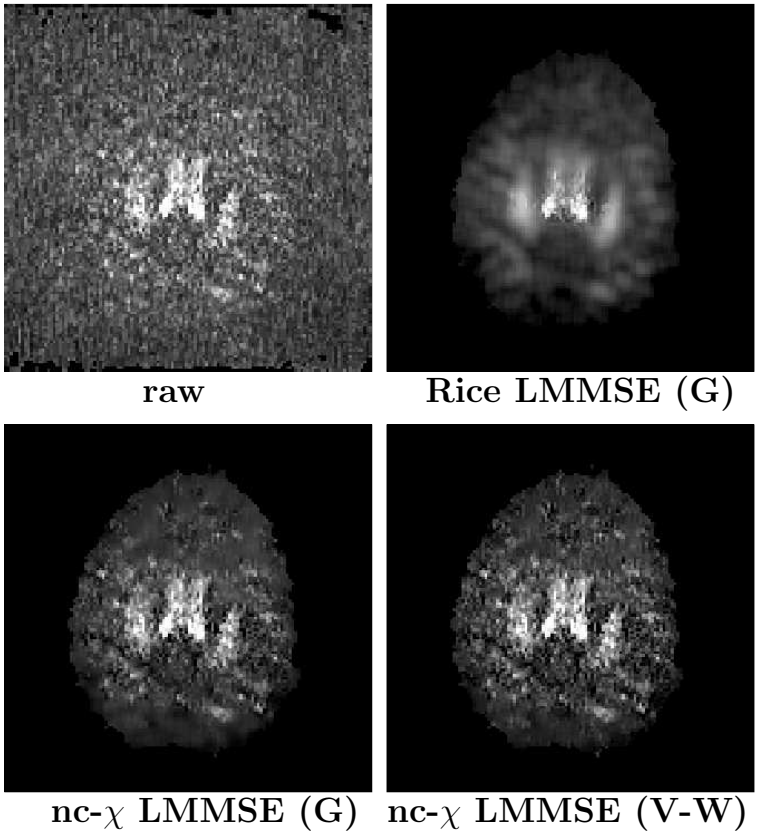


Figure 6.8: Comparison between the Rice LMMSE and the nc-χ LMMSE on real DW data at $b = 4500s \cdot mm^{-2}$ with either the global (G) or the voxel-wise (V-W) effective parameters.

homogeneous regions and is stopped in the borders to preserve the details. Then, we compared the GFA maps obtained either without or with regularization. We show in fig. 6.9 the comparison on the raw maps and the maps corrected using the nc-χ LMMSE with the voxel-wise effective parameters. The message is clear: the regularization improves a lot the visual quality of the maps. Similar results were obtained with the other techniques.

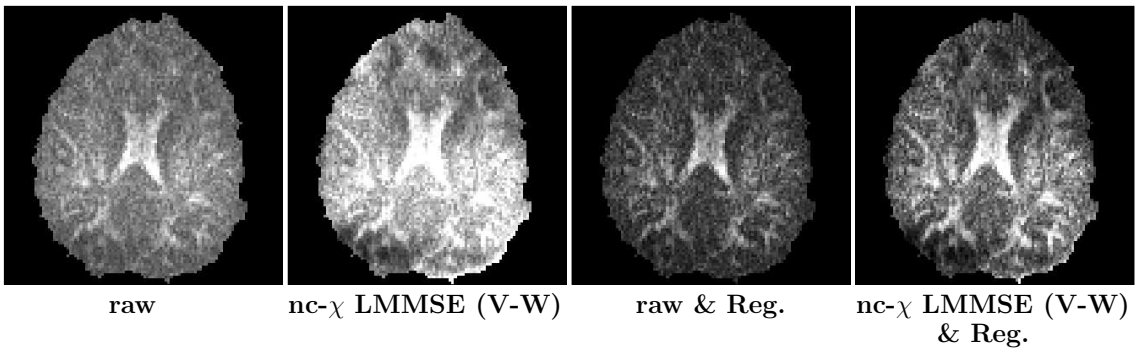


Figure 6.9: Comparison between the nc-χ LMMSE on real GFA maps at $b = 4500s \cdot mm^{-2}$ with either regularization ("& Reg.") or not.

Therefore, from now on, we only show the results obtained with regularization.

Finally, fig. 6.10 regroups the resulting GFA maps generated with the Rice and the nc-χ LMMSE with regularization. It also compares the use of the global and the voxel-wise estimation of the effective parameters in the case of the nc-χ LMMSE. It is again visible that the nc-χ LMMSE yields better corrected GFA maps than the Rice LMMSE, which oversmoothes the maps. Concerning the comparison between the global and the voxel-wise methods, it seems that there is no clear winner. Indeed, at $b = 1500s \cdot mm^{-2}$, the voxel-wise technique produces a smoothed GFA, whereas the global technique is much finer. For the other *b-values*, it is the opposite: the global technique produce GFA with a little more

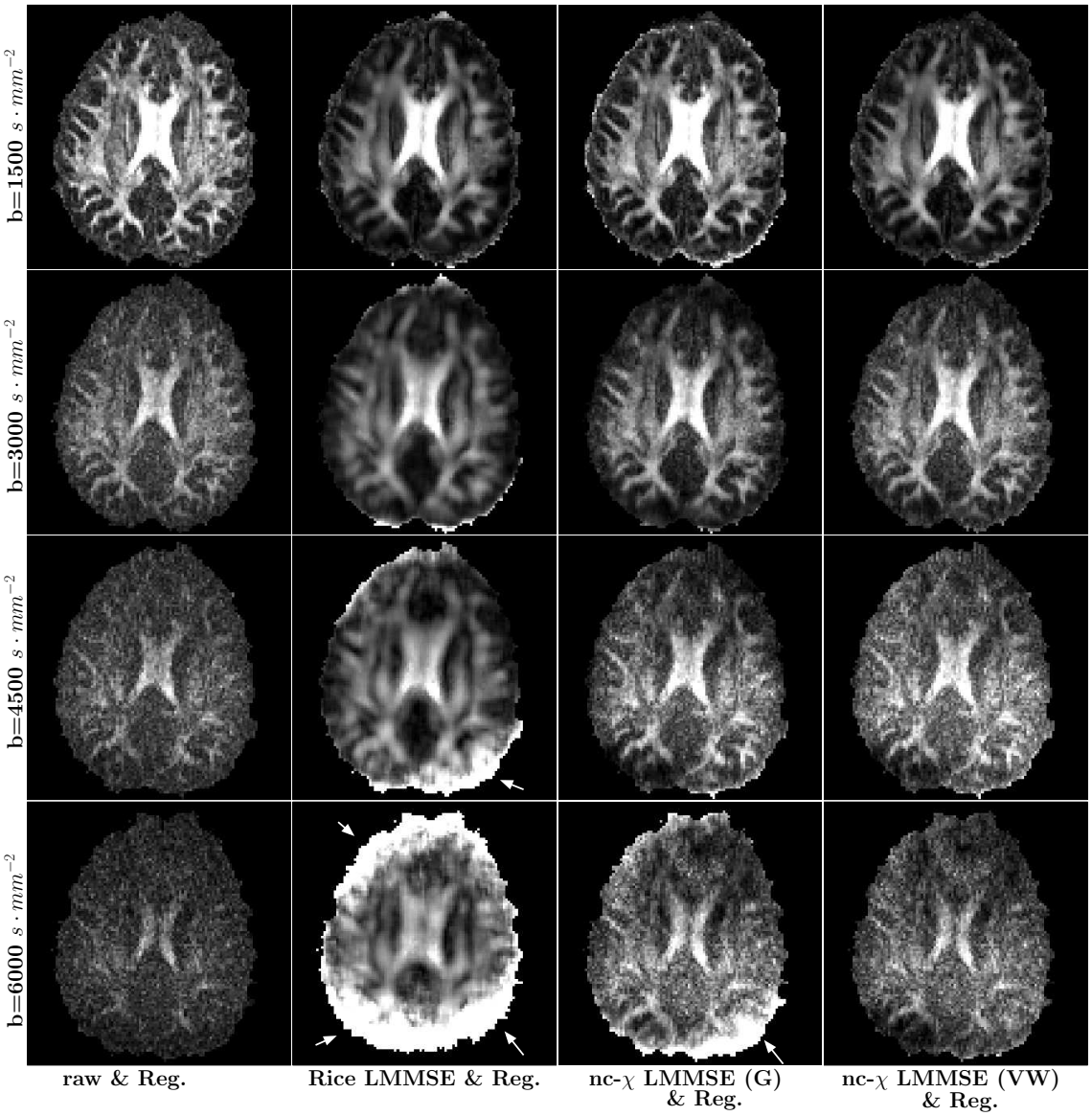


Figure 6.10: Comparison between the Rice and the nc-χ LMMSE on real GFA maps with regularization ("& Reg.") and with either the global (G) or the voxel-wise (V-W) effective parameters at $b = 1500/3000/4500/6000s \cdot mm^{-2}$. Hyperintensities in the maps are shown by white arrows.

smoothing, compared to the voxel-wise version. This is due to the fact that the global n_{eff} for these three *b-values* are lower than most of the voxel-wise n_{eff}, as we will analyze it later. We also notice in fig. 6.10 that there are hyperintensities located near the border of the mask in the case of the global technique (which are also present in worth with the Rice LMMSE). These hyperintensities are shown by white arrows. They are probably due to an intensity artifact in this area in the raw DW data. In the process of the LMMSE, the voxels located in this region may fall in the two first particular cases, described in subsection 6.2.3, for which the final LMMSE result is set to zero. Consequently, if the DW signals in a region are set to zero for most of the orientations and different from zero for very few of them, then the result is a very high anisotropy in that region, leading to white hyperintensities on the GFA maps. However, with the use of the voxel-wise method, these hypeintensities disappear thanks to the locally adapted noise variance estimation.

To quantitatively analyze the results presented in fig. 6.10, we calculated the GFA ratios on each configuration (fig. 6.11). The GFA ratio, as explained in the previous chapter, in subsection 5.3.2, is defined such that:

$$\text{GFA ratio} = \frac{\text{GFA}_\text{A}}{\text{GFA}_\text{B}}, \tag{6.20}$$

with GFA_A and GFA_B the GFA means calculated on the GFA map for the regions A and B, respectively. The ROI were chosen as shown in fig. 6.7. To obtain a GFA ratio indicator, which rightly measures the contrast change due to the method applied, it was important to choose for the region A an area of medium signal and not high signal. That is why we chose a region A located in a subcortical white matter (WM) area. For region B, we chose a low signal region in a CSF area located close to a WM region, so that a smoothing effect of the correction technique can be detected: if a smoothing effect appears (as in our previous presented results on simulated data), this region B will contain traces of WM signals, making the GFA ratio decrease. We can see in fig. 6.11 that the GFA ratio has indeed a tendency to decrease with the Rice LMMSE, which oversmoothes the map. Both global and voxel-wise nc-χ LMMSEs produce the highest GFA ratios, with a better result for the global technique. We have to keep in mind that this quality indicator is however local and has to be analyzed not alone, but together with the visual inspection of the quality of the images.

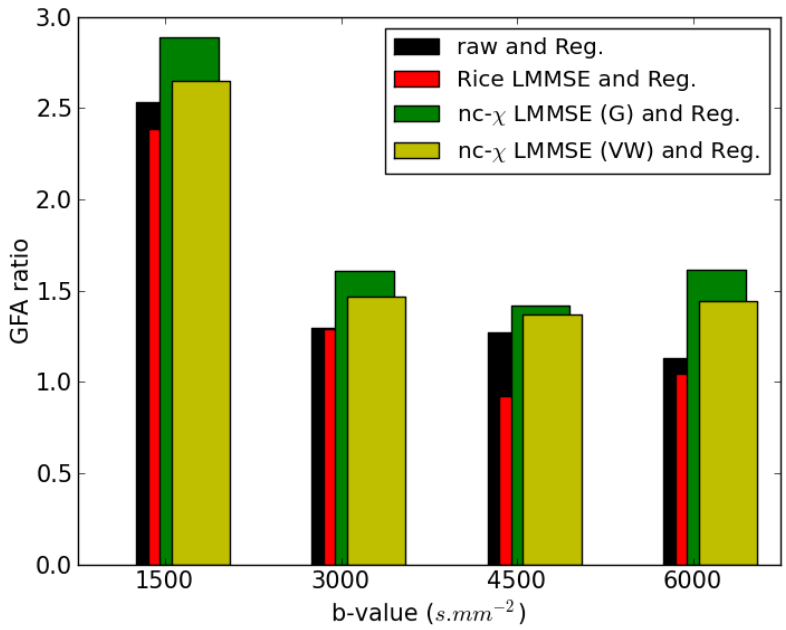


Figure 6.11: Comparison between the GFA ratios of the raw map and the map obtained after application of the Rice and the nc-χ LMMSE with regularization and with either the global or the voxel-wise effective parameters at $b = 1500/3000/4500/6000s \cdot mm^{-2}$. The ROI were chosen as shown in fig. 6.7.

To better understand the comparison between the global and the voxel-wise estimations of the effective parameters, we draw the histograms of the voxel-wise σ_{eff} and n_{eff} on the entire DW volumes. These histograms are shown in fig. 6.12. From these histograms, we can see that: first, the histogram of σ_{eff} is less spread when the *b-value* increases (for $b > 3000s \cdot mm^{-2}$). This means that for high *b-values* ($b > 4500s \cdot mm^{-2}$), the voxel-wise estimation of σ_{eff} tends to a global estimation with a nearly unique σ_{eff} for the whole volume. Secondly, concerning n_{eff}, it is almost one in the whole volume at $b = 1500s \cdot mm^{-2}$ and increases with b, yielding almost four (*i.e.* the number of channels) at $b = 6000s \cdot mm^{-2}$ for the whole DW volume. These histograms can be linked with the visual results in fig. 6.10 and to the values of σ_{eff} and n_{eff} obtained with the global method and given in table 6.1. Indeed, the voxel-wise estimation at $b = 1500s \cdot mm^{-2}$ generates too much smoothing because the values of n_{eff} are close to one. For the other *b-values*, the smoothing is subtler, yielding a small denoising, this corresponding to values of n_{eff} closer to four, as indicated by the histograms of n_{eff}. In comparison, the n_{eff} values obtained with the global solution (in table 6.1) are lower for $b > 3000s \cdot mm^{-2}$, yielding a stronger denoising with some smoothing side effect.

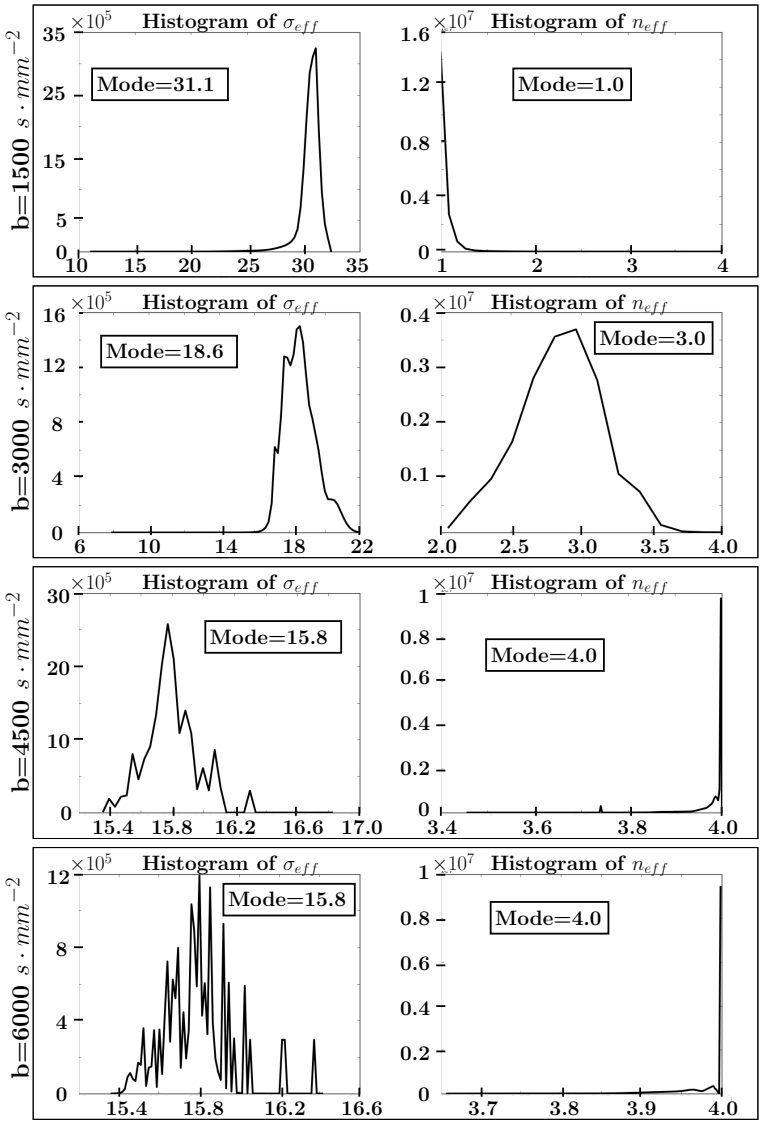


Figure 6.12: Histograms, performed inside the brain, of the voxel-wise parameters σ_{eff} and n_{eff} obtained with the voxel-wise technique by [Aja-Fernández et al. (2013)].

Global effective n_{eff} and σ_{eff} obtained for the nc-χ LMMSE				
b $(s \cdot mm^{-2})$	1500	3000	4500	6000
n_{eff}	2.0	2.2	3.2	3.2
σ_{eff}	24.1	22.3	18.2	18.0

Table 6.1: The global effective n_{eff} and σ_{eff} obtained for the nc-χ LMMSE at $b = 1500s \cdot mm^{-2}$ / $b = 3000s \cdot mm^{-2}$ / $b = 4500s \cdot mm^{-2}$ / $b = 6000s \cdot mm^{-2}$. Regularization was used to generate the GFA maps on which the GFA ratios necessary to obtain these global effective parameters were calculated.

Computational efficiency.

Finally, we also measured the computation time for the noise correction methods. For the global solution, we obtained a time of 2.2s on one CPU @ 2.8 GHZ, for a real DW volume of 128×128×70 at each orientation acquired. This time increased to 14.2s with the voxel-wise method, almost corresponding to the repetition time of the dMRI acquisition. These LMMSE computation times depend on the choice of the neighborhood used to calculate the LMMSE averages. In subsection 6.3.2 we explained that we chose a 5×5×5 neighborhood, as it was a good trade-off between a precise noise removal and a limited smoothing effect. Fig. 6.13 shows that the choice of this neighborhood size keeps the computation times lower than with higher neighborhood sizes.

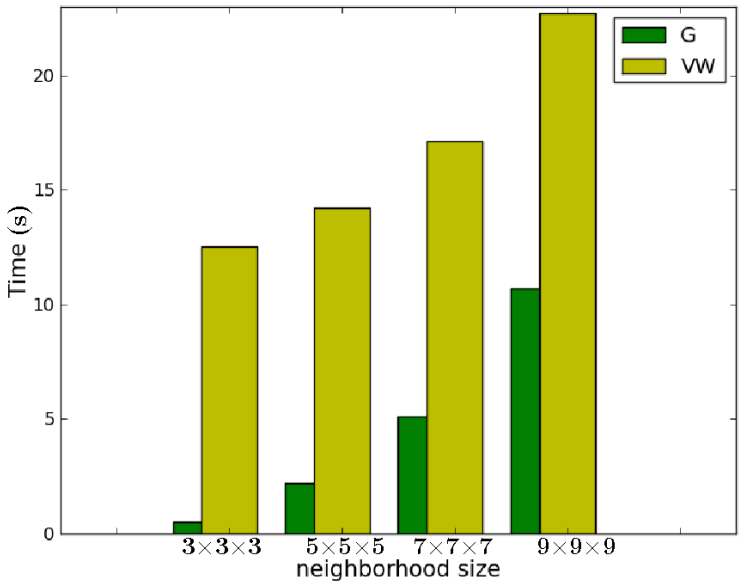


Figure 6.13: Impact of the neighborhood size on the LMMSE computation time. Both global (G) and voxel-wise (VW) versions of the LMMSE were tested.

It is important to notice here that the denoising of the global method, obtained without any parallelization on a cluster yet, is very low compared to other methods' times, *e.g.* the NLM filter time mentioned in 6.1.2 for an approximatively 6 times bigger data volume: for a 181×217×181 volume on a 3 GHz CPU, the NLM lasted several hours in its original version. This time was reduced, in parallelized versions with 8 CPU @ 3 GHz each, to around one minute with the accelerated method by [Coupé et al. (2008b)] and at most 35s with the accelerated method by [Tristán-Vega et al. (2012)].

Although the times of the global and voxel-wise methods are already very low, they can be further reduced. Indeed, for the RT process to be truly RT, it is essential that the denoising method delivers the corrected DW volume before the acquisition of the next DW volume. Therefore, the denoising process should last less than the repetition time T_R of typically a dozen of seconds. To achieve this time reduction, we used a cluster of 80 CPUs. The parallelization and distribution of the C++ code on this cluster permitted to

obtain times of 48ms and 307ms for the global and the voxel-wise nc-χ LMMSE, clearly far below the T_R.

6.3.5 Conclusion on the extended LMMSE adapted to nc-χ distributions

The extended LMMSE adapted to nc-χ distributions produced nice results on our simulated and real data. However, it caused a smoothing effect, clearly visible on our simulated data that is due to averages performed isotropically and that requires to be attenuated. This smoothing effect will be addressed in the next chapter.

The global solution presents the disadvantage of finding the optimum σ_{eff} and n_{eff} in an empirical way by testing the nc-χ LMMSE for several values of n_{eff}. This cannot be performed in RT. Consequently, our global method is proposed for a study of a database acquired with a unique set of constant parameters: the optimum n_{eff} can be tuned on one DW volume set, and its value can be used on the rest of the database, without any supplementary tuning. This disadvantage does not appear with the voxel-wise solution, however this latter technique was shown to yield sometimes less adequate effective parameters values than the global method.

Finally, the small time required by our method, as well as its good denoising results confirmed our choice to use LMMSE and LMMSE-derived algorithms for our RT noise correction objective.

6.4 Conclusion of this chapter

In this chapter, after providing an overview of denoising methods actually available to correct dMRI, we focused on LMMSE methods. We proposed an original extended LMMSE algorithm compatible in terms of computational efficiency with the objective of this thesis to run it in RT during the ongoing scan. To account for eventual noise correlations, we proposed an alternative to the voxel-wise determination of effective parameters of [Aja-Fernández et al. (2013)]. Indeed, we suggested to use a trade-off consisting of an empirical estimation of the effective parameters defined globally for the data volume. We analyzed the advantages and disadvantages proposed by such a method. We also qualitatively and quantitatively studied the performances of the LMMSE on simulated and real DW data and implemented a parallel version of it on a cluster of CPUs. The contributions of our work are listed below.

6.4.1 Contributions of this chapter

- A detailed state-of-the-art review on the noise correction methods.
- An extension of the LMMSE to nc-χ distributions. This nc-χ LMMSE filter was shown to outperform the Rice LMMSE on simulated and real data corrupted by nc-χ noise [Brion et al. (2011b,c,a)].

- A trade-off with the empirical estimation of effective parameters to account for eventual noise correlations in the data to correct. This solution was shown to be faster than its voxel-wise alternative proposed by [Aja-Fernández et al. (2013)] and to deliver accurate results.
- An analysis of the performances of the LMMSE (time and results) regarding our RT denoising purpose.

In this chapter, we highlighted the constraints to face when wanting to operate a noise correction in RT. The LMMSE respects these constraints and can be performed in RT. The next chapter presents the incorporation of this estimator in the RT framework based on a Kalman filter described in chapter 4. The next chapter also details other methods for an RT noise correction.

6.5 Appendix A: Calculation of the 2^{nd} and 4^{th} order moments of the measured magnitude for a single-channel acquisition

In this chapter, page 148, we used the expressions of the 2^{nd} and 4^{th} order moments of the measured magnitude for a single-channel acquisition with the assumption that noise is Rician. It is possible to obtain these expressions from the equation giving the measured magnitude $M = \sqrt{(S_r + \epsilon_r)^2 + (S_i + \epsilon_i)^2}$ (with $S = \sqrt{S_r^2 + S_i^2}$) and from the hypotheses detailed page 147 concerning the noise signals. Let us begin with the 2^{nd} order moment of M:

$$\begin{aligned} E(M^2) &= E\left((S_r + \epsilon_r)^2 + (S_i + \epsilon_i)^2\right), \\ &= E(S_r^2) + 2\underbrace{E(S_r\epsilon_r)}_{=0} + E(\epsilon_r^2) + E(S_i^2) + 2\underbrace{E(S_i\epsilon_i)}_{=0} + E(\epsilon_i^2), \\ &= E(S_r^2 + S_i^2) + 2\sigma^2. \end{aligned}$$

The latter equation leads to:

$$\boxed{E(M^2) = E(S^2) + 2\sigma^2} \tag{6.21}$$

Similarly, we calculate $E(M^4)$:

$$\begin{aligned} E(M^4) &= E\left([(S_r + \epsilon_r)^2 + (S_i + \epsilon_i)^2]^2\right), \\ &= E\left((S_r + \epsilon_r)^4 + 2(S_r + \epsilon_r)^2(S_i + \epsilon_i)^2 + (S_i + \epsilon_i)^4\right). \end{aligned}$$

For two reals a and b, we have: $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$. This enables to simplify the expression of $E(M^4)$ into:

$$\begin{aligned} E\left((S_r + \epsilon_r)^4\right) &= E\left(S_r^4 + 4S_r^3\epsilon_r + 6S_r^2\epsilon_r^2 + 4S_r\epsilon_r^3 + \epsilon_r^4\right) \\ &= E(S_r^4) + 6E(S_r^2)\sigma^2 + 3\sigma^4. \end{aligned}$$

Similarly, we obtain: $E\left((S_i+B_i)^4\right) = E(S_i^4) + 6E(S_i^2)\sigma^2 + 3\sigma^4$. It remains to calculate $E\left((S_r+\epsilon_r)^2(S_i+\epsilon_i)^2\right)$.

$$\begin{aligned}
&E\left((S_r+\epsilon_r)^2(S_i+\epsilon_i)^2\right)\\
&= E\left(S_r^2S_i^2 + 2S_i\epsilon_iS_r^2 + S_r^2\epsilon_i^2 + 2S_rS_i^2\epsilon_r + 4S_rS_i\epsilon_r\epsilon_i + 2S_r\epsilon_r\epsilon_i^2 + \epsilon_r^2S_i^2 + 2\epsilon_r^2S_i\epsilon_i + \epsilon_r^2\epsilon_i^2\right)\\
&= E(S_r^2S_i^2) + \sigma^2E(S_r^2) + \sigma^2E(S_i^2) + \sigma^4.
\end{aligned}$$

Gathering the three simplified terms together, we obtain:

$$\begin{aligned}
&E(M^4)\\
&= E(S_r^4) + 6E(S_r^2)\sigma^2 + E(S_i^4) + 6E(S_i^2)\sigma^2 + 6\sigma^4 +\\
&2\left(E(S_r^2S_i^2) + \sigma^2E(S_r^2) + \sigma^2E(S_i^2) + \sigma^4\right)\\
&= E(S_r^4 + 2S_r^2S_i^2 + S_i^4) + 8\sigma^2E(S_r^2 + S_i^2) + 8\sigma^4.
\end{aligned}$$

It results that:

$$\boxed{E(M^4) = E(S^4) + 8\sigma^2E(S^2) + 8\sigma^4}. \tag{6.22}$$

6.6 Appendix B: LMMSE calculation for a multiple-channel acquisition with an SoS or an SoS GRAPPA reconstruction

To calculate the LMMSE for nc-χ distributions for a multiple-channel acquisition (with an SoS or an SoS GRAPPA reconstruction), it is first necessary to rewrite the expression of the magnitude with the SoS reconstruction:

$$M = \sqrt{\sum_{c=1}^{n}(S_{r_c}+\epsilon_{r_c})^2 + (S_{i_c}+\epsilon_{i_c})^2}. \tag{6.23}$$

Then, we rewrite the expression of the noise-free magnitude:

$$\begin{aligned}
S &= \sqrt{\sum_{c=1}^{n} S_{r_c}^2 + S_{i_c}^2},\\
&= \sqrt{\sum_{c=1}^{n} S_c^2},
\end{aligned} \tag{6.24}$$

with S_{r_c} and S_{i_c} the real and imaginary components, respectively, of the complex noise-free signal on the channel c. ϵ_{r_c} and ϵ_{i_c} are the real and imaginary components, respectively, of the complex noise signal on channel c. The noises ϵ_{r_c} and ϵ_{i_c} are supposed to be uncorrelated, zero-mean Gaussian noises of standard deviation σ. Therefore, let us remind that we can write:

$$\begin{cases}
E(\epsilon^k) = 0, \text{if k is odd},\\
E(\epsilon^2) = \sigma^2,\\
E(\epsilon^4) = 3\sigma^4.
\end{cases} \tag{6.25}$$

The imaginary and real parts of the noise signal are supposed to be independent. The same hypothesis is applied between the noise ϵ and the noise-free signal S, thus:

$$\begin{cases} E(\epsilon_r \epsilon_i) = E(\epsilon_r)E(\epsilon_i) = 0, \\ E(\epsilon S) = E(\epsilon)E(S) = 0. \end{cases} \tag{6.26}$$

After these statements, let us now rewrite the estimation of the squared noise-free magnitude signal, noted as $\hat{S^2}$, using the LMMSE:

$$\hat{S^2} = E(S^2) + \mathrm{Cov}(S^2, M^2)\mathrm{Var}(M^2)^{-1} \times \left(M^2 - E(M^2)\right). \tag{6.27}$$

This latter equation requires to explicit the unknown terms with respect to known or measurable quantities. The expression of the variance term is easy to get, as in the case of a single-channel acquisition:

$$\begin{aligned} \mathrm{Var}(M^2) &= E\left([M^2 - E(M^2)]^2\right) \\ &= E\left(M^4 - 2M^2E(M^2) + E(M^2)^2\right). \end{aligned}$$

Therefore, we obtain:

$$\boxed{\mathrm{Var}(M^2) = E(M^4) - E(M^2)^2}. \tag{6.28}$$

Let us now calculate the covariance term: $\mathrm{Cov}(S^2, M^2)$:

$$\begin{aligned} \mathrm{Cov}(S^2, M^2) &= E\left([S^2 - E(S^2)][M^2 - E(M^2)]\right) \\ &= E\left(S^2M^2 - S^2E[M^2] - E[S^2]M^2 + E[S^2]E[M^2]\right) \\ &= E\left(S^2\left[\sum_{k=1}^{n}\left((S_{r_k} + B_{r_k})^2 + (S_{i_k} + B_{i_k})^2\right)\right]\right) - E(S^2)E(M^2) \\ &= E\left(S^2\left[\sum_{k=1}^{n} S_k^2\right]\right) + \underbrace{E\left(S^2\left[\sum_{k=1}^{n} 2(S_{r_k}B_{r_k} + S_{i_k}B_{i_k})\right]\right)}_{=0} + 2n\sigma^2E[S^2] - \\ &\quad E(S^2)E(M^2). \end{aligned}$$

Therefore, we obtain:

$$\boxed{\mathrm{Cov}(S^2, M^2) = E\left(S^4\right) + 2n\sigma^2E(S^2) - E(S^2)E(M^2)}. \tag{6.29}$$

Let us now incorporate eq. 6.28 and eq. 6.29 in the LMMSE model eq. 6.27:

$$\hat{S^2} = E(S^2) + \left(E\left(S^4\right) + 2n\sigma^2E[S^2] - E(S^2)E(M^2)\right)\left(E(M^4) - E(M^2)^2\right)^{-1} \times \left(M^2 - E(M^2)\right).$$

Then, we have to simplify the expressions of the 2^{nd} and 4^{th} order moments of S. These moments are demonstrated in the next appendix 6.7, to be equal to:

$$E(S^2) = E(M^2) - 2n\sigma^2, \tag{6.30}$$

$$E(S^4) = E(M^4) - 4(n+1)\sigma^2 E(M^2) + 4n(n+1)\sigma^4. \tag{6.31}$$

Using eq. 6.30 and eq. 6.31 in the previous LMMSE equation, we obtain the final result:

$$\boxed{\hat{S^2} = \langle M^2 \rangle - 2n\sigma^2 + \left(1 - \frac{4\sigma^2\left[\langle M^2\rangle - n\sigma^2\right]}{\langle M^4\rangle - \langle M^2\rangle^2}\right) \times \left(M^2 - \langle M^2\rangle\right)}. \tag{6.32}$$

We easily remark that when setting $n = 1$, eq. 6.32 simplifies to its Rician form given by eq. 6.12, page 148.

6.7 Appendix C: Calculation of the 2^{nd} and 4^{th} order moments of the measured magnitude for a multiple-channel acquisition

In the previous appendix 6.6, we used the expressions of the 2^{nd} and 4^{th} order moments of the measured magnitude for a multiple-channel acquisition with n channels. Here we give the demonstration of eq. 6.30 and eq. 6.31 using only the following relations: $M = \sqrt{\sum_{k=1}^{n}(S_{r_k} + B_{r_k})^2 + (S_{i_k} + B_{i_k})^2}$,and $S = \sqrt{\sum_{k=1}^{n} S_{r_k}^2 + S_{i_k}^2} = \sqrt{\sum_{k=1}^{n} S_k^2}$, as well as the assumptions regarding the noise signals detailed page 167. We begin with the 2^{nd} order moment of M:

$$\begin{aligned}
E(M^2) &= E\left(\sum_{k=1}^{n}(S_{r_k} + B_{r_k})^2 + (S_{i_k} + B_{i_k})^2\right) \\
&= E\left(\sum_{k=1}^{n} S_k^2\right) + \underbrace{E\left(\sum_{k=1}^{n} 2S_{r_k}B_{r_k} + 2S_{i_k}B_{i_k}\right)}_{=0} + \underbrace{E\left(\sum_{k=1}^{n} B_{r_k}^2\right)}_{=n\sigma^2} + \underbrace{E\left(\sum_{k=1}^{n} B_{i_k}^2\right)}_{=n\sigma^2} \\
&= E\left(\sum_{k=1}^{n} S_k^2\right) + 2n\sigma^2.
\end{aligned}$$

Consequently, we obtain:

$$\boxed{E(M^2) = E(S^2) + 2n\sigma^2}. \tag{6.33}$$

Then to compute $E(M^4)$, we first detail M^4:

$$
\begin{aligned}
M^4 &= \left(\sum_{k=1}^{n}(S_{r_k}+B_{r_k})^2+(S_{i_k}+B_{i_k})^2\right)^2 \\
&= \left(\sum_{k=1}^{n}S_k^2+2(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})+B_{r_k}^2+B_{i_k}^2\right)^2 \\
&= \left(S^2+2\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})+\sum_{k=1}^{n}B_{r_k}^2+\sum_{k=1}^{n}B_{i_k}^2\right)^2 \\
&= \underbrace{S^4+4S^2\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})}_{=A_1}+\underbrace{2S^2\left(\sum_{k=1}^{n}B_{r_k}^2+\sum_{k=1}^{n}B_{i_k}^2\right)}_{=A_2}+ \\
&\underbrace{4\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})\cdot\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})}_{=A_3}+\underbrace{4\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})\sum_{k=1}^{n}B_{r_k}^2}_{=A_4}+ \\
&\underbrace{4\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})\sum_{k=1}^{n}B_{i_k}^2}_{=A_5}+\underbrace{\left(\sum_{k=1}^{n}B_{r_k}^2\right)^2+2\sum_{k=1}^{n}B_{r_k}^2\cdot\sum_{k=1}^{n}B_{i_k}^2+\left(\sum_{k=1}^{n}B_{i_k}^2\right)^2}_{=A_6}.
\end{aligned}
$$

Let us calculate the expectation of each A_i term:

$$E(A_1) = E(S^4)+0,$$

$$E(A_2) = 2E(S^2)\times 2n\sigma^2 = 4n\sigma^2E(S^2),$$

$$
\begin{aligned}
E(A_3) &= 4E\left(\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})^2\right)+ \\
&\quad\underbrace{4E\left(\sum_{k=1}^{n}(S_{r_k}B_{r_k}+S_{i_k}B_{i_k})\cdot\sum_{\substack{j=1\\j\neq k}}^{n}\left(S_{r_j}B_{r_j}+S_{i_j}B_{i_j}\right)\right)}_{=0} \\
&= 4E\left(\sum_{k=1}^{n}S_{r_k}^2B_{r_k}^2+2S_{r_k}B_{r_k}S_{i_k}B_{i_k}+S_{i_k}^2B_{i_k}^2\right) \\
&= 4\sigma^2E\left(\sum_{k=1}^{n}S_{r_k}^2+S_{i_k}^2\right) \\
&= 4\sigma^2E(S^2),
\end{aligned}
$$

$$E(A_4) = E(A_5) = 0,$$

$$E(A_6) = E\left(\sum_{k=1}^{n}\left(B_{r_k}^4 + \sum_{\substack{j=1\\ j\neq k}}^{n} B_{r_k}^2 B_{r_j}^2\right)\right) + E\left(\sum_{k=1}^{n}\left(B_{i_k}^4 + \sum_{\substack{j=1\\ j\neq k}}^{n} B_{i_k}^2 B_{i_j}^2\right)\right) +$$
$$2E\left(\sum_{k=1}^{n} B_{r_k}^2 \cdot \sum_{k=1}^{n} B_{i_k}^2\right)$$
$$= 2 \times \left(3n\sigma^4 + n(n-1)\sigma^4\right) + 2n^2\sigma^4.$$

We add the six A_i terms together:

$$E(M^4) = E(S^4) + 4n\sigma^2 E(S^2) + 4\sigma^2 E(S^2) + 2\left(3n\sigma^4 + n(n-1)\sigma^4\right) + 2n^2\sigma^4.$$

After some simplification, the result is given by:

$$\boxed{E(M^4) = E(S^4) + 4(n+1)\sigma^2 E(S^2) + 4n(n+1)\sigma^4}. \tag{6.34}$$

From eq. 6.33 and eq. 6.34, we obtain the following 2^{nd} and 4^{th} order moments of S:

$$\boxed{E(S^2) = E(M^2) - 2n\sigma^2},$$

$$\boxed{E(S^4) = E(M^4) - 4(n+1)\sigma^2 E(M^2) + 4n(n+1)\sigma^4}.$$

Chapter 7

Noise correction methods dedicated to real-time dMRI

In this chapter, we address the feasibility of denoising in real-time (RT). First, we present an adaptation of the LMMSE in the RT framework based on a Kalman filter (KF) described in chapter 4. Second, we propose to implement a denoising filter to be applied in the complex Fourier domain, in order to take benefit of the Gaussian noise distribution present on the real and imaginary channels. Last, we propose an alternative solution based on the use of a parallel Kalman filter able to deal with non-Gaussian noise distributions. All three methods are compared in terms of quality of the denoising, time performances and practical aspects for the user.

7.1 LMMSE integration into the KF-based framework dedicated to RT

This section explains our first RT noise correction method. The latter relies on the integration of the LMMSE, presented in the previous chapter, into the KF-based framework detailed in chapter 4. This RT method constitutes one contribution of this thesis work [Brion et al. (2010, 2011d,f,e)].

7.1.1 Presentation of the RT method

We present here this first RT noise correction method for the aQBI, DTI and sa-aQBI models. The presentation is detailed for the aQBI model, but its extension to DTI and sa-aQBI is straightforward.

Detailed presentation adapted to the aQBI model

Our RT denoising method is represented by the diagram given in fig. 7.1 for the aQBI model. Let us describe what happens to the DW signal $M(\mathbf{v}, \mathbf{o}_i)$ measured at the voxel $\mathbf{v}$ and at the orientation $\mathbf{o}_i$. After having acquired a diffusion-sensitized volume at on

orientation $\mathbf{o}_i$, we get this measured signal $M(\mathbf{v}, \mathbf{o}_i)$ and inject it into the LMMSE. The LMMSE requires the estimation of the noise standard deviation σ. This estimation can be global or voxel-wise defined. For the first case, we performed the empirical tuning, as explained in subsection 6.2.4 of chapter 6, page 151. For the second case, we relied on the method introduced by [Aja-Fernández et al. (2013)] detailed in table 5.1, page 126. The LMMSE enables to obtain the noise corrected DW signal $\hat{S}(\mathbf{v}, \mathbf{o}_i)$ that we inject in the KF, after having divided it by the noise LMMSE corrected T_2-weighted signal. In our case on real data, we directly injected the ratio performed with the averaged T_2-weighted signal, which does not require a correction, as explained in the appendix A. Here, the KF is necessarily used as an OLS estimator because nothing is known about the residual noise after the LMMSE application, and nothing is known *a fortiori* about the noise variance. In this KF process, the residual noise is required to be assumed as a zero-mean Gaussian noise, as explained in chapter 4. The KF adapted to the linear aQBI model relies on the following linear equation:

$$\frac{\hat{S}(\mathbf{v}, \mathbf{o}_i)}{\hat{S}_0(\mathbf{v})} = \mathbf{b_i} \cdot \hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}} + \epsilon_r, \tag{7.1}$$

with $\mathbf{b_i}$ the i^{th} row of $\mathbf{B}$, the matrix of the modified SH basis defined by eq. 3.13, page 46. The index i is the iteration index of the RT process (or equivalently the orientation we are considering). The noise term ϵ_r corresponds to the residual noise after the LMMSE filtering. The KF is used in its regularization form with the adapted initialization of

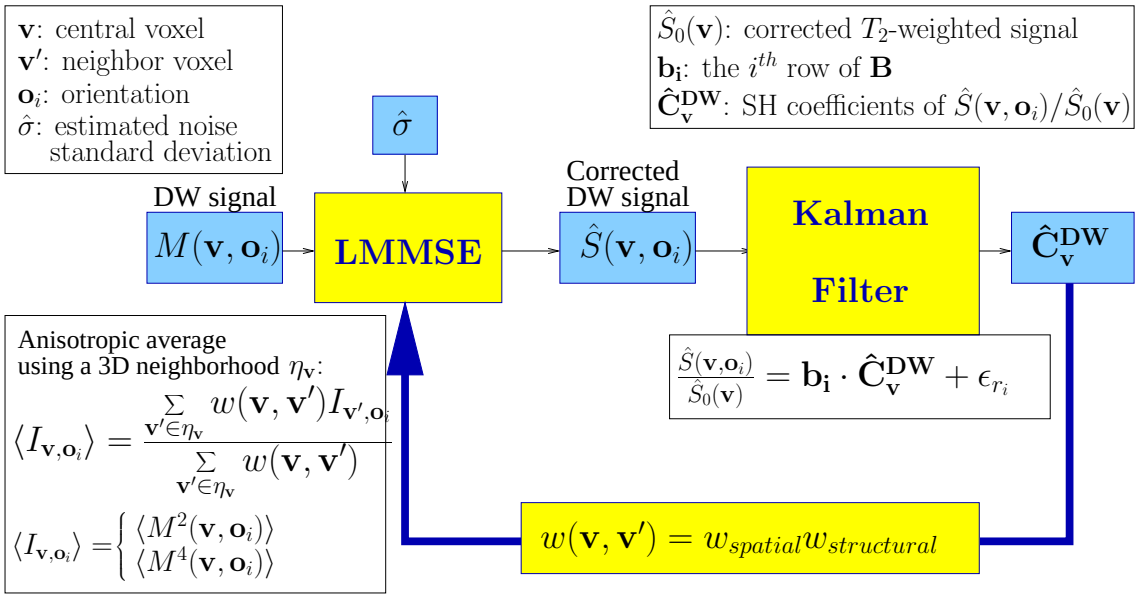


Figure 7.1: Diagram of the RT noise correction algorithm based on an LMMSE and a KF embedded together with a feedback loop. The blocks in yellow indicate the main steps of the algorithm with first the LMMSE, then the KF and finally the feedback loop, which calculates a weight to enable an anisotropic computation of the averages required by the LMMSE. Here the method is adapted for the aQBI model.

the estimation error covariance given in chapter 4. Finally the output of the KF is the coefficients' vector $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$ of the decomposition of the LMMSE estimated DW signals' vector on the modified SH basis. From these coefficients, we can calculate the aQBI maps, but this time, in opposition to what we saw in chapter 4, they will now contain a noise correction. At each new diffusion orientation acquired, the LMMSE followed by the KF will be performed and the vector $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$, as well as the maps, will be **iteratively refined**.

As mentioned previously in chapter 6, a disadvantage of the LMMSE is that it uses averages calculated on neighborhoods. This can lead to oversmoothing effects. As the estimation is embedded in an incremental system, we can use information stemming from the RT results that are refined at each new iteration in the LMMSE to reduce this oversmoothing effect. To this aim, we added a **feedback loop**. It calculates a weight that can be used for the average computations required by the LMMSE: instead of calculating isotropic averages as in the previous chapter, the idea here is to calculate **anisotropic averages** to make the LMMSE more edge-preserving. This weight constrains the LMMSE spatially and structurally for better accuracy. Its effect is similar to an anisotropic diffusion filter. We chose to define this weight as follows:

$$w(\mathbf{v},\mathbf{v}') = \underbrace{\exp\left(\frac{-(\mathbf{v}'-\mathbf{v})^2}{2\alpha^2}\right)}_{w_{SPATIAL}} \underbrace{\exp\left(\frac{-\sum_{j=1}^{N}\left(\hat{\mathbf{C}}_{\mathbf{v}'}^{\mathrm{DW}}(j)-\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}(j)\right)^2}{2\beta^2}\right)}_{w_{STRUCTURAL}}. \tag{7.2}$$

The $w_{SPATIAL}$ component modulates the influence of the neighboring voxels according to their distance from the central voxel, thus preserving the resolution of thin structures. The $w_{STRUCTURAL}$ component is based on an mean square error (MSE) of the SH coefficients between the neighbors and the central voxel. This MSE, presented in chapter 5, is similar to the metric of dODF similarities introduced in [Descoteaux (2008)]: it favors neighboring voxels that have a similar underlying structure. As the diffusion process encodes for the microstructure, the SH decomposition naturally embeds the structural information and any metric defined using the $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$ vector can be used to characterize the structural similarities between two voxels. It is essential that the T_2-weighted signal is incorporated in the $w_{STRUCTURAL}$ component to efficiently characterize the structural similarities between two voxels. The weight $w(\mathbf{v},\mathbf{v}')$ relies on two parameters α and β, that need to be tuned to get the best correction. Their tuning is explained in subsection 7.1.2.

Before going on with the presentation of this RT method for the DTI and sa-aQBI linear models, we want to highlight here that the feedback loop allows us to take advantage at some point of the **joint information** in the DW data along the several diffusion orientations. Our RT constraints do not allow to process all data along all diffusion orientations together, but the addition of the feedback loop offers an alternative enabling to take advantage of the previously acquired measurements to improve the filtering of the current measurement.

The RT method adapted to the DTI and sa-aQBI linear models

Adaptation to the DTI model. To adapt the diagram in fig. 7.1 to the DTI model, the linear equation on which the KF relies has to be replaced by the following linear equation of chapter 4, subsection 4.3.2:

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon_r}, \tag{7.3}$$

with, for iteration i, $y_i = ln\left(\frac{S_0}{\hat{S}(\mathbf{o}_i)}\right)$ (where $\hat{S}(\mathbf{o}_i)$ is the LMMSE estimate of $M(\mathbf{o}_i)$) and $\mathbf{a}_i = b_i[o_x^2, 2o_xo_y, 2o_xo_z, o_y^2, 2o_yo_z, o_z^2]$, b_i being the *b-value*. The state vector $\mathbf{x}$ corresponds to the six LMMSE & KF estimated coefficients of the symmetric diffusion tensor: $\mathbf{x} = [\hat{D}_{xx}, \hat{D}_{xy}, \hat{D}_{xz}, \hat{D}_{yy}, \hat{D}_{yz}, \hat{D}_{zz}]^T$. Finally, the noise residual ϵ_{r_i}, for iteration i, can be considered as before as a zero-mean Gaussian noise. The KF is then performed as an OLS. We want here to point out that [Casaseca-de-la-Higuera et al. (2012)] added some information about this noise to improve our original LMMSE & KF filter. Their approach relies on the empirical study by [Aja-Fernández et al. (2008b)] performed for $n = 1$ and showing that the LMMSE residual noise can be approximated as a Rician noise when $n = 1$. Therefore, they evaluated the variance σ_r of this residual noise and then computed the covariance matrix $\mathbf{R}$ used in the KF, such that, for each iteration i: $R_i = \left(\sigma_r/\hat{S}(\mathbf{o}_i)\right)^2$. Consequently, with this approach, the KF will behave as a WLS estimator, as the matrix $\mathbf{R}$ varies across the iterations. [Casaseca-de-la-Higuera et al. (2012)] demonstrated that the LMMSE & KF technique computed with this WLS filter outperformed the LMMSE & KF technique computed with an OLS filter.

As for the aQBI model, the KF is used in its regularization form with the adapted initialization of the estimation error covariance given in chapter 4. The feedback loop requires an adaptation to the DTI model, such that:

$$w_{STRUCTURAL}(\mathbf{v}, \mathbf{v}') = \exp\left(\frac{-\sum_{j=1}^{6}\left(\hat{\mathbf{D}}_{\mathbf{v}'}(j) - \hat{\mathbf{D}}_{\mathbf{v}}(j)\right)^2}{2\beta^2}\right), \tag{7.4}$$

where $\left(\hat{\mathbf{D}}_{\mathbf{v}'}(j) - \hat{\mathbf{D}}_{\mathbf{v}}(j)\right)^2$ is the l^2-norm of the vector of differences between the two tensors $\hat{\mathbf{D}}_{\mathbf{v}'}$ and $\hat{\mathbf{D}}_{\mathbf{v}}$.

Adaptation to the sa-aQBI model. The adaptation of the RT noise correction method to the sa-aQBI linear model requires also to modify the linear equation for the KF, following again the details given in chapter 4, subsubsection 4.3.3

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon_r}, \tag{7.5}$$

with, for iteration i, $y_i = ln\left(-ln\hat{E}(\mathbf{o}_i)\right)/S_0$ (where $\hat{E}(\mathbf{o}_i) = \hat{S}(\mathbf{o}_i)/S_0$ is the LMMSE & KF estimate of $M_E(\mathbf{o}_i)$)and $\mathbf{a}_i$ corresponding to the i^{th} row of the matrix $\mathbf{B}$ of the

modified SH basis. The state vector $\mathbf{x}$ corresponds to the LMMSE estimated coefficients' vector $\hat{\mathbf{C}}^{\mathrm{SA}}$. Finally, the noise residual ϵ_{r_i}, for iteration i, can be considered as before as a zero-mean Gaussian noise. As for the DTI model, the KF can be computed as an OLS assuming a non-varying variance of the residual noise across the iterations ($\mathbf{R} = \mathbf{I}$), or as an WLS estimator computing the covariance matrix $\mathbf{R}$, such that: $R_i = \frac{\sigma_r^2}{\hat{S}(\mathbf{o}_i)^2 ln\left(\hat{S}(\mathbf{o}_i)/S_0\right)^2}$, with σ_r the estimated variance of the residual noise. For $n = 1$, we can estimate the latter assuming that the residual noise is Rician, as done in [Casaseca-de-la-Higuera et al. (2012)] for the LMMSE & KF application on the DTI model.

As for the aQBI model, the KF is used in its regularization form with the adapted initialization of the estimation error covariance given in chapter 4. The feedback loop requires an adaptation to the sa-aQBI model, such that:

$$w_{STRUCTURAL}(\mathbf{v}, \mathbf{v}') = \exp\left(\frac{-\sum_{j=1}^{N}\left(\hat{\mathbf{C}}^{\mathrm{SA}}_{\mathbf{v}'}(j) - \hat{\mathbf{C}}^{\mathrm{SA}}_{\mathbf{v}}(j)\right)^2}{2\beta^2}\right). \tag{7.6}$$

Let us now explain how we tuned α and β.

7.1.2 Tuning of the α and β parameters

We performed the tuning of α and β empirically. For α, we simply looked at different configurations of $w_{SPATIAL}$ for different α, as represented in fig. 7.2. The tuning of α was performed considering a 5×5-voxel neighborhood in 2D. α was chosen to produce a good compromise between a too low spatial weight, leading to insufficient neighboring information, and a too high spatial weight, yielding excessive smoothing. We chose $\alpha = 2$.

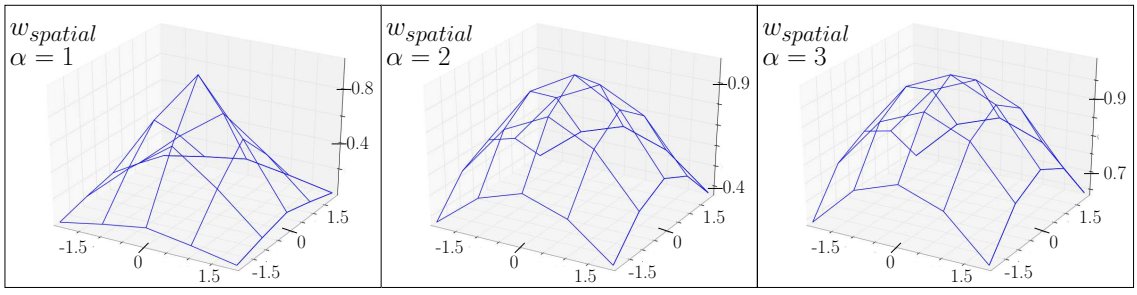


Figure 7.2: Configurations of $w_{SPATIAL}$ for different values of α. The central point of the mesh corresponds to the voxel, which is considered for the LMMSE step. Each other point of the mesh corresponds to a neighbor used in the average computation for the LMMSE.

Then, we tuned β on simulated DW data (for more details on the generation of these data, see appendix C, at the end of the manuscript) by choosing the value that produced the lowest MSE between the noise-free simulation and the corrected one. This MSE was calculated, as in the previous chapter, in the region shown in fig. 7.5. For the application

of our algorithm, which we call "LMMSE & KF", on these data, we used the real number of receiver channels n in the nc-χ LMMSE, as no noise correlation was involved in the simulated data. The true noise variance σ was used, so that the tuning of β only depends on the RT denoising algorithm and not on the variance estimation. The maximum SH order N was set to 8 and the Laplace-Beltrami regularization λ was first set to 0 (no regularization) and then to 0.006 (regularization) in the initialization of the estimation error covariance matrix defined by $\mathbf{P_0} = ((1/V)\mathbf{I} + \lambda\mathbf{L})^{-1}$ (as in chapter 4). The initial weight at the first iteration was defined such that: $w(\mathbf{v}, \mathbf{v}') = w_{spatial}(\mathbf{v}, \mathbf{v}')$.

Fig. 7.3 shows the curve of the MSE for several values of β tested on the simulation case at $b = 4500s \cdot mm^{-2}$ with the RT correction method adapted to the aQBI model without and with regularization. The minimum of MSE was obtained without regularization at $\beta = 0.13$. With regularization, it was obtained at $\beta = 0.05$. It is visible that there exists another local minimum for this latter configuration. We will come back to it later.

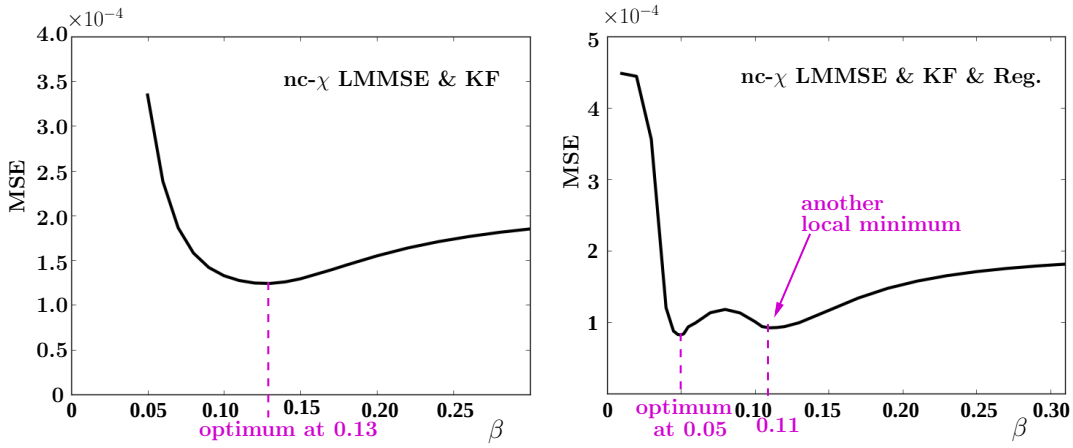


Figure 7.3: Tuning of β on a simulated data for the LMMSE & KF adapted to aQBI without (left) and with (right) regularization (indicated as "& Reg."). The LMMSE & KF method was applied on the simulated data corrupted by nc-χ noise at $b = 4500s \cdot mm^{-2}$ with $\sigma = 16$. A panel of β values were tested. The optimum β was found as the one yielding the smallest MSE.

Fig. 7.4 shows the optimum β found for different noise levels at $b = 4500s \cdot mm^{-2}$ without and with regularization. Without the regularization, for $\sigma \geq 25$, the optimum β was found much higher than at lower noise levels. We can see that the curve presents a jump between $\sigma = 22.5$ and $\sigma = 25$. For the highest noise levels, as β_{opt} increases, $w_{STRUCTURAL}$ approaches one: all neighbors are then considered equally, without *a priori* discrimination and the LMMSE & KF becomes equivalent to the LMMSE alone without any feedback loop. At this point, there is so much noise in the data that the LMMSE & KF requires as much neighboring information as available to denoise the data. The curve in fig. 7.4, right, obtained for the algorithm performed with regularization, does not contain a jump like for the curve on the left. It is likely that the regularization already decreases the noise level. In this case, the feedback is still used at high noise levels. The

σ values corresponding to the green dotted square refer to cases like the one presented in fig. 7.4 on the right: for these σ values, two local MSE minima were found. One was the global minimum and yielded β_{opt}. The other local minimum was found greater (equal to 0.11). The latter was shown to be more adapted for our real data, as we will discuss later.

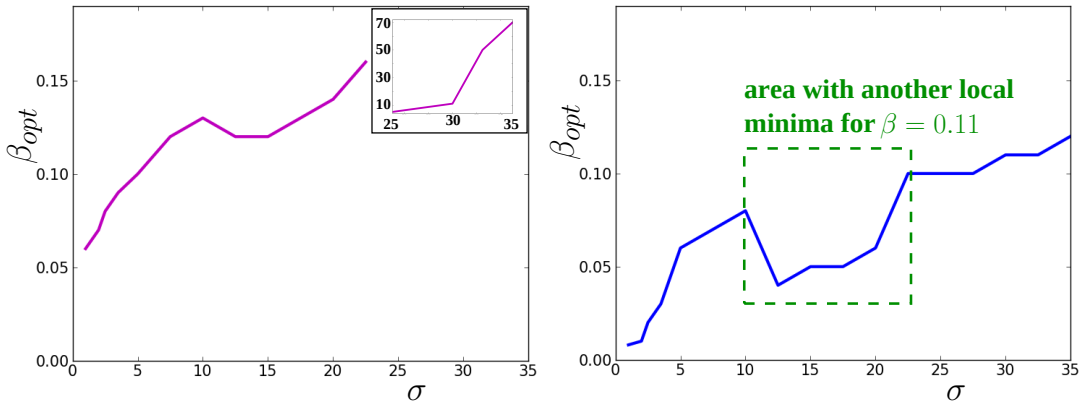


Figure 7.4: The optimum β, designed by β_{opt}, plotted against σ at $b = 4500s \cdot mm^{-2}$ for the LMMSE & KF without and with regularization.

Finally, we performed a similar tuning of β on other noisy configurations with $\sigma = 16$ and $n = 4$ at $b = 1500s \cdot mm^{-2}$, $b = 3000s \cdot mm^{-2}$, $b = 4500s \cdot mm^{-2}$ and $b = 6000s \cdot mm^{-2}$. The values of β are presented for the aQBI model in table 7.1 for the algorithm performed without regularization and in table 7.2 for the algorithm performed with regularization. It is visible that the β_{opt} value found without regularization at $b = 6000s \cdot mm^{-2}$ corresponds to the high noise level situation discussed earlier for fig. 7.4. As previously, when using the regularization, this value is much lower.

LMMSE & KF without regularization ($\alpha = 2$)				
b $(s \cdot mm^{-2})$	1500	3000	4500	6000
β_{opt} (aQBI) $(N = 8)$	0.24	0.21	0.13	10

Table 7.1: The optimum values of β for the first presented RT noise correction method called "LMMSE & KF" adapted to the DTI, aQBI and sa-aQBI models. The parameter α was set to 2 in all cases. The β values were obtained on simulated data corrupted by nc-χ noise with $\sigma = 16$ at the four following *b-values*: $b = 1500s \cdot mm^{-2}$ / $b = 3000s \cdot mm^{-2}$ / $b = 4500s \cdot mm^{-2}$ / $b = 6000s \cdot mm^{-2}$. No regularization was used in the RT process.

LMMSE & KF with regularization ($\alpha = 2$)				
b $(s \cdot mm^{-2})$	1500	3000	4500	6000
β_{opt} (aQBI) $(N = 8)$	0.10	0.11	0.05	0.06

Table 7.2: The optimum values of β for the first presented RT noise correction method called "LMMSE & KF" adapted to the DTI, aQBI and sa-aQBI models. The parameter α was set to 2 in all cases. The β values were obtained on simulated data corrupted by nc-χ noise with $\sigma = 16$ at the four following *b-values*: $b = 1500s \cdot mm^{-2}$ / $b = 3000s \cdot mm^{-2}$ / $b = 4500s \cdot mm^{-2}$ / $b = 6000s \cdot mm^{-2}$. Regularization was used in the RT process.

The next subsection shows the results obtained on these simulated data. The comparison between the use of regularization or not is presented.

7.1.3 Results on simulated data & discussion

We compared the RT anisotropic version of the nc-χ LMMSE introduced in this chapter with the previous isotropic nc-χ LMMSE presented in chapter 6. We refer to them as the "LMMSE & KF" and the "LMMSE", respectively. We computed both methods using a 5×5×5 neighborhood. Before running the two methods, the noise variance σ was estimated using the "mode M1-χ" method resumed by eq. 5.26 (page 124) with $n = 4$, as there was no correlation in our data. We used the same four different simulations as in the previous subsection. They are all corrupted by a nc-χ noise with $\sigma = 16$ at $b = 1500s \cdot mm^{-2}$, $b = 3000s \cdot mm^{-2}$, $b = 4500s \cdot mm^{-2}$ and finally $b = 6000s \cdot mm^{-2}$. We show the results for the aQBI model, for which the maximum SH order N was set to 8 and the Laplace-Beltrami regularization λ was first set to 0 (no regularization) and then to 0.006 (regularization) in the initialization of the estimation error covariance matrix defined by $\mathbf{P_0} = ((1/V)\mathbf{I} + \lambda\mathbf{L})^{-1}$ (as in chapter 4). Fig. 7.5 shows the results on the dODF fields at $b = 4500s \cdot mm^{-2}$. On these simulated data, the nc-χ LMMSE & KF, both without and with regularization, remarquably corrects the smoothing effect of the isotropic nc-χ LMMSE which we had already mentioned in chapter 5. With the LMMSE & KF, the crossing bundles have similar width as in the noise-free configuration. Moreover, as shown by the zoomed dODF, the crossing is retrieved closer to the noise-free crossing: it is

particularly visible that the two lobes of the zoomed dODF are more equally proportioned than on the dODF obtained with the nc-χ LMMSE alone. Finally, the result of the nc-χ LMMSE & KF performed with regularization is visually the closest to the noise-free image. The improvement brought by the regularization is higher compared to its effect on the nc-χ LMMSE. This is due to the fact that, with the LMMSE & KF, the regularization is incorporated during the RT denoising process and therefore does affect each iteration of the algorithm which was not the case with the LMMSE, for which the regularization was only applied at the end of the process to obtain the dODF results.

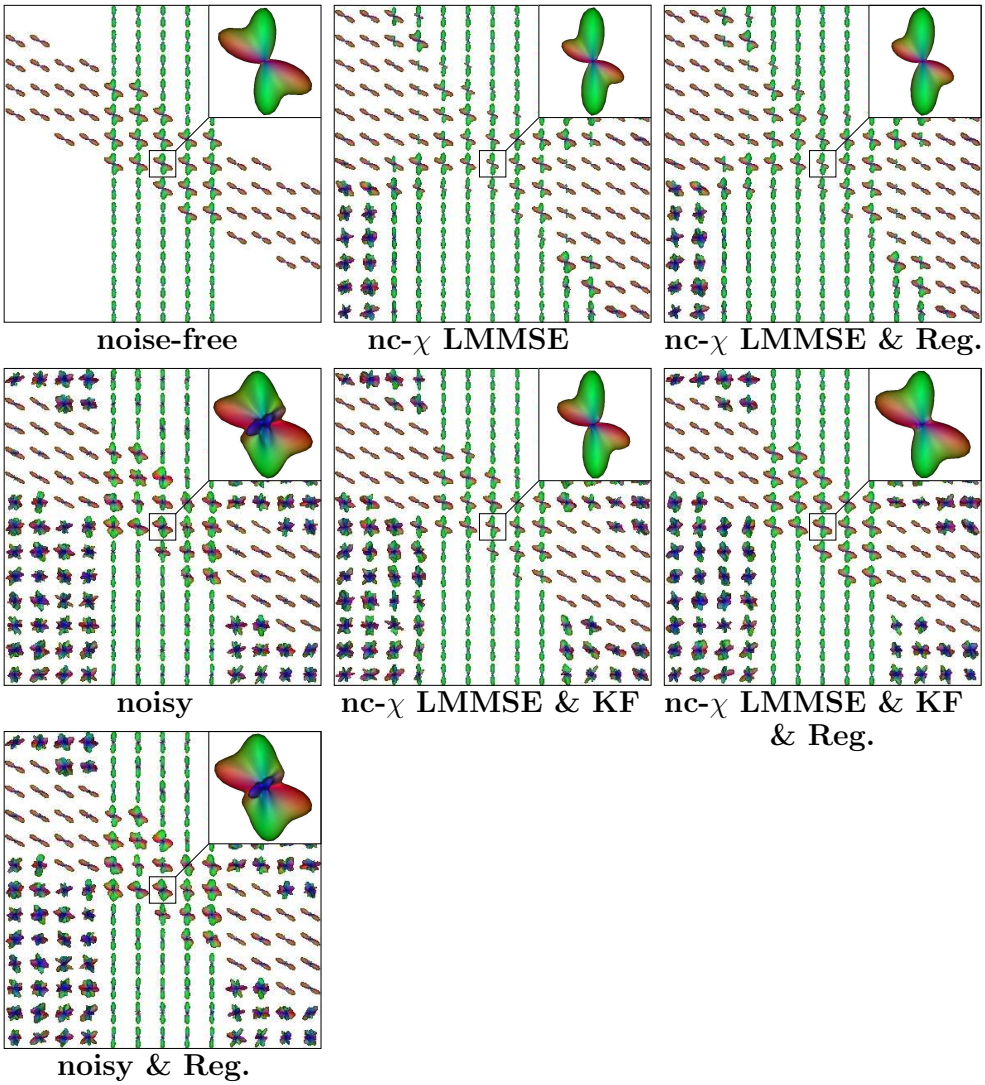


Figure 7.5: Comparison between the nc-χ LMMSE and the nc-χ LMMSE & KF on simulated DW data corrupted by nc-χ noise with $\sigma = 16$ at $b = 1500s \cdot mm^{-2}$. The dODF fields were generated from the noise-free, the noisy and the results of both methods without and with regularization. On each field a zoom of a dODF from the crossing region is displayed.

These qualitative results are confirmed by the graph in fig. 7.6. Concerning the nc-χ LMMSE & KF performed without regularization, it yields a lower MSE for $\sigma \leq 20$.

For higher noise levels, it does not bring much improvement —even no improvement sometimes— compared to the nc-χ LMMSE without regularization, and even less improvement compared to the nc-χ LMMSE with regularization. We want here to point out that the slightly chaotic aspect of all curves for $\sigma > 20$ is due to a less accurate estimation of σ in the correction methods. This less precise σ estimation also explains the fact that at $\sigma = 22.5$, the nc-χ LMMSE & KF yiels a very slightly higher error than the nc-χ LMMSE. Indeed, the β_{opt} found for $\sigma = 22.5$ is the not the most appropriate here, as the noise variance was not accurately estimated. Finally, this graph shows that the nc-χ LMMSE & KF performed with regularization outperforms all other methods.

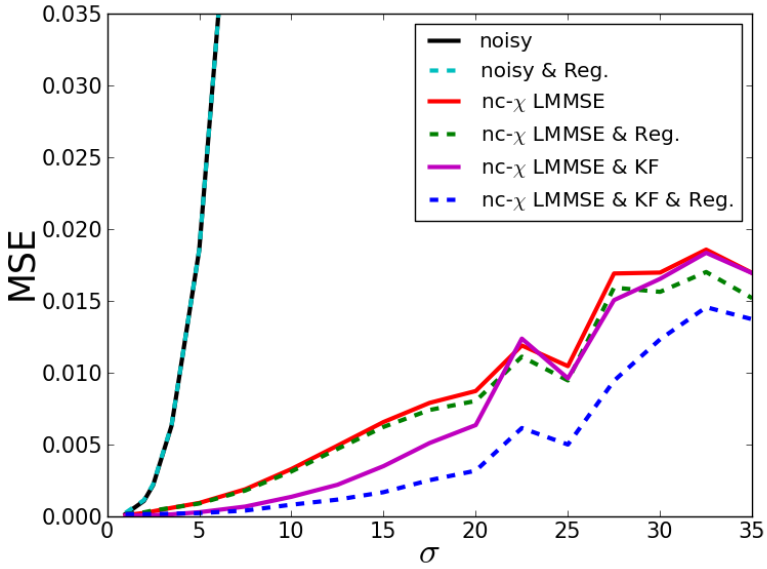


Figure 7.6: Comparison between the MSE of the noisy dODF field without and with regularization and the MSE of the dODF fields after the nc-χ LMMSE and after the nc-χ LMMSE & KF also without and with regularization. The noisy MSE is of 3000 at $\sigma = 35$.

We performed a similar comparison of the MSE on the four configurations at $b = 1500s \cdot mm^{-2}$, $b = 3000s \cdot mm^{-2}$, $b = 4500s \cdot mm^{-2}$ and finally $b = 6000s \cdot mm^{-2}$ with $\sigma = 16$, corresponding to real variance values measured on real data. These four simulated configurations are therefore close to real ones. Fig. 7.7 highlights again the improvement obtained by the anisotropic LMMSE & KF method, especially when regularization was used. The regularization brings a robustness regarding high *b-values*, as well as high noise levels.

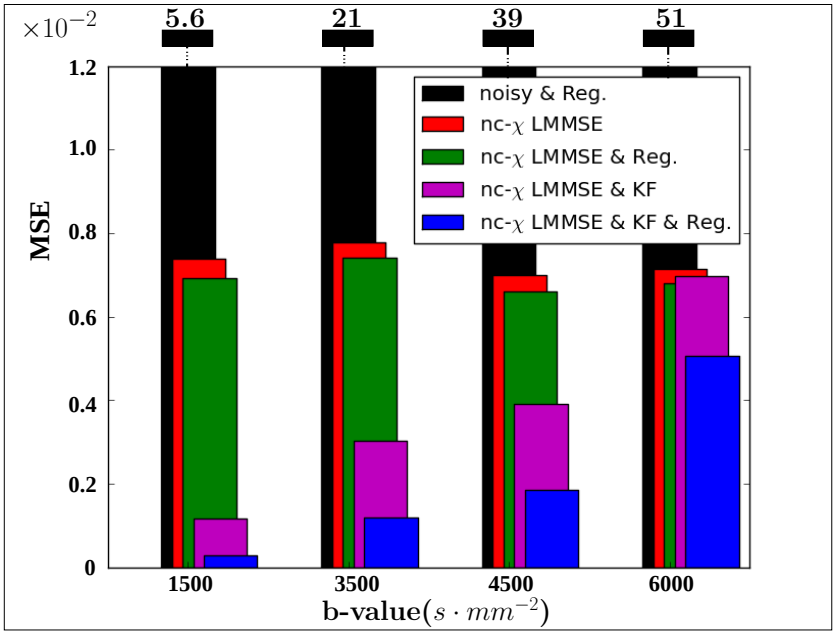


Figure 7.7: Comparison between the MSE of the noisy dODF field without and with regularization and the MSE of the dODF fields after the nc-χ LMMSE and after the nc-χ LMMSE & KF also without and with regularization. The comparison was performed at different *b-values* with $\sigma = 16$.

After having validated this first RT method on our simulated data, we applied it on our real data.

7.1.4 Results on real data & discussion

Results at the end of the RT process & discussion.

We performed a similar comparison on the real DW data acquired at $b = 1500s \cdot mm^{-2}$, $b = 3000s \cdot mm^{-2}$, $b = 4500s \cdot mm^{-2}$ and finally $b = 6000s \cdot mm^{-2}$. For more details on the acquisition parameters of these real data, we refer the reader to the appendix A at the end of the manuscript. As for the simulated data, we performed all denoising methods using a 5×5×5 neighborhood. The value of the α parameter was 2, similarly as before. Concerning the β parameter, we reported the values of table 7.2 for $b = 1500s \cdot mm^{-2}$ and $b = 3000s \cdot mm^{-2}$. For $b = 4500s \cdot mm^{-2}$ and $b = 6000s \cdot mm^{-2}$, the β_{opt} obtained for the simulated data were too low for our real data. Therefore, we prefered choosing the value of 0.11 for β, corresponding to the second local minimum described in the previous subsection (see fig. 7.4). To summarize, the values of β used for our real data are 0.10 / 0.11 / 0.11 / 0.11 for $b = 1500/3000/4500/6000s \cdot mm^{-2}$, respectively. We performed the LMMSE & KF technique with the maximum SH order N set to 8 and the Laplace-Beltrami regularization λ was set to 0.006 in the initialization of the estimation error covariance matrix of the KF. We run the algorithms with regularization only, as regularization was shown to increase the quality of the maps. To have an idea of the GFA maps obtained with the LMMSE & KF without regularization, we refer the reader to fig. 6.9 in the previous

chapter. The latter figure shows some intensity artifacts obtained without regularization that we similarly have when using the LMMSE & KF without regularization on our real data.

As in the previous chapter for the LMMSE method, we estimated the noise variance σ accounting for the possible noise correlations in the real data using effective parameters calculated globally for the whole DW volume. The resulting effective parameters for the four *b-values* are shown in table 7.3. We notice that the values of n_{eff} found are lower than the ones obtained for the isotropic LMMSE in table 6.1, page 163. We saw in the previous chapter that when n_{eff} decreases, the LMMSE method yields a higher smoothing effect. However, here, this effect will be much more controlled by the feedback loop inherent to the LMMSE & KF technique. Therefore, lower values of n_{eff} will not produce smoothed results, as it would have done with the isotropic LMMSE. We also compared our RT method using this trade-off with the nc-χ LMMSE and with the nc-χ LMMSE & KF, both performed using voxel-wise effective parameters.

Global effective n_{eff} and σ_{eff} obtained for the nc-χ LMMSE & KF				
b $(s \cdot mm^{-2})$	1500	3000	4500	6000
n_{eff}	1.8	1.8	2.6	2.4
σ_{eff}	26.0	25.5	20.7	21.4

Table 7.3: The global effective n_{eff} and σ_{eff} obtained for the nc-χ LMMSE & KF at $b = 1500s \cdot mm^{-2}$ / $b = 3000s \cdot mm^{-2}$ / $b = 4500s \cdot mm^{-2}$ / $b = 6000s \cdot mm^{-2}$. Regularization was used in the RT process.

First, we show in fig. 7.8 the results obtained with the nc-χ LMMSE and with the nc-χ LMMSE & KF in their global and voxel-wise versions performed with regularization. The DW images obtained with our RT method, either with the global or the voxel-wise solution, better reveal the underlying anatomical structure than the other ones. The features are better retrieved than with the non RT denoising methods (see the yellow circle) thanks to the feedback loop. This feedback loop uses the anatomical knowledge from the T_2-weighted signal to discriminate the neighbors in the average computations of the LMMSE. Therefore, the DW images obtained with the nc-χ LMMSE & KF gives a better perception of the anatomical information contained in the raw DW image. Between the global and the voxel-wise versions of our RT technique, there is more smoothing effect with the first one. This smoothing effect appears in homogeneous regions, without corrupting the features definitions.

Fig. 7.9 shows the same comparison as in fig. 7.8 on the GFA maps at $b = 1500/3000/4500/6000s \cdot mm^{-2}$. As in the previous figure, it is here visible that the use of the feedback loop improves the definition of the anatomical structures, especially for the global version of the nc-χ LMMSE & KF. Concerning the voxel-wise nc-χ LMMSE & KF, it yields more accurate results compared to the voxel-wise nc-χ LMMSE for $b = 1500s \cdot mm^{-2}$ and

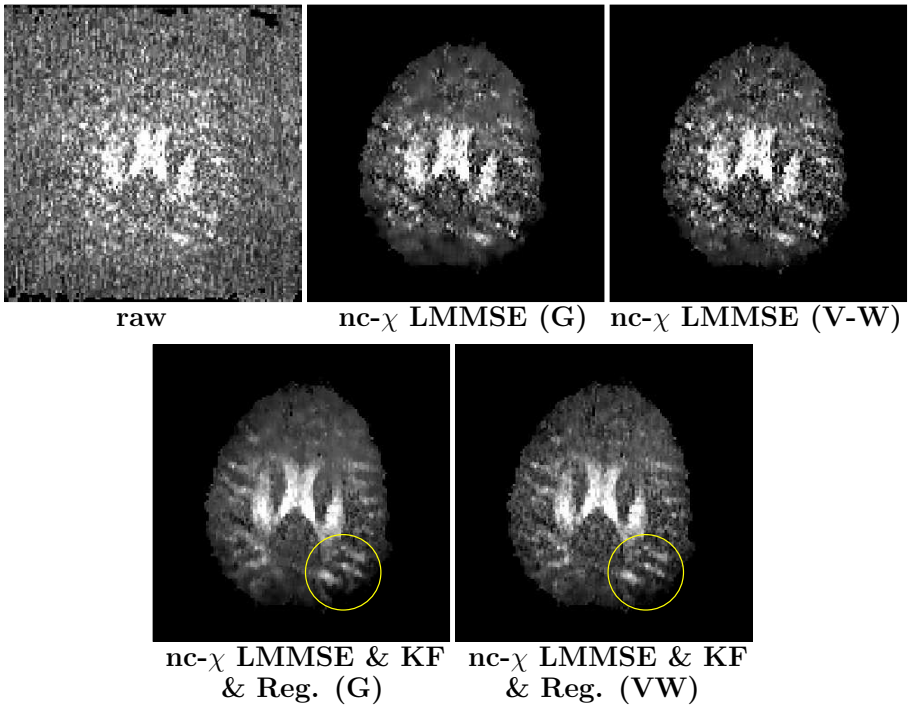


Figure 7.8: Comparison between the nc-χ LMMSE & KF (G) on real DW data at $b = 4500s{\cdot}mm^{-2}$ with either the global (G) or the voxel-wise (V-W) effective parameters with the nc-χ LMMSE. The yellow circle indicates the anatomical structures depicted by the nc-χ LMMSE & KF technique thanks to the feedback loop.

$b = 3000s \cdot mm^{-2}$. However at higher *b-values*, although the GFA signal is globally higher, the noise reduction is not much improved by the feedback loop. This is explained by the fact that the voxel-wise LMMSE analyzed in the previous chapter yielded a high n_{eff} value at $b \geq 4500s \cdot mm^{-2}$, which did not produce much smoothing effect. Therefore, the feedback loop, in this case, cannot bring much improvement, as there is originally not enough smoothing effect. The global nc-χ LMMSE & KF produced a higher contrast on all maps. The hyperintensity artifact appearing at $b = 6000s \cdot mm^{-2}$ for the global version of the nc-χ LMMSE is much reduced with the nc-χ LMMSE & KF. This latter method generated maps with finely preserved anatomical structures and noise removal yielding an improved visual contrast.

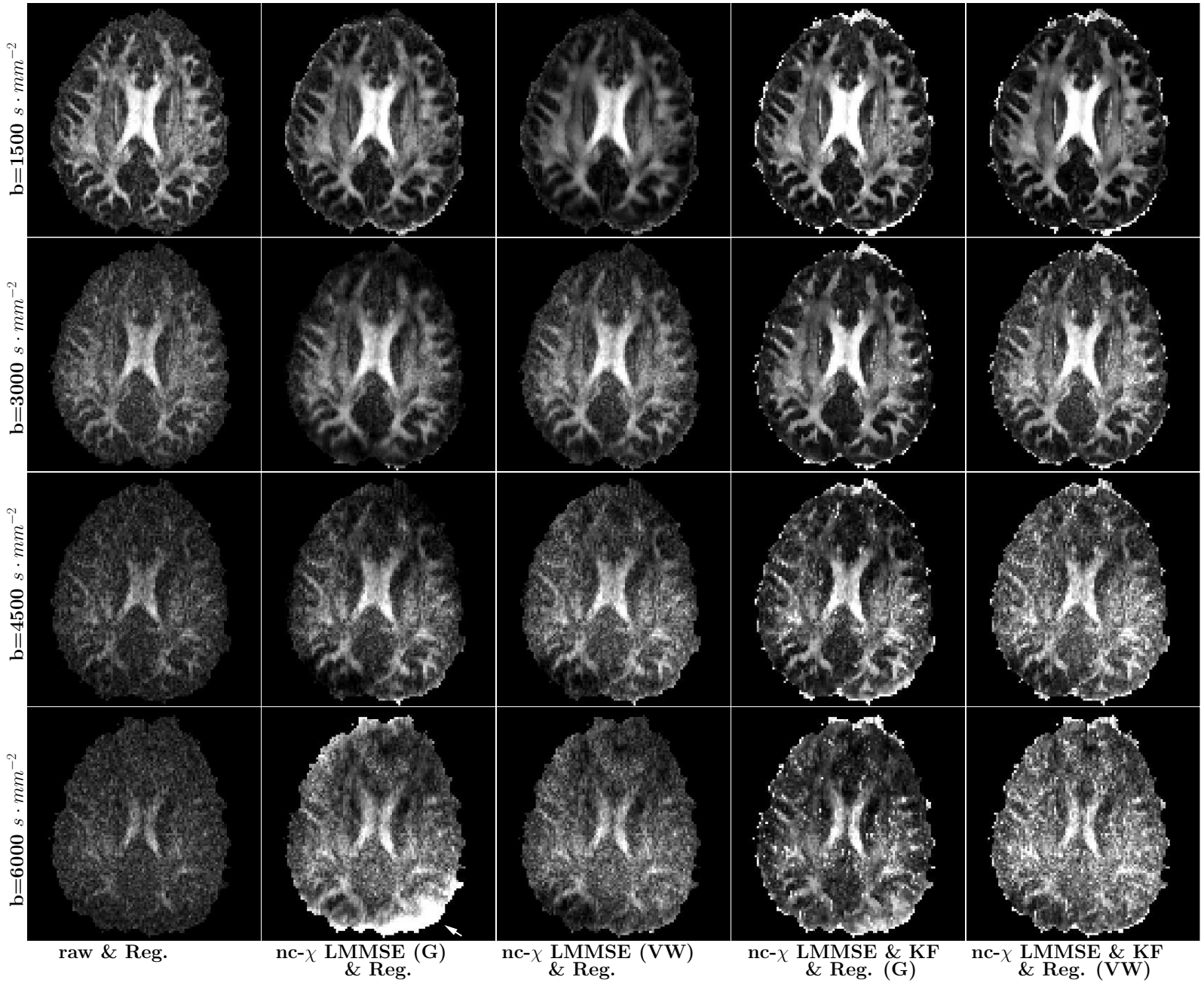


Figure 7.9: Comparison between the nc-χ LMMSE on real GFA maps with regularization ("& Reg.") and with either the global (G) or the voxel-wise (V-W) effective parameters and with the nc-χ LMMSE & KF at $b = 1500/3000/4500/6000 s \cdot mm^{-2}$. Hyperintensities in the maps are shown by white arrows.

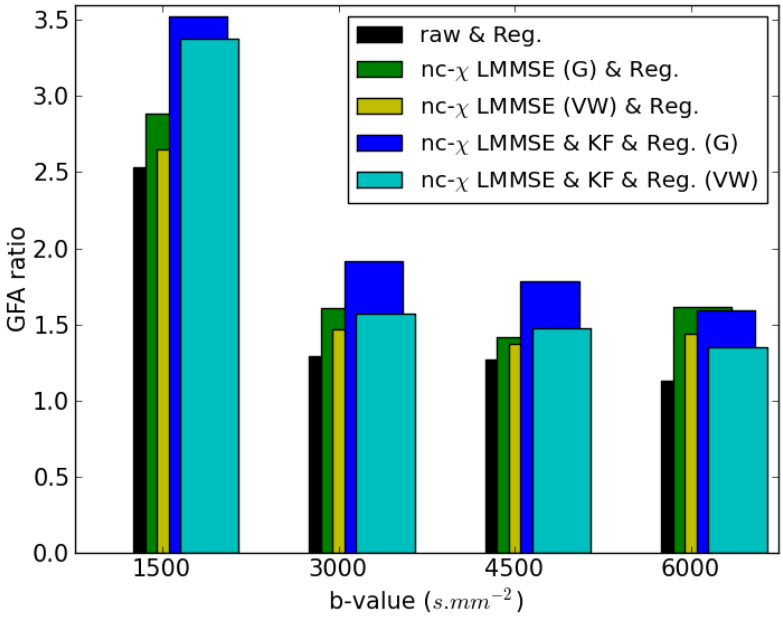


Figure 7.10: Comparison between the GFA ratios of the raw map and the map obtained after application of nc-χ LMMSE and the nc-χ LMMSE & KF with regularization and with the global effective parameters at $b = 1500/3000/4500/6000s \cdot mm^{-2}$. The ROI were chosen as shown in fig. 6.7, in the previous chapter.

To quantitatively confirm our results, we calculated the GFA ratios on each configuration (fig. 7.10). The ROI were chosen as before in fig. 6.7. This graph shows that all methods yielded at all *b-values* a higher GFA ratio than the raw one. For $b = 1500/3000/4500s \cdot mm^{-2}$, the highest GFA ratio is obtained for the nc-χ LMMSE & KF, in its global version, validating our qualitative results. At $b = 6000s \cdot mm^{-2}$, the GFA ratio of this method is very slightly smaller than the ratio obtained with the nc-χ LMMSE global technique. However, it is still very high and also confirms the gain in GFA contrast visible in fig. 7.9. Moreover, we have to keep in mind that the GFA ratio is a partial index of image quality, as it does not indicate if the structures are preserved or not after the application of the method which is essential to know to assess the quality of the denoising technique.

Fig. 7.11 compares again all methods on the RGB maps calculated in the aQBI model. This figure confirms the previous results. And finally, fig. 7.12 and fig. 7.13 present the effects of the global version of the LMMSE & KF method on zoomed dODF maps at $b = 1500/3000s \cdot mm^{-2}$ and $b = 4500/6000s \cdot mm^{-2}$, respectively. The results obtained with our RT denoising method are more coherent (see the black squares) than the raw dODF maps, while respecting the raw anatomy configuration.

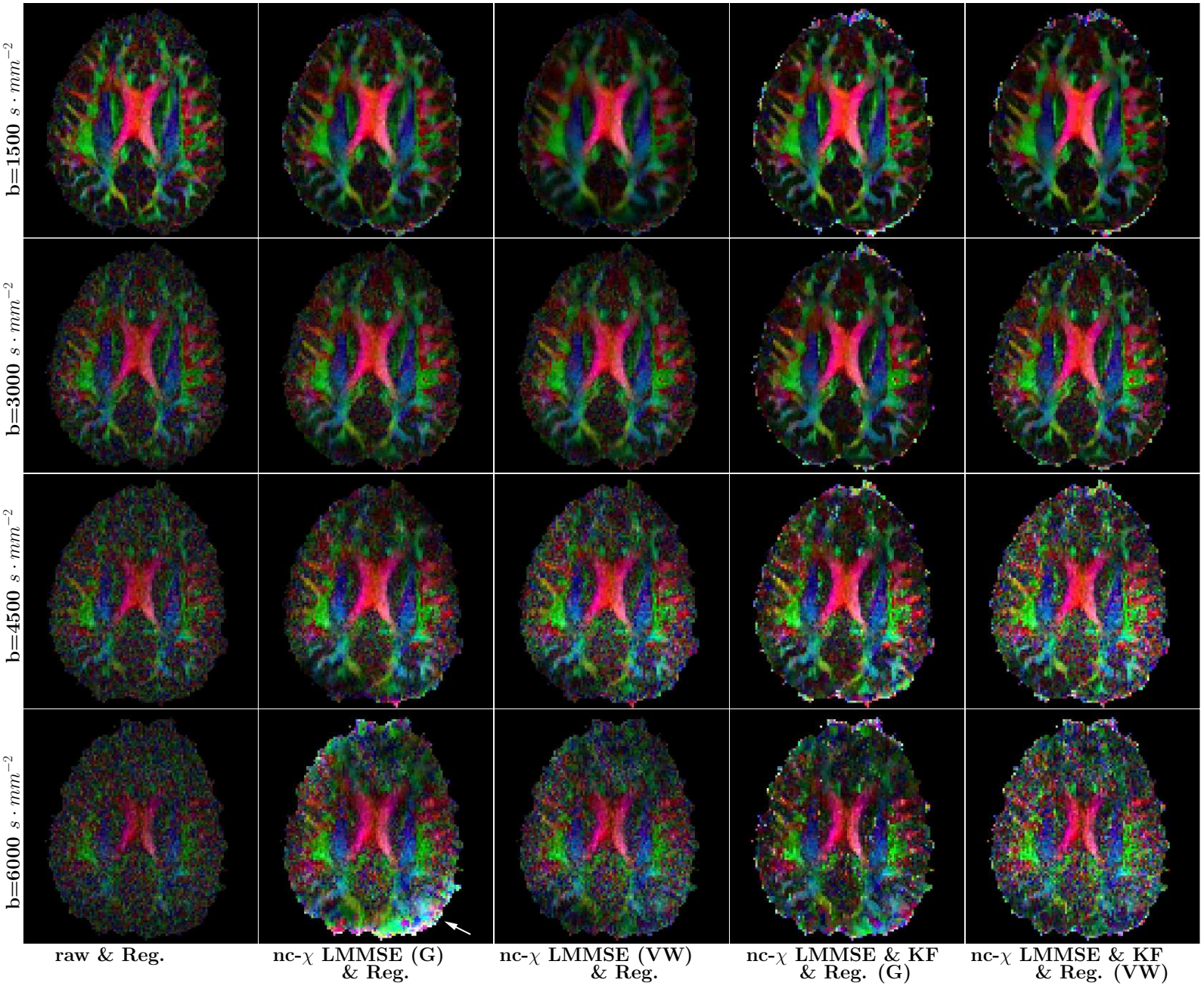


Figure 7.11: Comparison between the nc-χ LMMSE on real RGB maps with regularization ("& Reg.") and with either the global (G) or the voxel-wise (V-W) effective parameters and with the nc-χ LMMSE & KF at $b = 1500/3000/4500/6000 s \cdot mm^{-2}$. Hyperintensities in the maps are shown by white arrows.

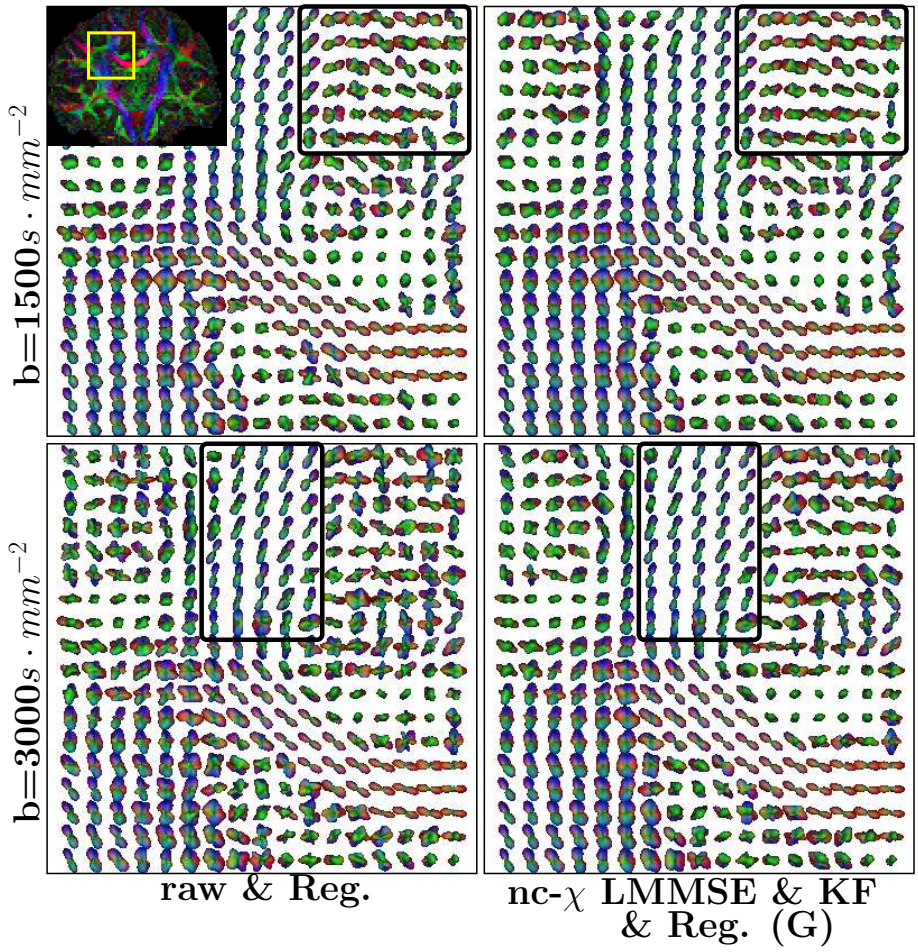


Figure 7.12: Comparison between the raw real dODF map with the dODF map obtained after our nc-χ LMMSE & KF with regularization and with the global effective parameters at $b = 1500/3000s \cdot mm^{-2}$. The maps were performed in the yellow region indicated in the RGB coronal slice.

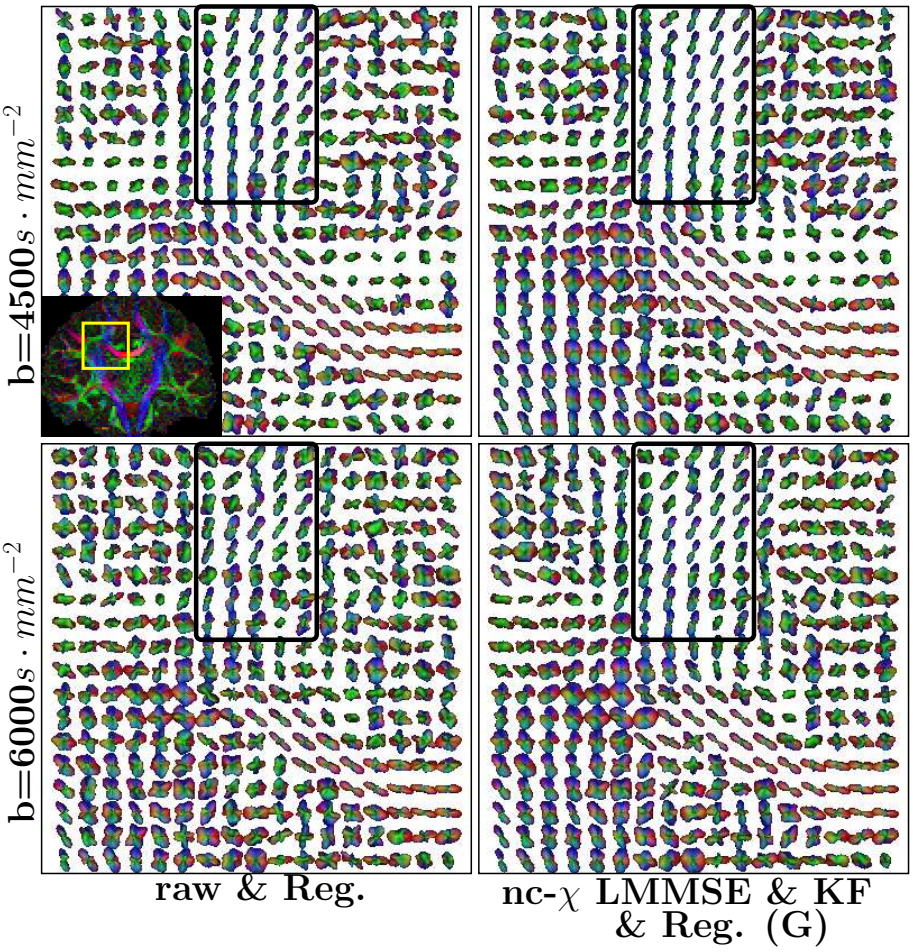


Figure 7.13: Comparison between the raw real dODF map with the dODF map obtained after our nc-χ LMMSE & KF with regularization and with the global effective parameters at $b = 4500/6000s \cdot mm^{-2}$. The maps were performed in the yellow region indicated in the RGB coronal slice.

Results along the RT process & discussion.

Here we show the results with our nc-χ LMMSE & KF method in its global version along the RT process on the real data at $b = 3000s \cdot mm^{-2}$. These results were obtained for the aQBI model with the maximum SH order N first set to 4 (fig. 7.14) and then to 6 (fig. 7.15) and to 8 (fig. 7.16). The Laplace-Beltrami regularization λ was set to 0.006, as previously. These results can be compared to the aQBI maps (GFA and dODF) without noise correction that were presented in fig. 4.14, page 91 (chapter 4).

From these three figures, it is visible that the RT noise correction visually better brings to light the structural information present in the GFA maps. At the maximum SH order of 4 (fig. 7.14), both raw and denoised RT processes achieve a good quality of the GFA map from early iterations. It is particularly visible on this figure that the RT noise correction brings a visual improvement, which can also be observed on the dODF map. The latter map contains more coherent and denser fiber bundles with the noise correction, while respecting the raw anatomy configuration. Concerning the fig. 7.15 and 7.16, to obtain accurate GFA and dODF maps, one has to wait until the 29^{th} and the 60^{th} iterations, respectively, without or with denoising. This requirement for more iterations to obtain accurate maps is due to the fact that the number of components (the *unknowns*) in the coefficients' vector $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$ is equal to 28 and 45 for a maximum SH order set to 6 and 8, respectively. Therefore, we need enough measurements —at least 28 and 45 iterations, plus the T_2-weighting measurement— to rightly estimate $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$. The noise correction process therefore does not permit to accelerate the process of getting maps of sufficient quality. However, the denoising is still efficient, as the structural information is better depicted with the application of the "LMMSE & KF" filter.

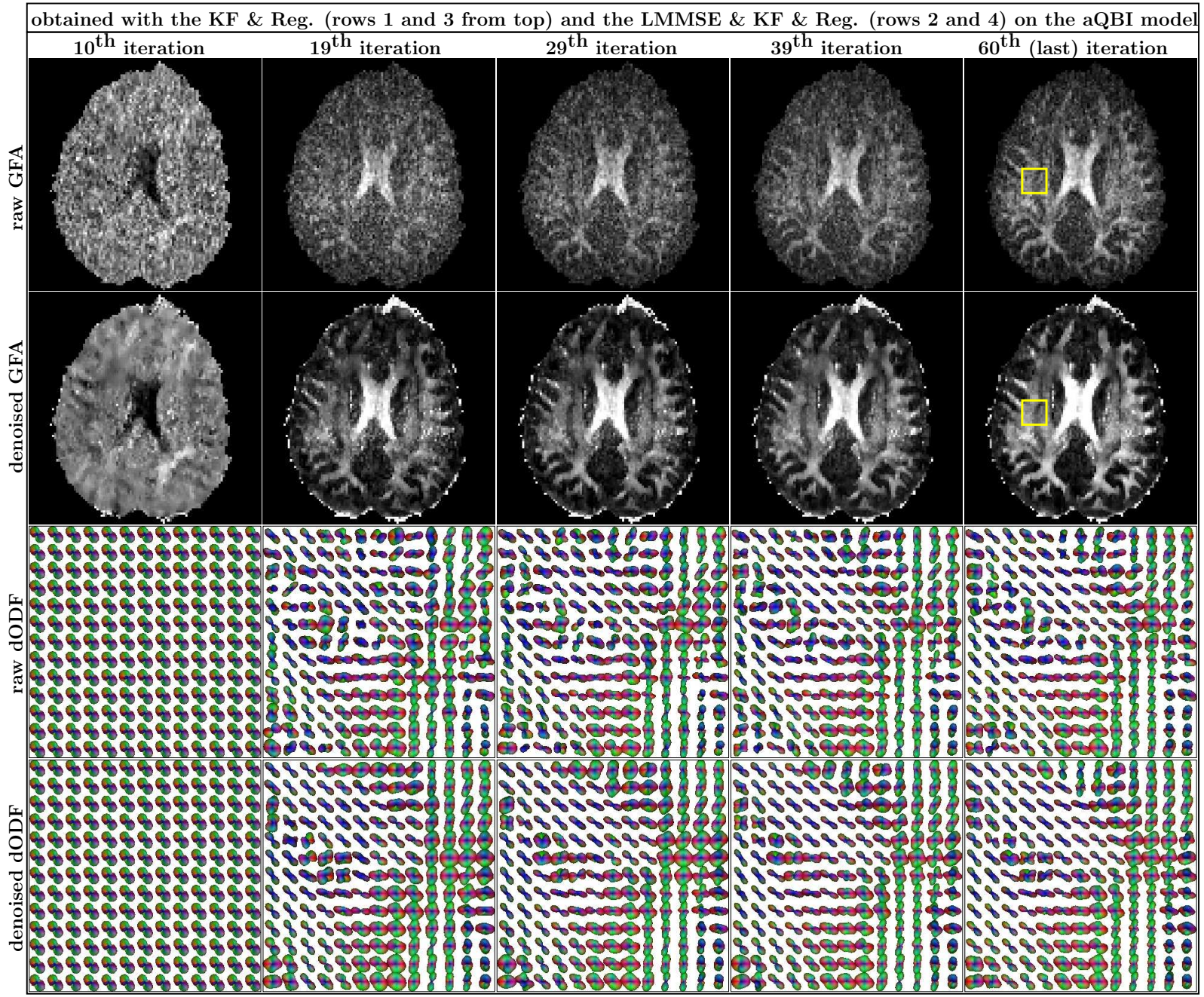


Figure 7.14: LMMSE & KF results along the RT process at $b = 3000s \cdot mm^{-2}$ for the aQBI model. The method was performed with $N = 4$ and $\lambda = 0.006$ (regularization).

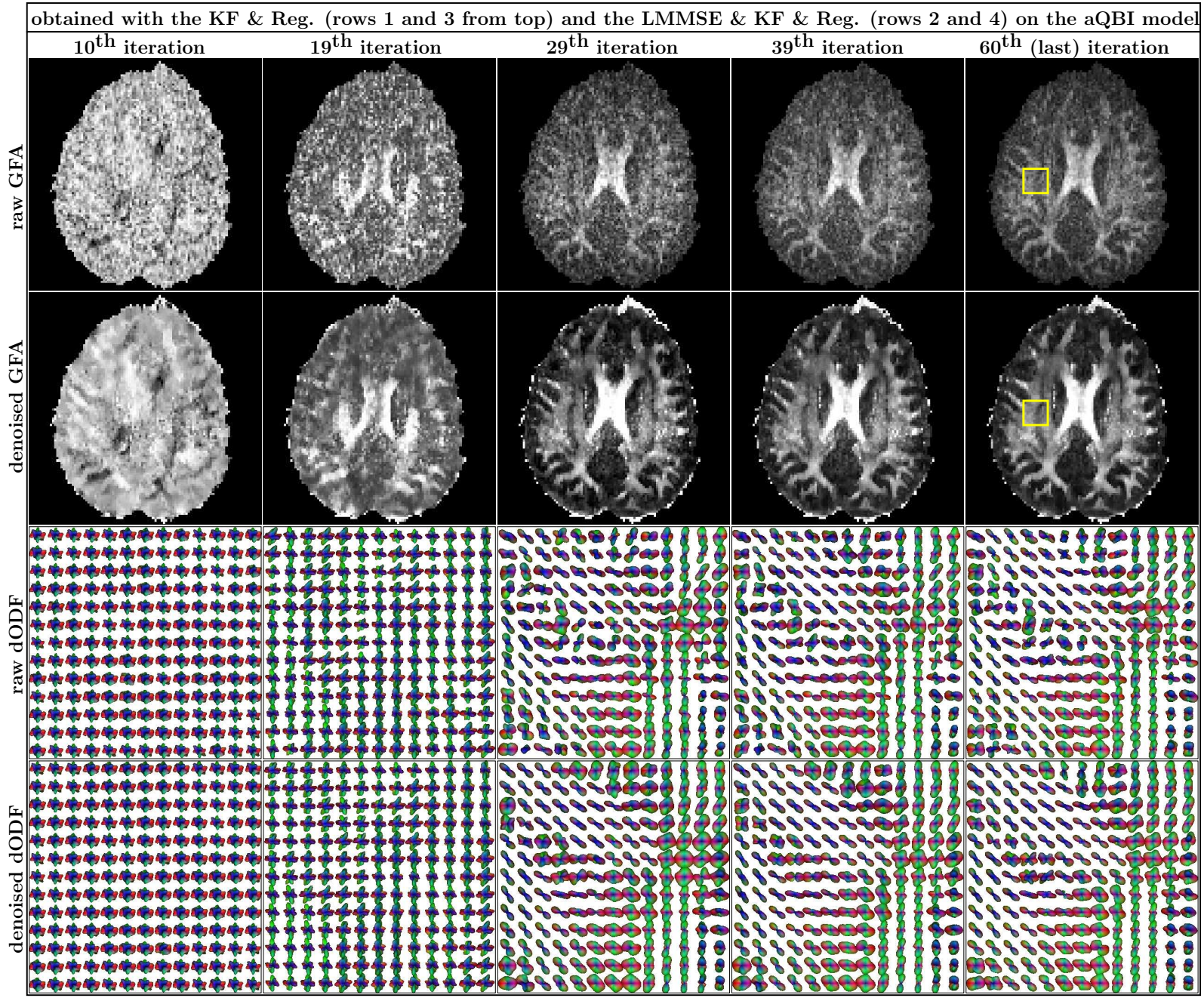


Figure 7.15: LMMSE & KF results along the RT process at $b = 3000s \cdot mm^{-2}$ for the aQBI model. The method was performed with $N = 6$ and $\lambda = 0.006$ (regularization).

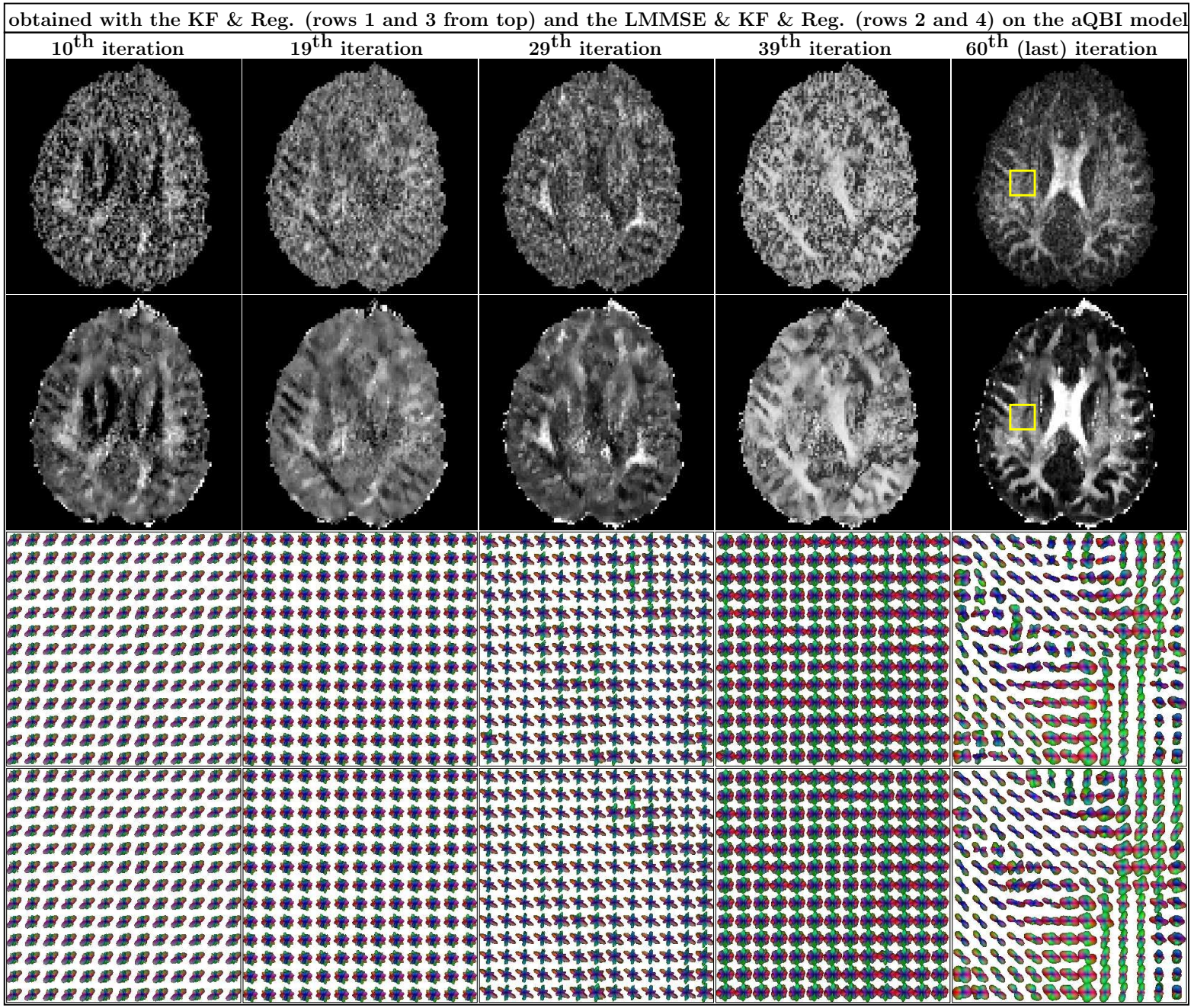


Figure 7.16: LMMSE & KF results along the RT process at $b = 3000s \cdot mm^{-2}$ for the aQBI model. The method was performed with $N = 8$ and $\lambda = 0.006$ (regularization).

Computational efficiency.

We also measured the duration of our RT noise correction method (the global version only). To meet the RT constraints, this denoising processing for one DW encoded volume has to be entirely performed before the acquisition of a new diffusion sensitized volume. It is consequently required that the denoising processing duration does not exceed the repetition time T_R of the dMRI sequence. We measured different durations involved in the RT denoising process: t_{LMMSE} is the time required to compute the LMMSE at one iteration i. This time does not include the feedback loop calculation. t_{KF} is the time accounting for the KF computation, as well as the feedback loop. Finally, t_{map} and $t_{display}$ are, as in chapter 4, the time for the GFA or dODF map reconstructions and the time to display these maps, respectively. As in chapter 4, concerning the dODF case run on the cluster, an additional transfer time is accounted for. Our time measurements are shown in fig. 7.17.

Thanks to the parallelization and distribution on a cluster (80 CPUs), the whole processing duration per volume decreased far below the T_R. Because of the communication time required to transfer the data from the master CPU to the nodes and to collect the results, the processing time after the parallelization and distribution is slightly higher than the time spent on only one CPU divided by 80. In fig. 7.17, we can notice that the LMMSE computation time (without the feedback loop) is much lower than others. It is again visible that the choice of the LMMSE algorithm permitted to perform a RT noise correction for DW data. When we performed the correction online during an exam with a subject, we could also verify that these very short durations easily allowed to achieve the correction methods before the acquisition of the next DW volume.

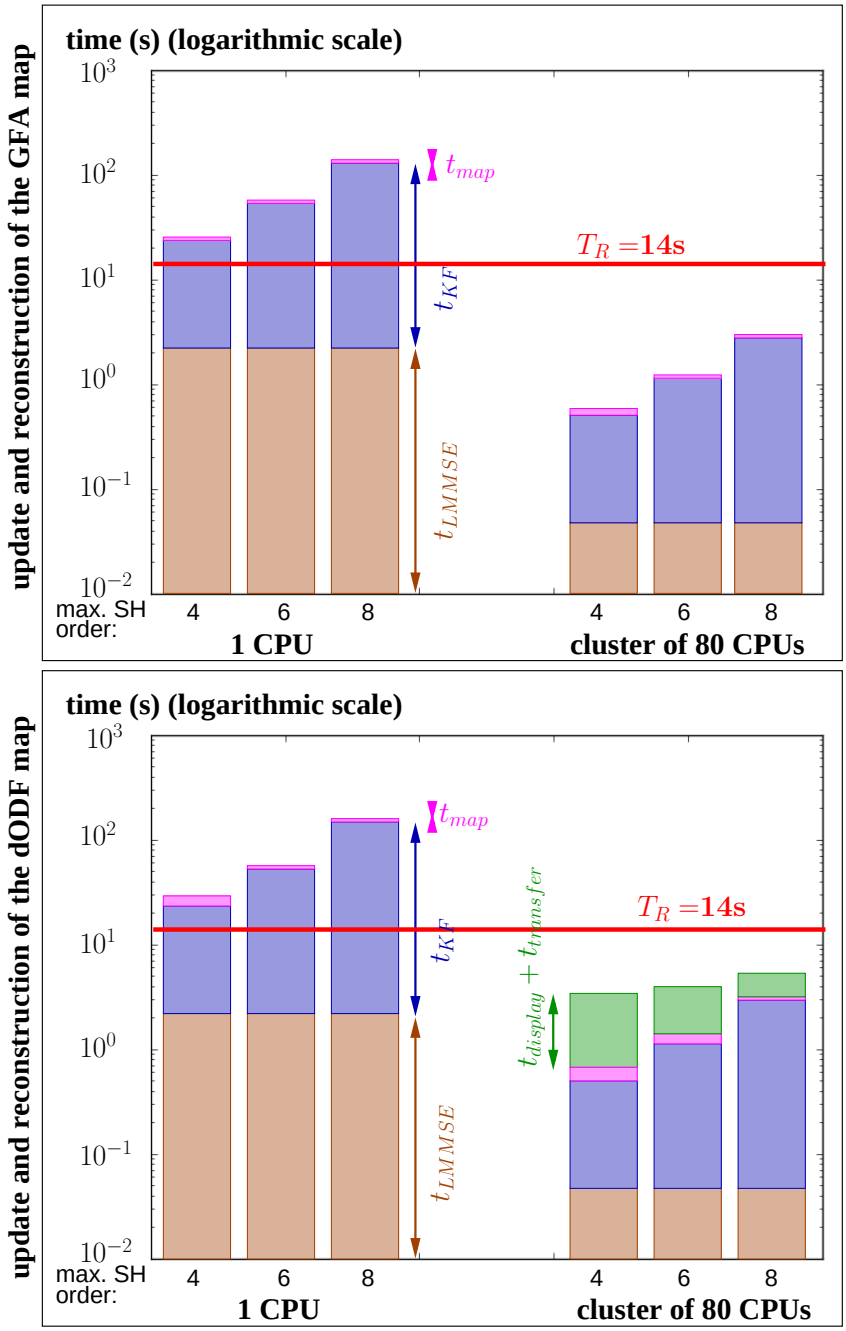


Figure 7.17: Times to get the GFA map and dODF map denoising update after a new iteration: t_{LMMSE} (brown box) is the LMMSE computation time for a neighborhood size of 5×5×5 voxels. t_{KF} (blue box) is the time required for one iteration of the KF algorithm and t_{map} (magenta box) is the time for processing the GFA map for the aQBI model (top) and the dODF for the aQBI model (bottom). Finally, $t_{display}$ (green box) is the time required to display the maps. It additionally contains the time required for the transfer of the data from the CPUs of the cluster to the server, when using the cluster. $t_{display}$ is not shown when it is very negligible compared to the other times. These times are compared to $T_R = 14s$. Pay attention to the logarithmic scale giving the feeling that t_{LMMSE} is higher than other times, which is actually the opposite.

7.1.5 Conclusion

This first RT noise correction method satisfied us for its quality and time performances. It is dependent on the choice of the parameters α and β. In the method, α was chosen with no dependence on the data. However, it was not the case for β, which was tuned using our simulated data. We want to point out here that our simulated DW data were generated using parameters of the diffusion process and of the fiber geometry which were very close to the real ones. Therefore, we can rely on the values we found for our configurations at the different *b-values* and for the σ chosen. But, it must not be forgotten that if a data is obtained with a completely different noise level, the β parameter may not be the optimum one and has to be tuned again.

Another possible flaw of the method concerns its theory. After having applied the LMMSE to the measured DW signal, there is a residual noise, called ϵ_r. This noise appears in the linear eq. 7.1 required for the KF. From the LMMSE &KF perspective, this noise is supposed to be zero-mean Gaussian, as required in the KF model detailed in chapter 4. However, this assumption was not verified. [Aja-Fernández et al. (2008b)] studied the distribution of the LMMSE filter output and empirically showed that it could be considered as a Rician distribution in the case of $n = 1$. Therefore, the residual noise may not follow a zero-mean Gaussian distribution.

The next section proposes another way to perform RT noise removal which deals with a simpler noise distribution.

7.2 Gaussian noise filtering in the complex Fourier space

7.2.1 Presentation of this reconstructor-integrated RT method

We propose a noise correction method that we integrated in the Siemens reconstruction system. This method is defined as a functor inserted before the SoS reconstruction in the Siemens reconstruction pipeline shown in fig. 7.18. Therefore, the advantage of this technique is that it directly corrects the **k**-space data acquired on each channel before their combination and consequently the noise to consider in these data is assumed to be **uncorrelated zero-mean Gaussian** distributed. Such a noise is much simpler to correct than a nc-χ noise. Furthermore, with this technique incorporated directly in the Siemens reconstruction pipeline, the denoising task is performed close to the acquisition step and before any partially parallel MRI reconstruction technique, thus yielding a **generic** correction method: whatever the parallel algorithm chosen, the correction is independent on this choice. We can call this algorithm an *online pre-processing* technique of denoising.

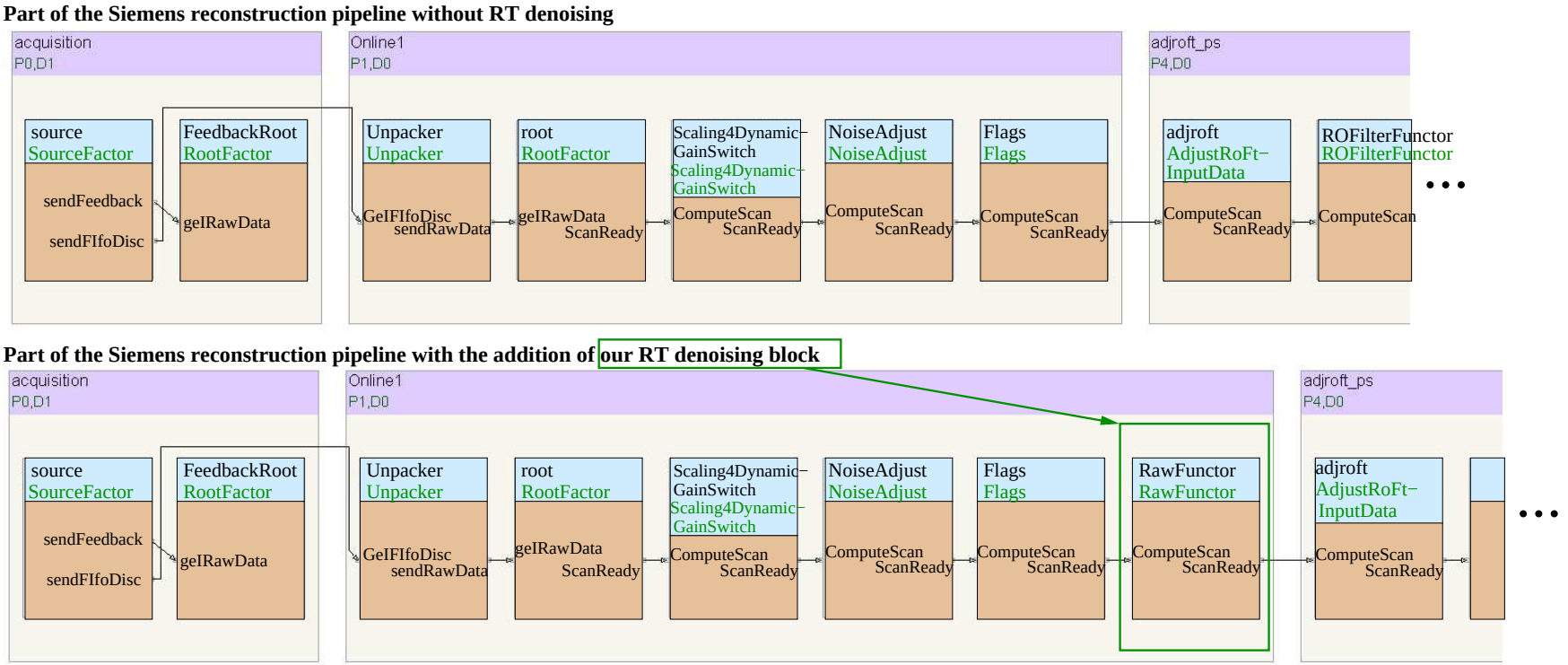


Figure 7.18: Location of the RT correction block in the Siemens reconstruction pipeline.

However, one **challenge** of such a technique is **to access** to the reconstruction pipeline of the manufacturer and to develop in his environment. This was doable with our Siemens Tim Trio system, and we implemented our denoising algorithm in the ICE programming environment. The code corresponds to a **functor** inserted in the pipeline shown in fig. 7.18.

When dealing with noise correction in the **k** space, as the noise is contained in the high frequencies, the common idea is to discard these high frequencies. Several low-pass filters have been developed to this aim. The simplest low-pass filter is called the ideal low-pass filter and simply eliminates all frequencies above a *cutoff frequency*, which has to be defined. A more sophisticated filter, the Butterworth filter is a low-pass filter relying on the use of normalized Butterworth polynomials and proposes a smoother transition to partially eliminate the high frequencies. The Chebyshev filter, named after the Chebyshev polynomials used in this filter, gives a similar filter with this difference compared to the Butterworth filter that it can contain some undulations in its magnitude frequency response. Other filters like the Bessel filter (with no undulation) and the elliptic filter (with undulations) were also proposed. Their gain-magnitude frequency responses are compared in fig. 7.19. Some of these filters (like the ideal low-pass filter or the Butterworth filter) can generate reverberation or undulation effects in the structural edges of the image in the spatial domain.

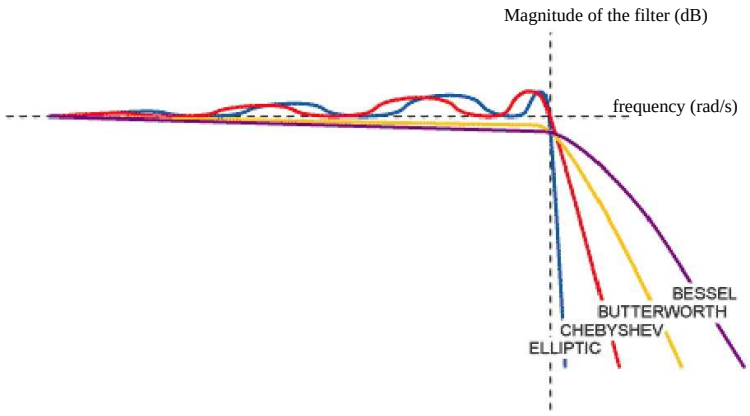


Figure 7.19: Gain-magnitude frequency responses of different low-pass filters. Fig. extracted from aune.lpl.univ-aix.fr/~fulltext/2758.pdf.

Another simple low-pass filter is the Gaussian filter, which avoids undulation effects in the borders of the structures contained in the image. This is due to the fact that the Fourier transform of a Gaussian PDF remains a Gaussian PDF, whereas for example the ideal filter becomes a sinc function after the inverse Fourier transform, leading to undulations effects in the image in the spatial domain. For our RT noise correction, we focused on such a Gaussian low-pass filter. Its shape is given by a 3D Gaussian PDF (fig. 7.20).

The Gaussian filter is based on the following equation adapted to our 1D filtering at

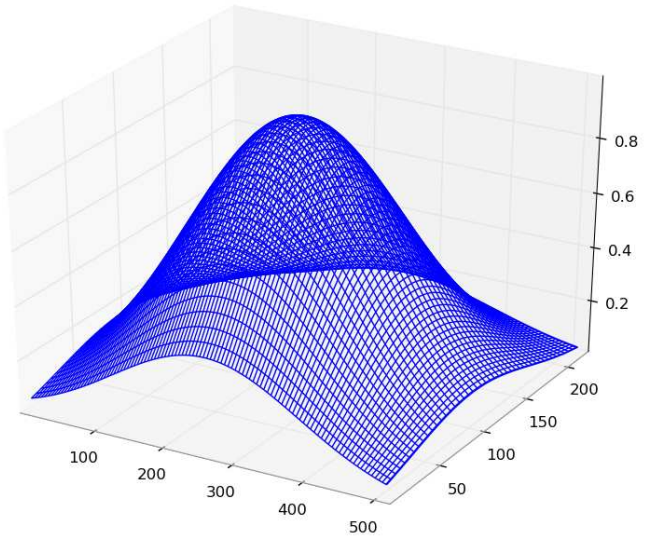


Figure 7.20: Shape of 3D Gaussian PDF with $\sigma_{Gauss} = 0.015$.

the position (k_x, k_y) on the **k**-space free induction decay (FID) line, considering the raw real or imaginary signal m_{k_x,k_y}:

$$\boxed{\hat{s}_{k_x,k_y} = m_{k_x,k_y} \times G(k_x, k_y, \sigma_{Gauss})}, \tag{7.7}$$

where $\hat{s}_{k_x,k_y}$ is the noise-free estimated signal and $G(k_x, k_y, \sigma_{Gauss}) = e^{-\frac{(k_x-o_x)^2+(k_y-o_y)^2}{2}\sigma^2_{Gauss}}$ is a 3D Gaussian PDF of standard deviation σ_{Gauss} in the **k**-space centered at (o_x, o_y). The parameter σ_{Gauss} has to be tuned to get the optimum noise removal, with a limited smoothing effect. This low-pass filter is applied on each FID line (for each real and imaginary received lines at each channel).

We show in the next subsection the results obtained on real data, which are different from the previous data from the "Archi database", as we did not have their corresponding raw data anymore.

7.2.2 Results on real data & discussion

Presentation of the real DW data used here.

The real DW data, which we used here, were collected on a human brain using a Magnetom Tim Trio 3T MRI system (Siemens Medical Solutions, Erlangen, Germany), employing a spherical direction sampling of 60 uniformly distributed over a shell at $b = 1400s \cdot mm^{-2}$. A T_2-weighted volume was acquired at $b = 0s \cdot mm^{-2}$. The acquisition parameters were as follows: $T_E/T_R = 92ms/12s$, field of view FOV= $256mm$, matrix 128×128, 60 slices, resolution $2{\times}2{\times}2mm^3$, GRAPPA factor of 2. The acquisition was performed with a 12-element head coil available on the Tim Trio, for which the 12 coil elements are combined

into 4 groups of 3 coil elements. These groups are received through 4 distinct receiver channels, yielding $n = 4$.

We applied the Gaussian low-pass (L-P) filter with the functor on the T_2-weighted data (as they were not averaged) and the DW data with three values for σ_{Gauss}: $\sigma_{Gauss} = 0.01/0.02/0.03$. We also applied the nc-χ LMMSE & KF method with the same parameters used in subsection 7.1.4 for the DW volume at $b = 1500s \cdot mm^{-2}$. These parameters are defined as follows: $n_{eff} = 1.8$, $\alpha = 2$ and $\beta = 0.10$. Concerning the correction by this filter on the T_2-weighted data, we only used a spatial weighting in the feedback loop with $\alpha = 2$, without any structural weighting. Fig. 7.21 shows the results obtained with each method on a DW image, a GFA and an RGB map.

From fig. 7.21, we can first see that the Gaussian low-pass filter rightly removes noise with $\sigma_{Gauss} = 0.02/0.03$. For $\sigma_{Gauss} = 0.01$, the effects of the filter are very small. At $\sigma_{Gauss} = 0.02$, a certain smoothing effect is visible on all images. This effect increases at $\sigma_{Gauss} = 0.03$ and has a tendency to appear isotropically in the image, independently of the structural features contained in the image. The smoothing effect is nevertheless still very low. When we compare the images obtained with the Gaussian low-pass filter with the images obtained with the LMMSE & KF, it appears that the LMMSE & KF has a higher smoothing effect in homogeneous regions which is however better controlled in the borders between different structural regions thanks to the spatial and structural feedback loop. Anyway, the maps present an efficient noise removal, with a slightly better contrast and a slightly more accurate definition of the anatomical structures than with the Gaussian low-pass filter.

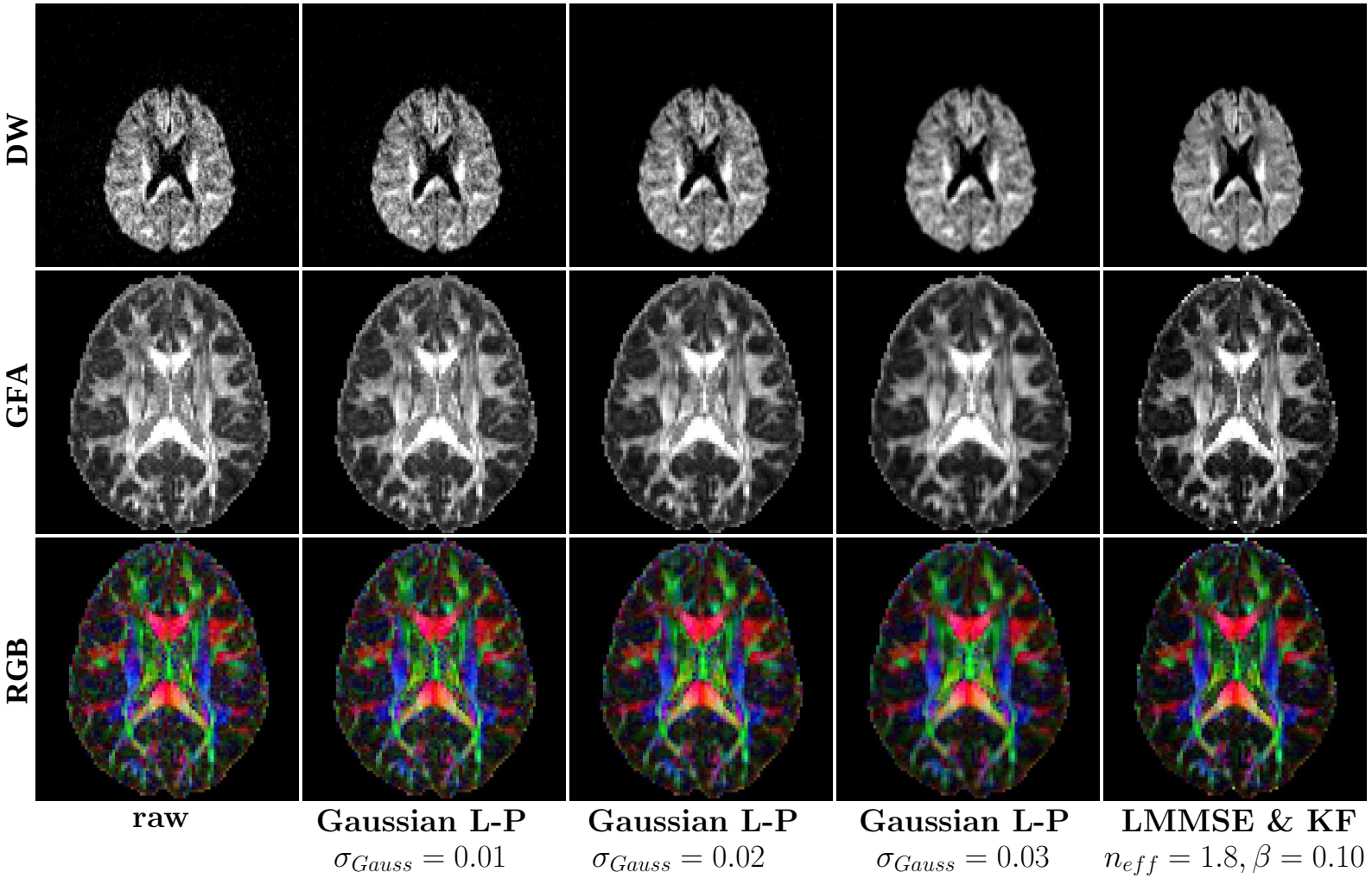


Figure 7.21: Comparison between the Gaussian low-pass (L-P) filter, applied with $\sigma_{Gauss} = 0.01/0.02/0.03$ and the LMMSE & KF method applied with $n_{eff} = 1.8$, $\alpha = 2$ and $\beta = 0.10$ on DW data acquired at $b = 1400s \cdot mm^{-2}$.

These qualitative results are confirmed in fig. 7.22, showing the comparison between the GFA ratios, calculated as indicated in the figure, on the raw GFA and the GFAs obtained with all four methods. The GFA ratio is increased after each method, with the highest ratio obtained for the LMMSE & KF technique.

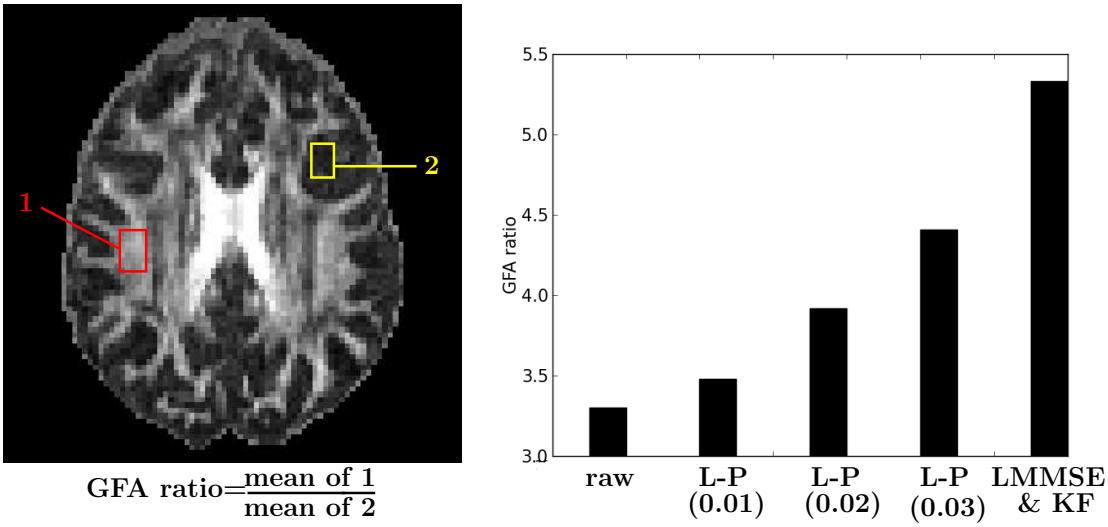


Figure 7.22: Comparison between the GFA ratios of the raw map and the maps obtained after application of the Gaussian low-pass (L-P) filter with $\sigma_{Gauss} = 0.01/0.02/0.03$ (number indicated in brackets), and the map obtained after application of the LMMSE & KF with $n_{eff} = 1.8$, $\alpha = 2$ and $\beta = 0.10$. The GFA ratio computation was performed for a data acquired at $b = 1400 s \cdot mm^{-2}$.

Finally, we studied the computational efficiency of the Gaussian low-pass filter integrated in the Siemens reconstruction pipeline. We could measure the global time to reconstruct a DW volume at one iteration. This time therefore accounts for the entire Siemens reconstruction pipeline (including the Gaussian low-pass filter functor), but does not account for the dMRI maps' generation. This time is equal to 0.534s per iteration. When performing the same measurement on the raw pipeline (not including the Gaussian low-pass filter functor), we could not measure a noticeable duration difference. Consequently, the functor is executed in a very short time (less than 10 ms). Finally, this pipeline integrated method is definitely a RT technique, as its computation time is far below the repetition time T_R of 14s.

7.2.3 Conclusion

To conclude on this RT noise correction, we can say that it is a theoretical simple method. Its difficulty lies in the fact that it is not straightforward to change the Siemens reconstruction pipeline. Furthermore, this was possible in our case, but it may not be always the case, depending on the manufacturer. Concerning the results of this technique, they were shown to improve the quality of the maps. We want to highlight here that our evaluation was performed at a quite low *b-value*, and therefore does not represent high *b-values*'

cases with a lower SNR. This method present the risk of blurring the images. To avoid such an effect, the σ_{Gauss} parameter has to be carefully chosen. Finally, compared to the LMMSE & KF technique, the Gaussian low-pass filter did not achieve a similar contrast. A perspective regarding denoising methods applied on the **k**-space data would be to test other low-pass filter and compare their performances. Finally, we want to highlight that such methods cannot incorporate a structural feedback loop as proposed in the LMMSE & KF technique. Indeed, as they work in the **k**-space, a structural information measured in the spatial domain is not easily integrated.

7.3 Correction relying on a parallel Kalman filter (PKF)

After having developed the "LMMSE & KF" incremental denoising technique, we wanted to propose a more rigorous RT noise correction technique. For that, we looked more deeply in the literature about incremental solvers accounting for a non-Gaussian noise. Concerning incremental solvers like the KF, methods were developed to create a similar framework enabling filtering of non-Gaussian noise. We made a quick review of such techniques in chapter 4, subsection 4.2.1 (page 63). Our intention was here to build a simple incremental framework (close to the KF simplicity) and incorporate in this framework the non-Gaussian distribution information about noise. For that, we eliminated computationally demanding methods like point-mass approaches or particle filters. We also eliminated the extended KF (EKF), as well as the unscented KF, which are not always reliable methods (especially the EKF). We preferred Gaussian mixture approximations.

Among this latter group of techniques, some methods like [Sorenson and Alspach (1971); Alspach and Sorenson (1972); Masreliez (1975)] either present increasing computational complexity in time or cannot operate in a changing noise environment ([Plataniotis et al. (1997)]) and thus would not be convenient for an RT nc-χ noise correction. Nevertheless, one technique, namely the **parallel Kalman Filter** (PKF) of [Plataniotis et al. (1997)], shows computational efficiency without the necessity of similar assumptions. It was demonstrated to be efficient in the field of narrowband interference suppression. The work presented here is based on the PKF. It is the first adaptation of the PKF to this particular RT nc-χ noise correction issue for rtdMRI. It also constitutes one contribution of this thesis work [Brion et al. (2012b,a,c)].

7.3.1 Presentation of this RT method based on a PKF

In this subsection, we first give insights to understand on which principles the PKF relies. Then, we present this RT noise correction method for the aQBI, DTI and sa-aQBI linear models. The presentation is detailed for the aQBI model.

The principle of the PKF.

In the previous "LMMSE & KF" method, the entire noise correction step is performed by the bayesian LMMSE filter, which contains the knowledge on the MRI nc-χ noise (fig. 7.23). The KF is used as a recursive least squares technique to iteratively obtain the coefficients of the decomposition of the normalized DW signal on the modified SH basis. The KF therefore does not contain any noise removal process and is only used for its incremental framework.

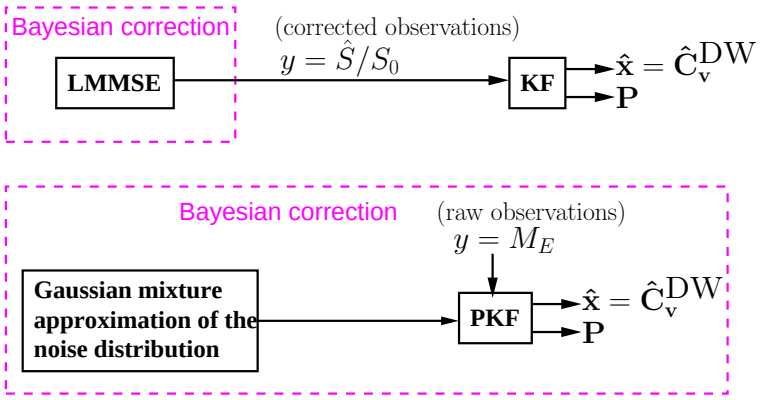


Figure 7.23: Comparison between the "LMMSE & KF" and the PKF principles.

The method, which we propose here, first contains a Gaussian mixture approximation of the noise distribution. Second, this mixture is injected in a PKF to propose an incremental noise-free signal estimation. This new method is very different from the previous "LMMSE & KF" method, because it deals with a Bayesian noise correction from its beginning to its end (fig. 7.23). Whereas in the previous technique, the performance of the algorithm was entirely dependent on the LMMSE, here, there are several required steps dedicated to noise correction.

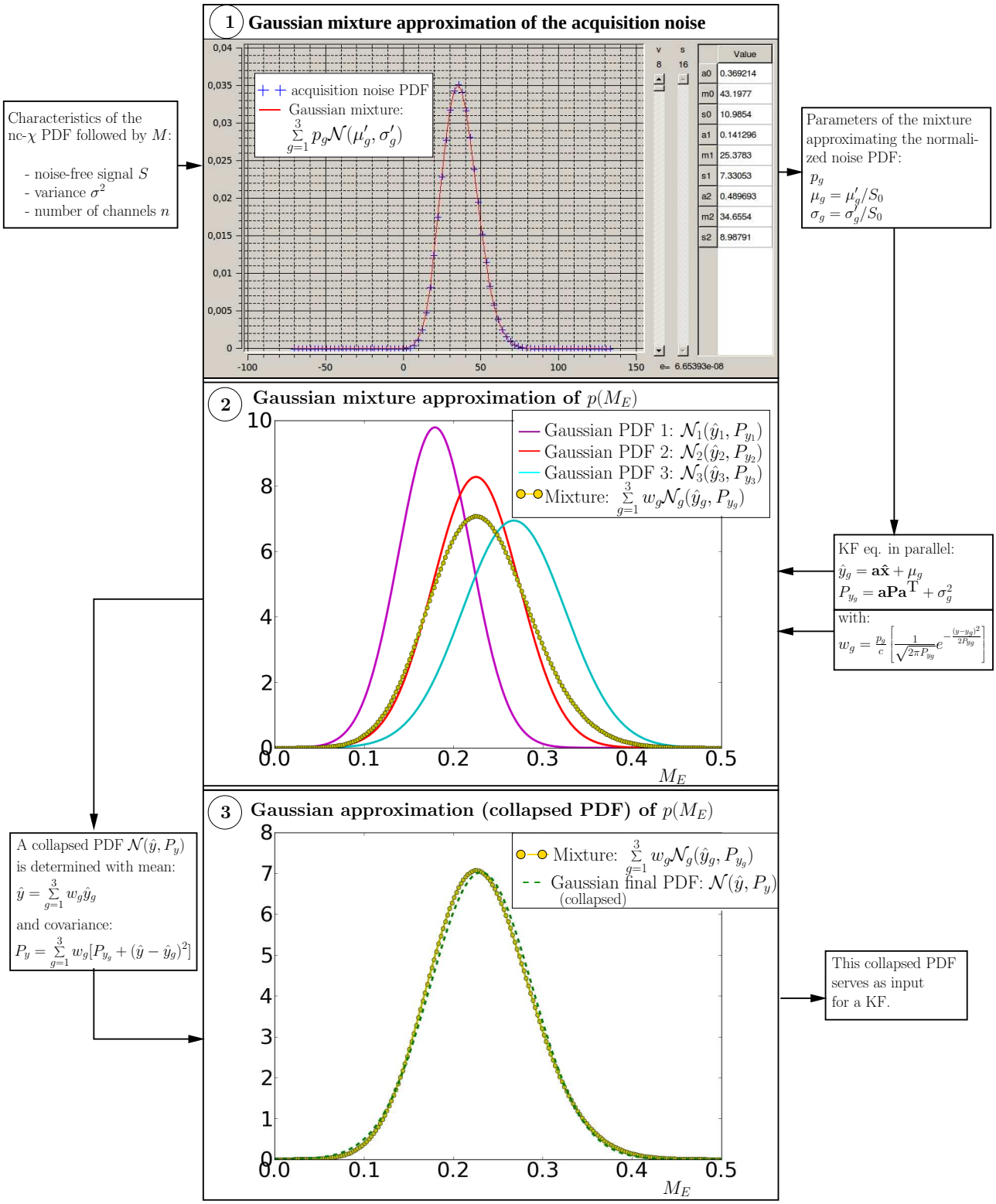


Figure 7.24: Principle of the PKF.

Fig. 7.24 summarizes in three steps the principle of the PKF. First, using the characteristics of the nc-χ PDF followed by the noisy DW signal M, the acquisition noise PDF

is approximated by a Gaussian mixture (first graph in fig. 7.24). From the latter, the parameters of the Gaussian mixture approximating the noise normalized by the noise-free T_2-weighted signal are determined. These parameters are the weight p_g, the mean μ_g and the standard deviation σ_g of each Gaussian PDF of the mixture. Then, KFs' equations give the expression of the mean $\hat{y}_g$ and the covariance P_{y_g} of each new Gaussian PDF used in a mixture to approximate the PDF of the observations $p(M_E)$ (second graph in fig. 7.24). We also call this mixture as the "pre-collapsed PDF". Finally, a collapsed PDF is calculated. It is a Gaussian PDF, which approximates the previous mixture. As this collapsed PDF is a Gaussian distribution, it can be used as input for a simple KF.

The reason of this final collapse to obtain a Gaussian PDF approximating $p(M_E)$, is to be able to perform a final KF to determine the state vector $\mathbf{x}$ and the error covariance $\mathbf{P}$. With this collapsing process, we see that the PKF has a computational complexity that is nearly equivalent to the one of a simple KF. There is no increasing complexity in time with the PKF, in opposition to other Gaussian mixture filters. Although the Gaussian mixture is reduced to a Gaussian PDF at the end of the process, before the injection into the KF, the PKF does not reduce to a KF adapted to a Gaussian noise with a non-zero-mean. As we will see later, the final collapsed density is not equal to the Gaussian PDF we would get with a simple KF.

Detailed presentation adapted to the aQBI linear model.

Fig. 7.25 synthesizes the whole algorithm based on a PKF and designed to obtain from the measured DW signals M the resulting noise corrected coefficients $\hat{\mathbf{C}}^{\mathrm{DW}}$. These coefficients define the noise corrected normalized signal $\hat{S}/S_0$ in the modified SH basis introduced in subsection 3.2.4. The RT method is performed for each voxel $\mathbf{v}$ of the input DW volume and is repeated at each new diffusion orientation $\mathbf{o}_i$ during the acquisition. In fig. 7.25, the blocks in orange show the different steps of the RT method. In our case, the PKF purpose is to **simplify the nc-χ noise correction problem to a Gaussian one**, using the idea that the distribution of a signal corrupted by any non-Gaussian noise can be approximated by a Gaussian mixture [Plataniotis et al. (1997)]. The global method can be resumed as follows:

1. The **LMMSE** estimates the noise-free DW magnitude $\hat{S}(\mathbf{v}, \mathbf{o}_i)$, which, with the noise standard deviation σ completely defines the noise PDF.

2. Then, a **fit** approximates this noise distribution by a Gaussian mixture.

3. Finally, the **PKF** [Plataniotis et al. (1997)] accounts for each Gaussian noise through linear systems working in parallel. It joins the results in a resulting collapsed density with only one Gaussian term, and then, as a classical KF, gives the optimal coefficients $\hat{\mathbf{C}}^{\mathrm{DW}}$, corrected from noise.

4. Also, a **feedback loop** is added, as in the previous RT method presented in section 7.1, to calculate a weight $w(\mathbf{v}, \mathbf{v}')$ using the results $\hat{\mathbf{C}}^{\mathrm{DW}}$. The weight is then injected

in the LMMSE to limit smoothing effects.

To resume, the LMMSE and the fit enable the noise estimation and the PKF realizes the main correction step using the measured magnitude as an input. All algorithms involved in this PKF-based method are generic and require only two conditions to be used:

- the LMMSE depends on the knowledge of the two moments of the PDF of the noisy measurement,
- the state estimation problem must be linear to be solved by the PKF.

Such a method could be applied to any kind of noise as long as these two conditions are respected.

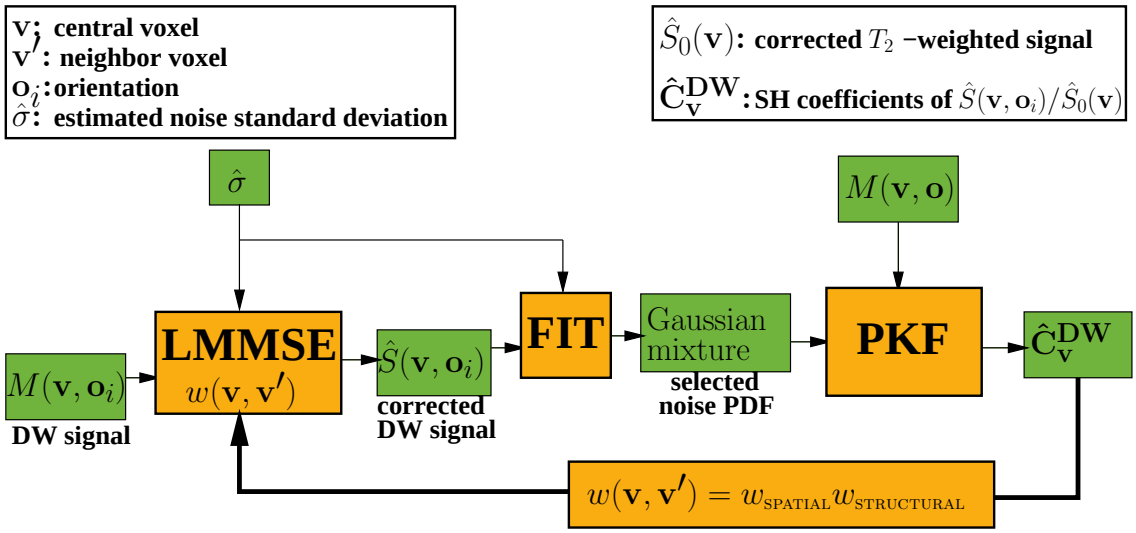


Figure 7.25: The RT noise correction method based on a PKF in a global view. The blocks in orange show the different steps of the RT method.

Fig. 7.25 synthesizes the global method which can be summarized as follows: first, the LMMSE is applied to the measured DW signal $M(\mathbf{v}, \mathbf{o}_i)$, at position $\mathbf{v} \in R^3$ for the orientation $\mathbf{o}_i \in S^2$, in order to give an estimate of the noise-free signal $\hat{S}(\mathbf{v}, \mathbf{o}_i)$. Also, the noise standard deviation is estimated either globally or voxel-wise for the DW volume. For the first case, we performed the empirical tuning, as explained in subsection 6.2.4, of chapter 6, page 151. For the second case, we relied on the method by [Aja-Fernández et al. (2013)] detailed in table 5.1, page 126. Both estimations $\hat{\sigma}$ and $\hat{S}(\mathbf{v}, \mathbf{o}_i)$, with the measured magnitude $M(\mathbf{v}, \mathbf{o}_i)$ completely define the nc-χ distribution of $M(\mathbf{v}, \mathbf{o}_i)$. From this distribution, we can also characterize the distribution of the noise ϵ which corrupts $M(\mathbf{v}, \mathbf{o}_i)$. Indeed, we saw in chapter 5, how the noise PDF is written. It is given by eq. 5.18, page 114. This allows, in a second stage, to approximate the noise distribution by a Gaussian mixture using a Levenberg-Marquardt non-linear fit. It corresponds to the FIT block in the fig. 7.25.

We have to pay attention that this Gaussian mixture approximates the acquisition noise on the DW signal M. However, the linear system on which relies the PKF is the following one:

$$\frac{M(\mathbf{v},\mathbf{o}_i)}{S_0(\mathbf{v})} = \mathbf{b_i} \cdot \mathbf{C}_{\mathbf{v}}^{\mathrm{DW}} + \epsilon, \tag{7.8}$$

with $\mathbf{b_i}$ the i^{th} row of $\mathbf{B}$, the matrix of the modified SH basis defined by eq. 3.13, page 46. The index i is the iteration index of the RT process. The noise term ϵ corresponds to the acquisition noise normalized by the noise-free T_2-weighted signal. Consequently, the Gaussian mixture has to be adapted to this normalized noise. To this aim, we used the **Theorem 1.**, page 112, and deduced, from the original Gaussian mixture $\sum_{g=1}^{f} p_g \mathcal{N}(\mu'_g, \sigma'_g)$ approximating the acquisition noise, the following Gaussian mixture approximating the normalized noise ϵ: $\sum_{g=1}^{f} p_g \mathcal{N}(\mu_g, \sigma_g)$ with $\mu_g = \mu'_g/\hat{S}_0$ and $\sigma_g = \sigma'_g/\hat{S}_0$. This latter mixture of f Gaussian distributions is injected in a PKF, which calculates, in RT, the coefficients characterizing the noise-corrected signal in the modified SH basis.

The PKF details are given in fig. 7.26, which shows the refinement, also called update, after the acquisition of the i^{th} diffusion orientation, of the previously corrected results of the $(i-1)^{th}$ diffusion orientation. The job of the PKF is to account for each Gaussian noise through linear systems working in parallel, using the results of the approximation done by the Levenberg-Marquardt algorithm and using the vector $\mathbf{a}_i = [A_{i1}, ..., A_{iR}]$ corresponding to the i^{th} row of the $M \times R$ matrix $\mathbf{B}$ of the modified SH basis, defined by eq. 3.13, page 46. These linear systems (fig. 7.26), calculated in the PKF, are then gathered in a single density being a Bayesian *a posteriori* estimate of the Gaussian mixture. This collapsed density has only one Gaussian term and can be injected in a Kalman-like filter (fig. 7.26). The Kalman-like filter equations give the expressions of the innovation ν_i, the Kalman gain $\mathbf{k}_i$, the normalized covariance matrix $\mathbf{P}_i$ of the state vector $\mathbf{x}$ and the updated estimate of the state vector $\hat{\mathbf{x}}_i$. The Kalman-like filter delivers the updated estimates of the state vector, corresponding to the optimal corrected coefficients' vector $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$. From it, the corrected DW signals can be obtained, as well as any map of the aQBI model. More information about the derivation of the PKF can be found in appendix 7.6 at the end of this chapter.

As for the first RT noise correction method relying on the LMMSE followed by a KF, a feedback loop is used to limit the smoothing effect induced by the LMMSE. This loop defines weighting coefficients that are composed of a spatial and a structural term that are defined for the aQBI model as:

$$w_{spatial}(\mathbf{v},\mathbf{v}') = exp\left(-\frac{(\mathbf{v}'-\mathbf{v})^2}{2\alpha^2}\right), \tag{7.9}$$

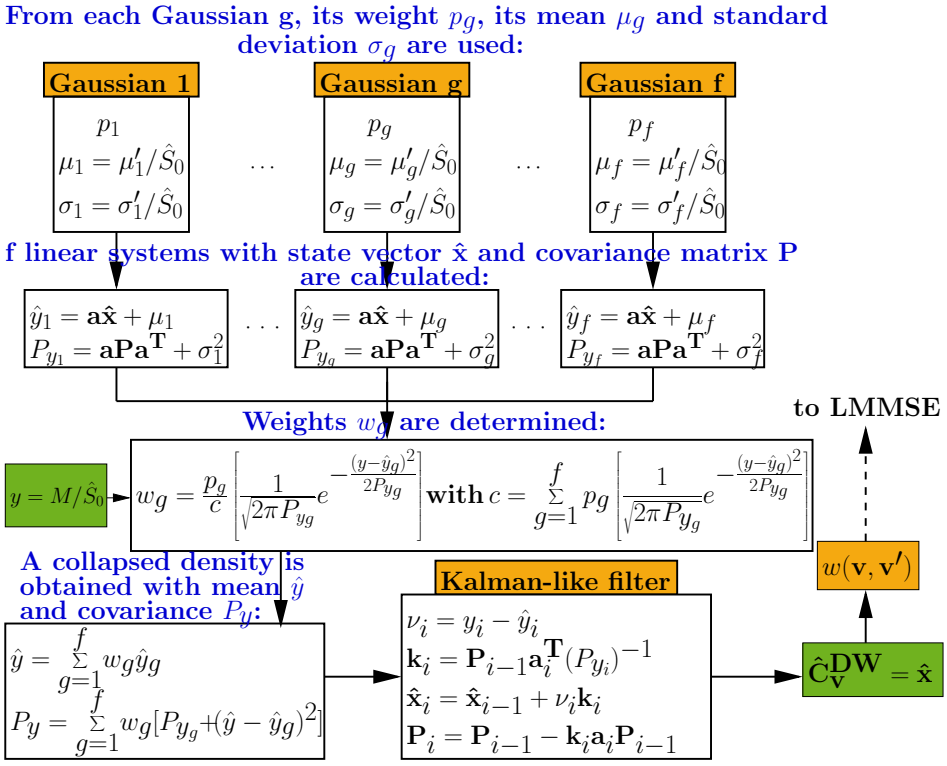


Figure 7.26: The detailed PKF adapted from [Plataniotis et al. (1997)].

and

$$w_{structural}(\mathbf{v}, \mathbf{v}') = exp\left(-\frac{\sum_{j=1}^{N}\left(\hat{\mathbf{C}}_{\mathbf{v}'}^{\mathrm{DW}}(j) - \hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}(j)\right)^2}{2\beta^2}\right), \tag{7.10}$$

where $\hat{\mathbf{C}}_\mathbf{v}^{\mathrm{DW}}$ is the current value of the PKF state vector at voxel $\mathbf{v}$ (N being the length of the state vector), and α and β are two constants to be tuned.

The PKF was initialized by setting the initial guess $\hat{\mathbf{x}}_0$ to the null vector. The initial covariance matrix was set to $\mathbf{P}_0 = (c\mathbf{I} + \lambda\mathbf{L})^{-1}$ with $c = 10^{-6}$, as suggested by [Deriche et al. (2009)] for the KF, to optimize its convergence. $\mathbf{I}$ represents the identity matrix, λ is a regularization factor, and $\mathbf{L}$ is the Laplace-Beltrami operator. This operator regularizes the solution and prevents remaining negative peaks to contribute to the final solution. The initial weight was defined such that: $w(\mathbf{v}, \mathbf{v}') = w_{spatial}(\mathbf{v}, \mathbf{v}')$.

The RT method adapted to the DTI and sa-aQBI linear models.

Adaptation to the DTI linear model. To adapt this RT method to the DTI linear model, the linear equation on which the PKF relies has to be replaced by the following

linear equation of chapter 4, subsection 4.3.2:

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}, \tag{7.11}$$

with, for iteration i, $y_i = ln\left(\frac{S_0}{M(\mathbf{o}_i)}\right)$ and $\mathbf{a}_i = b_i[o_x^2, 2o_xo_y, 2o_xo_z, o_y^2, 2o_yo_z, o_z^2]$, b_i being the *b-value*. The state vector $\mathbf{x}$ corresponds to the six LMMSE estimated coefficients of the symmetric diffusion tensor: $\mathbf{x} = [\hat{D}_{xx}, \hat{D}_{xy}, \hat{D}_{xz}, \hat{D}_{yy}, \hat{D}_{yz}, \hat{D}_{zz}]^T$. The noise is expressed, for iteration i, such that: $\epsilon_i = -ln\left(1 + \frac{\mu_i}{S_0} e^{b_i \mathbf{o}_i^T \hat{\mathbf{D}} \mathbf{o}_i}\right)$, with μ_i the acquisition noise. Therefore, the Gaussian mixture approximating the acquisition noise μ_i has again to be adapted to this different noise signal ϵ_i. Finally, the feedback loop also requires an adaptation to the DTI linear model, such that:

$$w_{STRUCTURAL}(\mathbf{v}, \mathbf{v}') = \exp\left(\frac{-\sum_{j=1}^{6}\left(\hat{\mathbf{D}}_{\mathbf{v}'}(j) - \hat{\mathbf{D}}_{\mathbf{v}}(j)\right)^2}{2\beta^2}\right). \tag{7.12}$$

Adaptation to the sa-aQBI linear model. The adaptation of the RT noise correction method to the sa-aQBI linear model requires also to modify the linear equation for the PKF, following again the details given in chapter 4, subsubsection 4.3.3

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}, \tag{7.13}$$

with, for iteration i, $y_i = ln\left(-lnM_E(\mathbf{o}_i)\right)/S_0$ and $\mathbf{a}_i$ corresponding to the i^{th} row of the matrix $\mathbf{B}$ of the modified SH basis. The state vector $\mathbf{x}$ corresponds to the LMMSE estimated coefficients' vector $\hat{\mathbf{C}}^{\mathrm{SA}}$. The noise is expressed, for iteration i, such that: $\epsilon_i = ln\left(-ln\mu_i\right)/S_0$, with μ_i the acquisition noise. Therefore, the Gaussian mixture approximating the acquisition noise μ_i has again to be adapted to this different noise signal ϵ_i. Finally, the feedback loop also requires an adaptation to the sa-aQBI linear model, such that:

$$w_{STRUCTURAL}(\mathbf{v}, \mathbf{v}') = \exp\left(\frac{-\sum_{j=1}^{N}\left(\hat{\mathbf{C}}^{\mathrm{SA}}_{\mathbf{v}'}(j) - \hat{\mathbf{C}}^{\mathrm{SA}}_{\mathbf{v}}(j)\right)^2}{2\beta^2}\right). \tag{7.14}$$

7.3.2 Evaluation of the FIT performances

In the PKF-based method, it is essential that the fit performs well. One parameter to adjust is the number f of Gaussian PDFs used to fit the PDF of ϵ. To this aim, we tested the PKF-based method on our simulated data at $b = 4500s \cdot mm^{-2}$ with a nc-χ noise with $\sigma = 16$ and $n = 4$. To restrict the evaluation to the FIT performances themselves, we injected the true noise-free signal S, instead of the LMMSE estimate in the FIT block represented in fig. 7.25. We can therefore call this algorithm: "FIT & PKF".

Additionally, we used the true noise standard deviation σ instead of the estimated one. Therefore, the correct nc-χ noise PDF could be reconstructed with its true parameters (S, σ). The Gaussian mixture approximation (fit) of the nc-χ noise PDF was performed under these conditions. Then, the PKF was run, using the fit outputs. We tested different values for the number f of Gaussian PDFs that we used to make the Gaussian mixture approximation of the noise distribution. We computed the MSE between the noise-free configuration and the "FIT & PKF" processed configuration. The MSE calculation was performed for each number f of Gaussian PDFs on the five slices containing the crossing of the two fiber bundles of the simulated volume. Therefore, we got an MSE measurement computed on many samples of voxels contained ($27\times31\times5\times500$) in this region of interest in our data. Our results are shown in fig. 7.27.

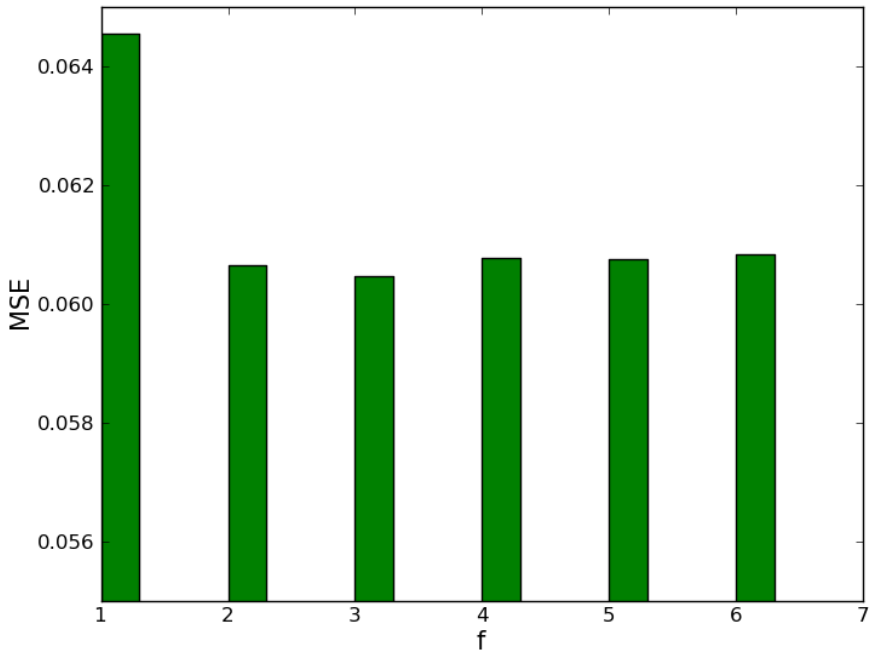


Figure 7.27: Evaluation of the optimum number of Gaussian PDFs to be used in the FIT. As a reference, the MSE calculated on the noisy regularized configuration is equal to 7.84×10^{-2}.

From this fig. 7.27, it is visible that the MSE decreases from the configuration with 1 Gaussian PDF used for the fit —then, the PKF is equivalent to a simple KF incorporating a Gaussian noise knowledge (with non-zero-mean)— to the other configurations. We remark that $f = 3$ is a local optimum. This led us to choose for the next experiments a Gaussian mixture approximation performed with 3 Gaussian PDFs.

Fig. 7.28 shows this approximation of the acquisition noise PDF (red curve) by a Gaussian mixture with $f = 3$ (blue points) for a noise-free signal $S = 8$ and a noise standard deviation $\sigma = 16$. The quadratic error between both curves is equal to 6.6×10^{-8}.

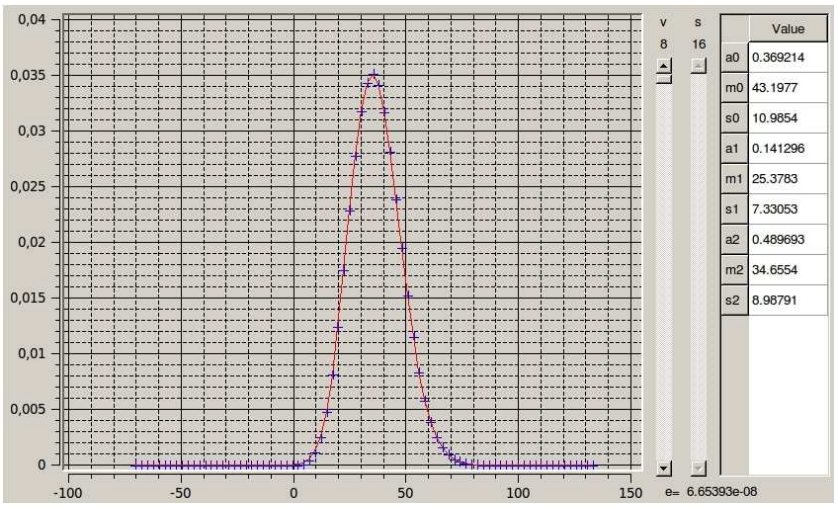


Figure 7.28: Fit curve result.

7.3.3 Evaluation of the PKF performances

In this subsection, we again injected the true noise-free signal S, instead of the LMMSE estimate in the FIT block represented in fig. 7.25. Therefore, we limited our experiments to the observation of the PKF algorithm performed with the true nc-χ distribution knowledge.

The goal of the PKF is to estimate the state vector $\mathbf{x}$ at each iteration i. To account for the non-Gaussian noise, it incorporates the Gaussian mixture approximating the noise PDF. From this Gaussian mixture, another Gaussian mixture is deduced in the PKF process which is used to estimate the PDF of the observation y, corresponding to the normalized DW signal M_E. This new mixture is called pre-collapsed PDF. The latter is used to obtain the collapsed PDF, which is a Gaussian PDF and serves as input in a KF. This collapsed distribution is the final approximation of the true $p(M_E)$.

Fig. 7.29 shows several PDFs at two iterations among the 500 iterations of the RT process performed on our simulated data. First, at $i = 50$, we can see that the PKF is at its beginning: the state vector $\mathbf{x}$ is estimated with a high error covariance $\mathbf{P}$ yielding large covariances for the curves involved in the approximation of the true $p(M_E)$. In opposition, at $i = 350$, the different curves have lower variances and better fit the true $p(M_E)$. In fig. 7.29, we also show the comparison between the PKF performed with mixtures using 3 Gaussian PDFs and a PKF performed with only one Gaussian PDF. In the latter case, the PKF is equivalent to a KF accounting for Gaussian noise with a non-zero mean. It is visible that this collapsed PDF resulting from the case $f = 1$ is different from the collapsed PDF obtained through Gaussian mixtures with 3 Gaussian PDFs ($f = 3$). Thefore, we can conclude that although there is a collapse process in the PKF algorithm which leads to account for a unique final Gaussian PDF to serve as input of a KF, this collapse does not make the PKF algorithm equivalent to the KF algorithm. Furthermore, we see in fig. 7.29 that the final collapsed density with $f = 3$ better approximates the pre-collapsed

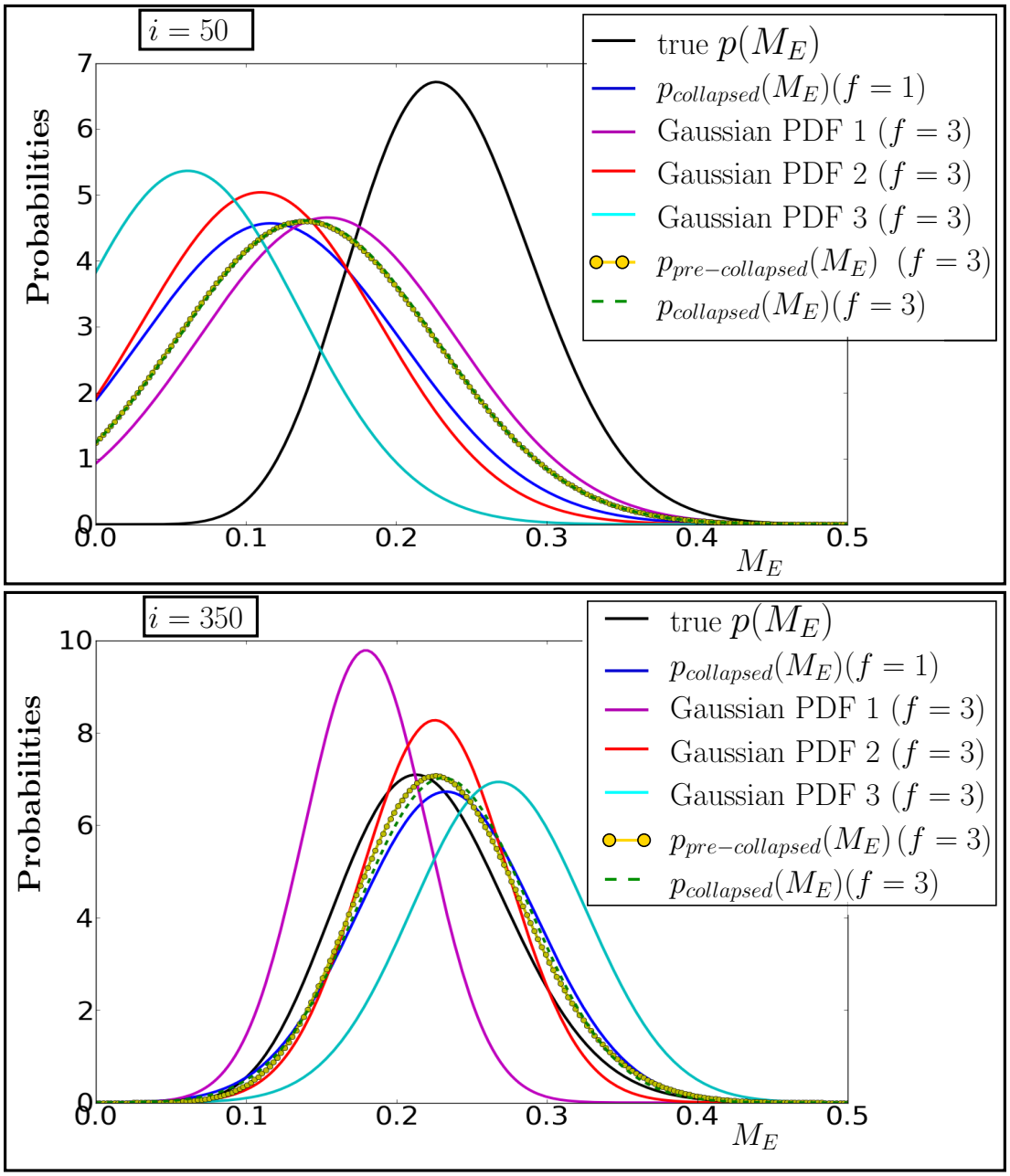


Figure 7.29: Evolution of the PDF curves involved in the PKF algorithm along the iterations.

density, as well as the true $p(M_E)$.

Fig. 7.30 shows the dODF results of this "FIT & PKF" process obtained on our simulated data. It is visible that the PKF-denoised maps retrieve an angular information closer to the noise-free reference than the noisy dODF map. Here, we do not remark any improvement in taking three Gaussian PDFs ($f = 3$) instead of one. This can be explain by the fact that the nc-χ PDF, we are dealing with, is very close to a Gaussian PDF, except for its skewness at low SNRs and low number of channels n. In our case, the PKF results are equivalent compared to those obtained with a simple KF accounting for the

non-zero-mean of a Gaussian noise approximating the nc-χ noise.

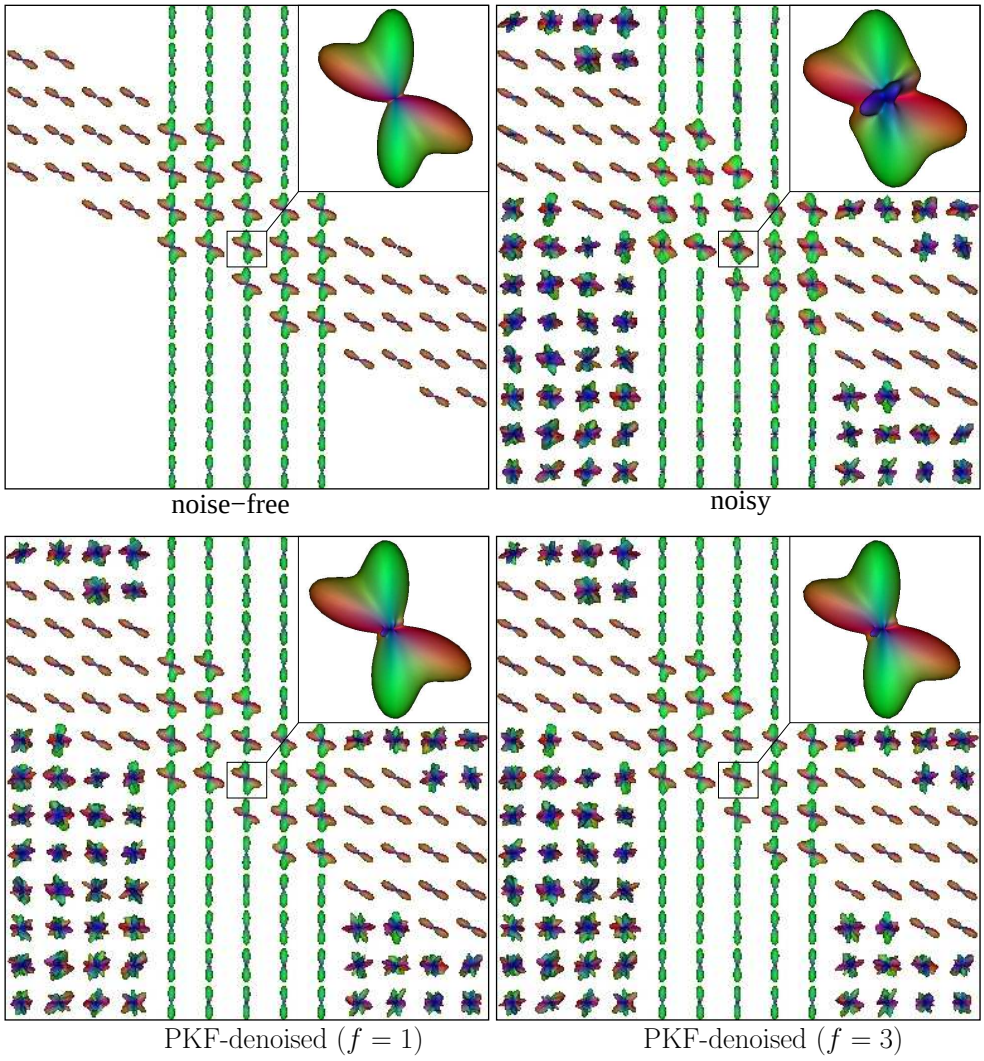


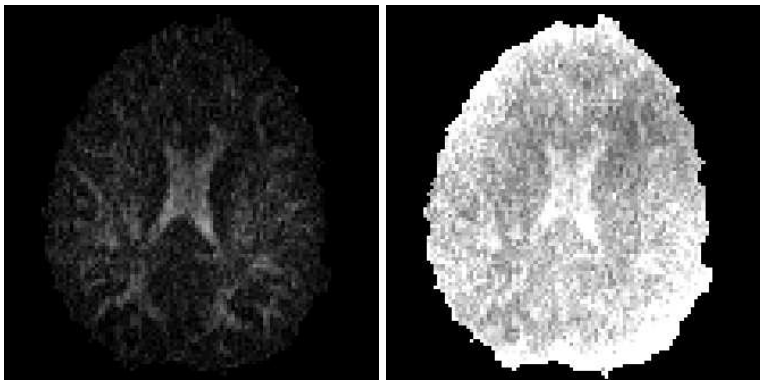
Figure 7.30: Comparison on dODF maps of the PKF performed with the true nc-χ PDF knowledge as input (the LMMSE is not used).

The rest of our results are shown with $f = 3$.

7.3.4 Improving robustness of the PKF

In this subsection, we show in fig. 7.31 the result obtained with the PKF on the real GFA map at $b = 4500s \cdot mm^{-2}$. This result surprised us, as it was not similar to what we had obtained previously on the simulated data.

Fig. 7.31 highlights that the GFA signal after application of the PKF global method is very much increased with a loss of some anatomical structures' perception in the image. Such a result led us to think that there is some mathematical indeterminacy in the PKF process. As a temporary solution, we proposed a solution that improves the robustness of the PKF-based method.



raw GFA GFA after PKF

Figure 7.31: Application of the PKF on real data. The raw GFA at $b = 4500s \cdot mm^{-2}$ is shown with the GFA after the application of the entire PKF-based method explained in subsection 7.3.1.

To make the PKF more robust, we propose to inject a spatial and structural mean of the raw magnitude M, instead of M as input for the PKF. Therefore the observations to inject in the PKF are no longer the normalized signals M_E, but they become: $y = \frac{\langle M \rangle}{S_0}$, with the mean $\langle M \rangle$ defined for the voxel $\mathbf{v}$ and the orientation $\mathbf{o}_i$, such that: $\langle M(\mathbf{v}, \mathbf{o}_i) \rangle = \dfrac{\sum\limits_{\mathbf{v}' \in \eta_\mathbf{v}} w_B(\mathbf{v}, \mathbf{v}') M(\mathbf{v}', \mathbf{o}_i)}{\sum\limits_{\mathbf{v}' \in \eta_\mathbf{v}} w_B(\mathbf{v}, \mathbf{v}')}$, with $w_B = w_{B_{spatial}} \times w_{B_{structural}}$. Both components of this weight are expressed such that:

$$w_{B_{spatial}}(\mathbf{v}, \mathbf{v}') = exp\left(-\frac{(\mathbf{v}' - \mathbf{v})^2}{2\alpha_B^2}\right), \tag{7.15}$$

and

$$w_{B_{structural}}(\mathbf{v}, \mathbf{v}') = exp\left(-\frac{\sum\limits_{j=1}^{N}\left(\hat{\mathbf{C}}_{\mathbf{v}'}^{\mathrm{DW}}(j) - \hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}(j)\right)^2}{2\beta_B^2}\right), \tag{7.16}$$

where $\hat{\mathbf{C}}_{\mathbf{v}}^{\mathrm{DW}}$ is the current value of the PKF state vector at voxel $\mathbf{v}$ (N being the length of the state vector), and α_B and β_B are two constants to be tuned. We write them with the index B, in opposition to the parameters used in the feedback loop injected in the LMMSE which we will write with the index A.

This mean to be injected as input in the PKF is not an ideal solution, as it will necessarily induce some smoothing effect. However, it will give the PKF the required neighborhood information to achieve more stable results. Furthermore, the parameters α_B and β_B have to be tuned to limit the smoothing effect and produce a edge-preserving result. We show in the next subsection how we tuned the four parameters α_A, β_A, α_B and β_B.

7.3.5 Tuning of four parameters α_A, β_A, α_B and β_B

Concerning the tuning of the two parameters α_A and α_B, we chose to set $\alpha_A = 2$ and $\alpha_B = 2$, for the same reasons as the ones explained in subsection 7.1.2.

Concerning the tuning of the two parameters β_A and β_B, we performed similarly as for the "LMMSE & KF" method. Here, we tested the global PKF method (with the LMMSE step) with the knowledge of the true σ for different values of the two parameters and calculated the corresponding MSE on the same region as previously for the "LMMSE & KF" method. Fig. 7.32 shows that the minimum MSE for our simulated data at $b = 4500s \cdot mm^{-2}$, with a nc-χ noise of parameters $\sigma = 16$ and $n = 4$, was found for $\beta_A = 0.11$ and $\beta_B = 0.17$.

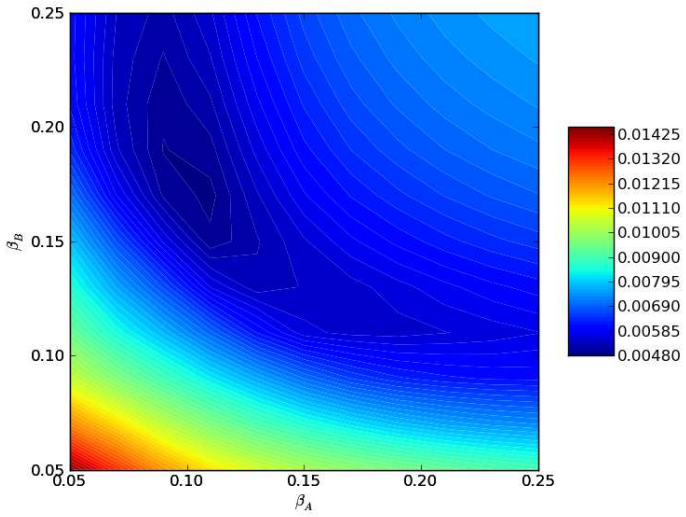


Figure 7.32: Tuning of β_A and β_B for the PKF-based method on our simulated DW volume at $b = 4500s \cdot mm^{-2}$, with a nc-χ noise of parameters $\sigma = 16$ and $n = 4$.

7.3.6 Results on simulated data & discussion

We applied the PKF-based method with $f = 3$ and with the previously tuned parameters α_A, β_A, α_B and β_B for the configuration at $b = 4500s \cdot mm^{-2}$, with a nc-χ noise of parameters $\sigma = 16$ and $n = 4$. The Laplace-Beltrami regularization λ was set to 0.006 (regularization) in the initialization of the estimation error covariance matrix defined by $\mathbf{P_0} = ((1/V)\mathbf{I} + \lambda\mathbf{L})^{-1}$ (as in chapter 4). The initial weight at the first iteration was defined such that: $w(\mathbf{v}, \mathbf{v}') = w_{spatial}(\mathbf{v}, \mathbf{v}')$.

We show the results on the dODF fields in fig. 7.33. We also compare the PKF-based method with the "LMMSE & KF" technique. We can see that the PKF-based method rightly corrects the noisy data with a small smoothing effect. The zoomed dODF shows two finely retrieved crossing orientations. The "LMMSE & KF" technique produces a result close to the noise-free configuration with less smoothing effect than the PKF-based method.

The zoomed dODF presents two equivalent lobes, compared to the dODF generated by the PKF-based technique, although the difference is very subtle. Finally, the calculated MSE is smaller with the "LMMSE & KF" filter than with the PKF-based technique, but is of the same order of magnitude. Compared to the noisy MSE, the gain of the LMMSE & KF technique over the PKF-based method is less than 0.4 %. We can therefore conclude that the PKF-based method achieves comparable performances compared to the LMMSE & KF technique, while being more rigorous in the way the non-Gaussian noise nature is accounted.

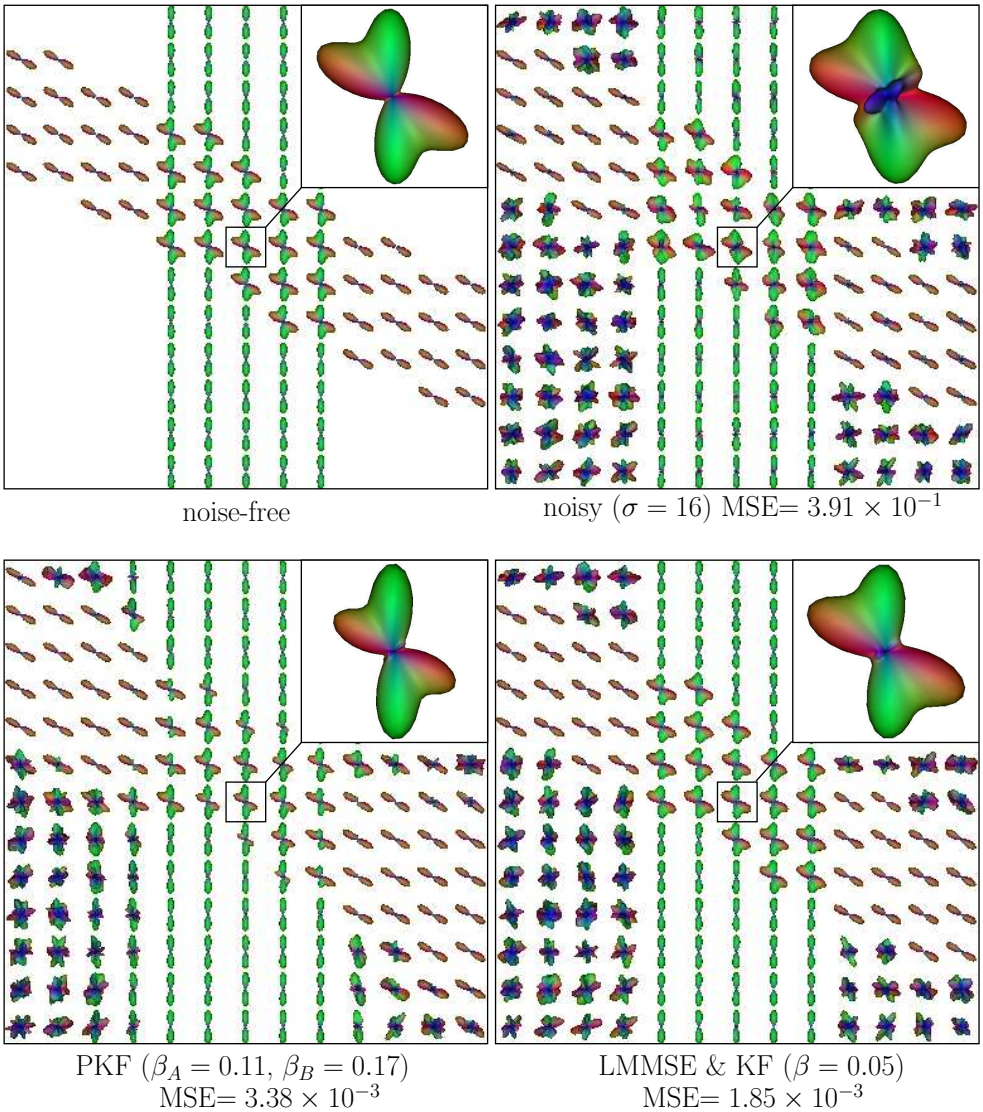
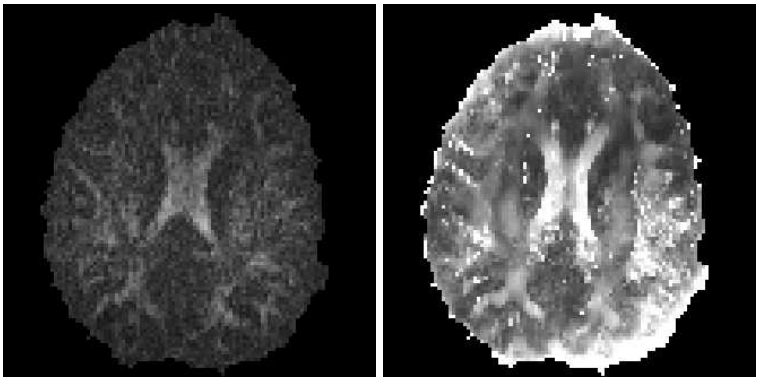


Figure 7.33: Comparison between the PKF-based method and the "LMMSE & KF" technique on our simulated data at $b = 4500s \cdot mm^{-2}$, with a nc-χ noise of parameters $\sigma = 16$ and $n = 4$.

7.3.7 Results on real data & discussion

We applied the PKF-based method with $f = 3$ and with the previously tuned parameters α_A, β_A, α_B and β_B on real data at $b = 4500s \cdot mm^{-2}$. To account for the eventual noise correlations, we used $n_{eff} = 2.6$, as in the "LMMSE & KF" method.

Fig. 7.34 gives the result of the PKF-based method on the real GFA at $b = 4500s \cdot mm^{-2}$. We remark that after the application of the PKF technique, the GFA is enhanced with a smoothing effect in homogeneous regions. Some white pixels appear, which are due to too strong constraints imposed by w_B. Indeed, after several trials, we concluded that the PKF global method performed at its best without any feedback loop injected in the LMMSE and with $\alpha_B = 2$, $\beta_B = 0.3$ for w_B, and $n_{eff} = 3$. We compared the result obtained using these parameters with the "LMMSE & KF" method on GFA maps in fig. 7.35.



raw GFA

GFA after PKF
($n_{eff} = 2.6, \beta_A = 0.11, \beta_B = 0.17$)

Figure 7.34: Application of the PKF based method with the previously tuned β_A and β_B on a real data at $b = 4500s \cdot mm^{-2}$.

Fig. 7.35 presents an improved PKF result on the GFA map. The map is more consistent and does not contain any hyperintense voxel. Furthermore, the GFA values are enhanced in anisotropic regions (fiber bundles). The drawback of the method is the smoothing effect generated by our temporary solution of injecting a mean instead of the voxel-wise measured magnitude as input in the PKF. The results of this method are promising and the technique should be deeper investigated to understand the gap of performances between the applications of the technique on simulated data and real data. From fig. 7.35, it is visible that the "LMMSE & KF" technique is better than the PKF global method on real data: both resulting "LMMSE & KF" GFA maps are more accurate than the smoothed GFA map obtained with the PKF-based method. Although the GFA is as high or even higher in anisotropic regions after the PKF global method, compared to the "LMMSE & KF" GFA, the contrast in the map is not as good as with the latter method, especially because of the smoothing effect. This is confirmed in fig. 7.36, which shows the highest GFA ratio —measured as previously for fig. 7.10— obtained for the

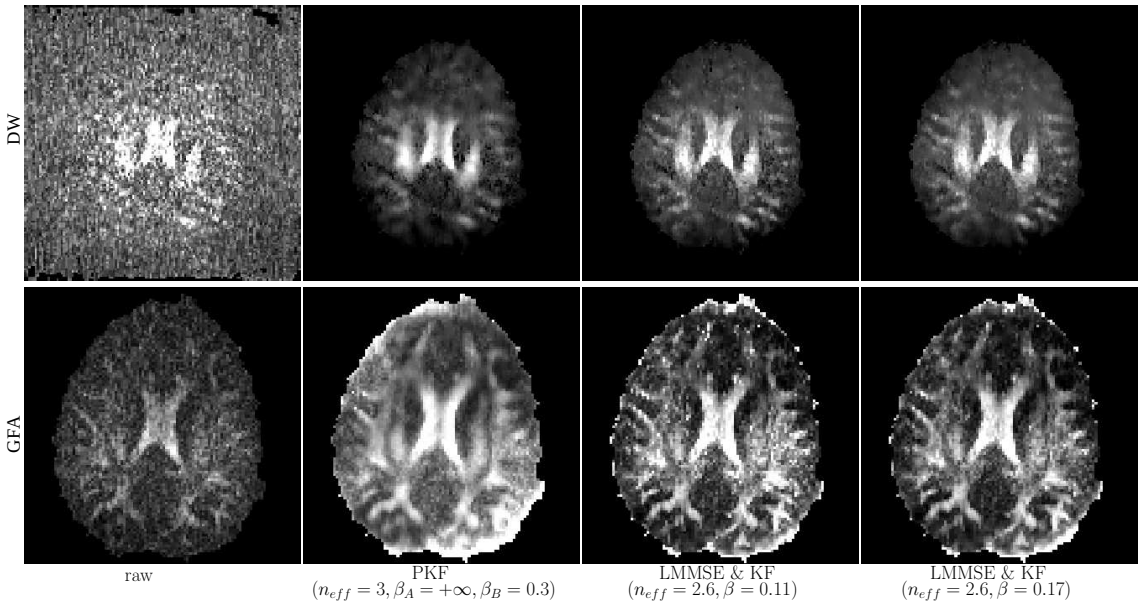


Figure 7.35: Application of the PKF based method with other parameters on a real data at $b = 4500s \cdot mm^{-2}$ and comparison with two configurations of the "LMMSE & KF" technique.

LMMSE & KF technique.

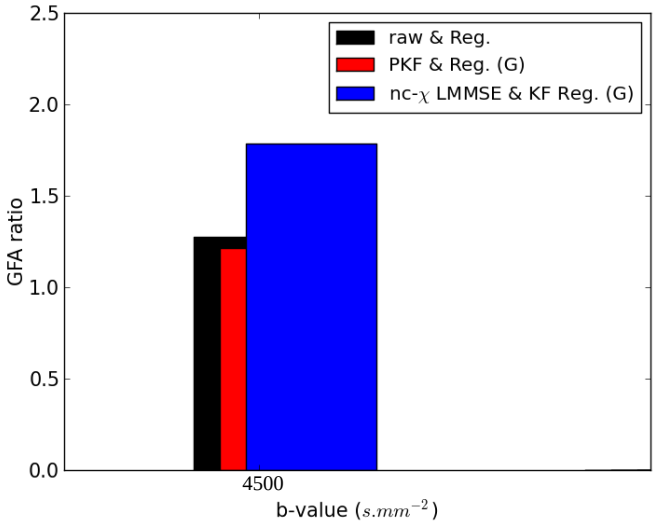


Figure 7.36: Comparison between the GFA ratios of the raw map and the map obtained after application of the PKF-based method and the LMMSE & KF at $b = 4500s \cdot mm^{-2}$ with the same parameters as in fig. 7.35.

Finally, concerning the computational efficiency of the PKF-based method, it was nearly as good as the LMMSE & KF technique. Indeed, for a maximum SH order set to 8, we compared the LMMSE & KF and the PKF durations. Whereas the LMMSE

& KF takes 131s, the PKF takes 138s. We did not perform the parallelization of the latter technique on the cluster, because of its weaker denoising results than the LMMSE & KF. Nevertheless, it could also be parallelized yielding a computation duration below the repetition time T_R.

7.3.8 Conclusion

The RT noise correction that we presented here relies on another incremental denoising technique than the LMMSE & KF technique. Its theoretical aspects seem more rigorous than the ones of the LMMSE & KF technique. However, from our practical results we showed that the PKF is equivalent to the LMMSE & KF on simulated data, but its results on real data are less satisfaying. This can be due to the fact that the nc-χ noise we are dealing with is better evaluated and outlined with the LMMSE rather than with the Gaussian mixture used in the PKF global method. We need further investigations on the PKF tricks to improve its quality on real data.

7.4 Comparison between the three RT noise correction techniques

7.4.1 Comparison

This subsection is dedicated to a comparison of the performances of the three RT noise correction techniques we proposed. This comparison is summarized in fig. 7.37.

In this figure, we gathered aspects that concerned us for our RT denoising objective. The theoretical aspects are different between the Gaussian low-pass filter and the two other techniques. Indeed, in the first case, the noise removal is performed on the **k**-space data, assuming an uncorrelated zero-mean Gaussian noise, whereas the two other techniques are performed on the magnitude data, accounting for a non-stationary nc-χ noise process. This is a fundamental difference which leads to different mathematical approaches in the denoising algorithms.

Another essential aspect of the techniques is their dependence on a critical parameter tuning or not. The Gaussian low-pass filter is a simple method, which only involves one parameter (σ_{Gauss}) to tune. However, as we saw in section 7.2, this parameter has a huge impact on the result: a too high parameter could blur the results. Moreover, as the Gaussian low-pass filter is directly integrated in the Siemens reconstruction pipeline, we had to extract raw data from the MRI scanner to test the method. Additionally, we could not test it on simulated data. Therefore, it was impossible to perform the same tuning approach of this parameter as for the β parameter used in the LMMSE & KF technique. Finally, σ_{Gauss} may have to be tuned for different dMRI sequence configurations, e.g. with different *b-values*. In comparison to the Gaussian low-pass filter, both LMMSE & KF and PKF-based methods rely on the tuning of two or three parameters. To tune the n_{eff} parameter, we proposed an heuristic based on several trials on the real data. This

Aspects	**LMMSE & KF**	**Gaussian low-pass filter**	**PKF-based method**
THEORY			
-targetted image	magnitude in spatial domain	real & imaginary **k**-space data	magnitude in spatial domain
-noise nature	non-stationary nc-χ	assumed stationary zero-mean Gaussian	non-stationary nc-χ
-number of parameters to be tuned (not easily)	2 (n_{eff}, β)	1 (σ_{Gauss})	2 or 3 (n_{eff}, (β_A), β_B)
PRACTICAL USE			
On our simulated data:			
-MSE reduction	+++	N/A	++
-visual noise reduction	+++	N/A	+++
-control of the smoothing effect	+++	N/A	++
-retrieved angular crossing information	+++	N/A	++
On our real data:			
-visual noise reduction	+++	++	++
-control of the smoothing effect	+++	+	-
-gain in the GFA ratio	+++	++	-
-small computation time	++	+++	++
-easily implemented & tested	+++	- -	+++

Figure 7.37: Comparison between theoretical and practical aspects of the three RT noise correction techniques.

heuristic is not optimal, as it is not an automatic RT tuning of the parameter. However, it could be used for a study on a database acquired with the same parameters (like the "Archi database"). Concerning the tuning of β, we proposed in the LMMSE & KF and PKF-based techniques to tune it on simulated data first and to report the values of β found for the simulated data on the real data. Indeed, for the LMMSE & KF technique, the β value had to be increased from the application on the simulated data to the real data. A similar increase had to be performed for the PKF-based method. Nevertheless, we noticed that a value of around 0.11 for β was shown to yield accurate results on the real data, independently from the *b-value*. A certain caution has still to be kept to avoid white pixels to appear, due to a too low β value at high *b-values*. A similar analysis can be done for the PKF global method.

Concerning the practical use, we could compare both LMMSE & KF and PKF techniques on our simulated data with the noise-free simulated reference. Different aspects of the noise removal were compared: the MSE reduction, the visual noise reduction, the visual evaluation of the control of the smoothing effect, the retrieval of the angular crossing information. Both methods were shown to propose accurate results regarding these criteria, with a slightly better performance with the LMMSE & KF technique.

The comparison between the three RT methods performed on the real data showed that all methods achieved a visual noise reduction, with a higher smoothing effect for the Gaussian low-pass filter and the PKF-based method. The measurement of the GFA ratio before and after each correction confirmed this visual comparison. The smallest computation time was obtained with the Gaussian low-pass filter. Both other methods respected the RT constraints (a computation time lower than the repetition time), after a parallelization and distribution on a cluster of 80 CPUs. Finally, these latter methods were easily implemented and tested as they did not require any implementation inside the reconstruction pipeline of the manufacturer like for the Gaussian low-pass filter.

7.4.2 Future prospects

The comparison between the three different RT methods brought several ideas for future prospects.

Concerning the Gaussian low-pass filter, a future prospect would first be to test the Gaussian low-pass filter on DW data with high *b-values* to better probe the performances of this solution at low signal levels. Concerning the LMMSE & KF and PKF techniques, we could also further investigate very high *b-values* for all studies focused on a very high diffusion resolution imaging. For the PKF-based method, as its results are promising, especially on the simulated data, it would be required to first understand and correct the weaker performances obtained on the real data.

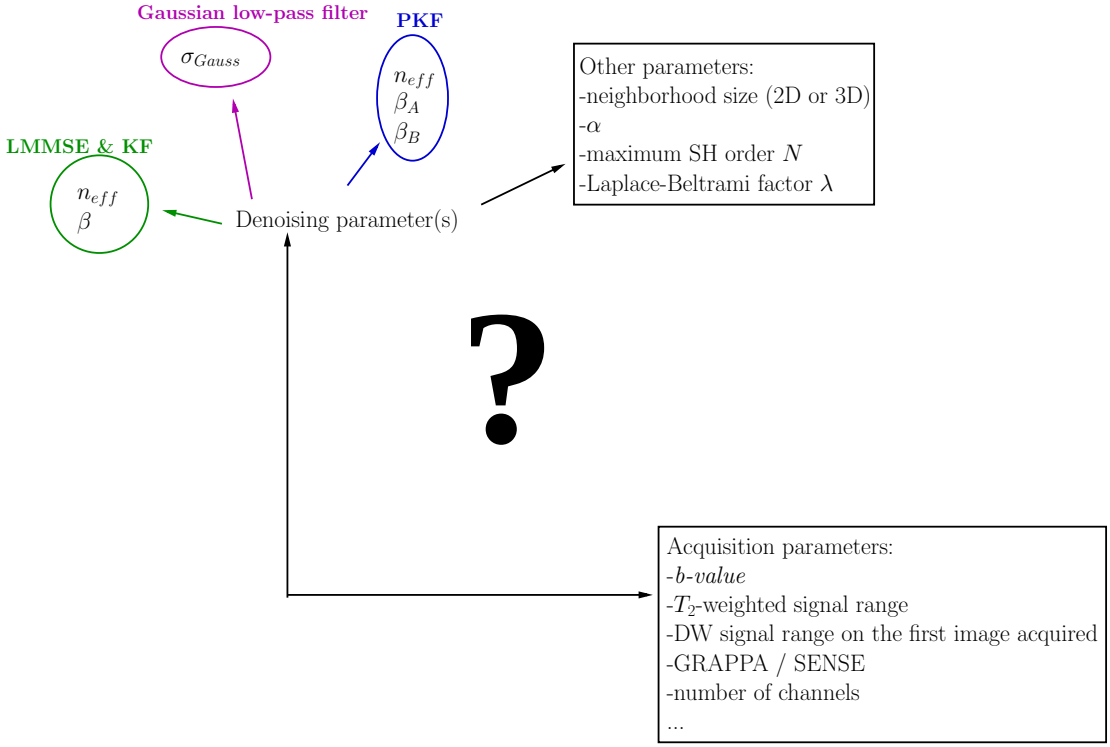


Figure 7.38: Future prospects about the tuning of image-dependent parameters.

An other prospect would be to attempt finding an automatical solution for tuning the parameters used in all methods. This is shown by fig. 7.38. An idea would be to find a relation between the parameters of the denoising method and the acquisition parameters defining the quality level of the raw data (the *b-value*, the range of intensities in the T_2-weighted image, the range of intensities in the first DW volume...), as well as its noise characteristics (the use of GRAPPA/ SENSE..., the number of channels...). The ideal solution would be to propose an autocalibration of the denoising parameters with respect to these acquisition parameters.

Finally, concerning the methods performed on the magnitude data (LMMSE & KF and PKF), it would be interesting to further investigate the voxel-wise estimation of the effective parameters (n_{eff},σ_{eff}) introduced by [Aja-Fernández et al. (2011); Aja-Fernández and Tristán-Vega (2012); Aja-Fernández et al. (2013)] to accurately account for the true noise distribution at each voxel of the data and the non-stationarity of the noise across the voxels.

7.5 Conclusion of this chapter

This chapter attempted to answer the requirement of a RT noise correction. It proposed three methods. The LMMSE & KF, embedding an LMMSE filter and a KF with a feedback loop, was shown to yield very accurate and improved results. The Gaussian low-pass filter, which is integrated in the Siemens reconstruction pipeline represents an interesting alternative closer to the acquired signal. Finally, the PKF propose a rigorous mathematical framework which achieved accurate results on simulated data. This latter technique would require a further study to improve its performances on real data. Finally, through this chapter, we tackled the typical problem of tuning the parameters and of decreasing the computation times to respect RT goals. We proposed some future prospects aiming at improving the RT noise removal. The contributions of our work are listed below.

7.5.1 Contributions of this chapter

- An LMMSE integration in the KF-based framework dedicated for RT purpose. This integration includes a feedback loop enabling to compute anisotropic averages required by the LMMSE to limit the smoothing effects. This contribution can be found in [Brion et al. (2010, 2011d,f,e)].
- A correction scheme located before the SoS reconstruction and which has the advantage to only deal with a zero-mean Gaussian noise. This scheme was incorporated to the Siemens reconstructor.
- A parallel Kalman filter as an alternative solution for denoising the MRI magnitude images. This contribution can be found in [Brion et al. (2012a,b,c)].
- A comparison between the methods.

7.6 Appendix A: More details about the derivation of the PKF filter

In the appendix of [Plataniotis et al. (1997)], the authors detailed the derivation of the PKF. They proceeded from analogy with the KF idea. We follow their mathematical steps here.

Let us write our linear model, introduced in chapter 4:

$$\boxed{\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{\epsilon}}, \text{ with } \begin{cases} \mathbf{y}\text{: the K}\times 1 \textbf{ observation vector}, \\ \mathbf{A}\text{: the K}\times\text{N design matrix}, \\ \mathbf{x}\text{: the N}\times 1 \textbf{ state vector}, \\ \boldsymbol{\epsilon}\text{: the K}\times 1 \text{ vector of } \textit{i.d.} \text{ noises distributed according to } \mathcal{N}(0, \mathbf{R}), \end{cases} \tag{7.17}$$

A Bayesian KF can be applied on this model, as seen in chapter 4, if the initial state vector $\mathbf{x}_0$ follows a normal distribution with known mean and variance. Concerning the demonstration of this KF, it relies on the Bayes' theorem, which expresses the *a posteriori* density $p\left(\mathbf{x}_i|y_i, Y^{i-1}\right)$. This **posterior** is the PDF of the **parameters** $\mathbf{x}_i$ given the **data** y_i at iteration i and given the data $Y^{i-1} = [y_0, y_1, ..., y_{i-1}]^{\mathrm{T}}$ of all the previous iterations before i. This posterior PDF is expressed following the Bayes' rule such that:

$$\begin{aligned} p\left(\mathbf{x}_i|y_i, Y^{i-1}\right) &= \frac{p\left(y_i|\mathbf{x}_i, Y^{i-1}\right) p\left(\mathbf{x}_i|Y^{i-1}\right)}{p\left(y_i|Y^{i-1}\right)} \\ &= \frac{p\left(y_i|\mathbf{x}_i, Y^{i-1}\right) p\left(\mathbf{x}_i|Y^{i-1}\right)}{\int p\left(y_i|\mathbf{x}_i, Y^{i-1}\right) p\left(\mathbf{x}_i|Y^{i-1}\right) d\mathbf{x}_i}, \end{aligned} \tag{7.18}$$

where the term $p\left(y_i|\mathbf{x}_i, Y^{i-1}\right)$ is designated as the **likelihood**, which corresponds to the conditional density of the data y_i given the parameters $\mathbf{x}_i$ and the previous data Y^{i-1}. The term $p\left(\mathbf{x}_i|Y^{i-1}\right)$ is the **prior**, *i.e.* the distribution of the parameters in the absence of any data at the considered iteration i. Finally the denominator in eq. 7.18 is just a normalization factor to ensure that $p\left(\mathbf{x}_i|y_i, Y^{i-1}\right)$ integrates to 1. This Bayes' rule is used to determine the posterior, which then enables to obtain the MMSE estimator:

$$\begin{aligned} \hat{\mathbf{x}}_i &= E\left(\mathbf{x}_i|y_i, Y^{i-1}\right) \\ &= \int \mathbf{x}_i \cdot p\left(\mathbf{x}_i|y_i, Y^{i-1}\right) d\mathbf{x}_i. \end{aligned} \tag{7.19}$$

From this MMSE estimator, the KF can be derived using the hypotheses inherent to the linear model on which the KF relies.

Consequently, the KF only depends on the estimation of the likelihood $p\left(y_i|\mathbf{x}_i, Y^{i-1}\right)$ and on the prior $p\left(\mathbf{x}_i|Y^{i-1}\right)$ to determine the posterior. The KF theory requires that the initial state vector $\mathbf{x}_0$ follows a normal distribution with known mean and variance. This hypothesis permits to explicit the prior. Concerning the likelihood, it can be expressed

such that:

$$\begin{aligned} p\left(y_i|\mathbf{x}_i, Y^{i-1}\right) &= p\left(y_i|\mathbf{x}_i\right) \text{assuming that the data } y_i \text{ are independent for all } i \\ &= p_{\epsilon_i}\left(y_i - \mathbf{b_i}\cdot\mathbf{x}_i|\mathbf{x}_i\right) \text{ as } y_i = \mathbf{b_i}\cdot\mathbf{x}_i + \epsilon_i, \\ &= p_{\epsilon_i}\left(y_i - \mathbf{b_i}\cdot\mathbf{x}_i\right), \end{aligned} \tag{7.20}$$

assuming that ϵ_i is independent of $\mathbf{x}_i$ [Kay (1993); Plataniotis et al. (1997)]. Then, using the hypothesis of a zero-mean Gaussian noise, which is a requirement to use the KF, the likelihood is equal to:

$$p\left(y_i|\mathbf{x}_i, Y^{i-1}\right) = \mathcal{N}\left(\hat{y}_i, \mathbf{P}_{y_i}\right), \tag{7.21}$$

with the mean of the latter Gaussian defined using the KF equations such that:

$$\begin{aligned} \hat{y}_i &= E\left(y_i|Y^{i-1}\right), \\ &= E\left(\mathbf{b_i}\cdot\mathbf{x}_i + \epsilon_i|Y^{i-1}\right), \\ &= \mathbf{b_i}\cdot\hat{\mathbf{x}}_i, \end{aligned} \tag{7.22}$$

as $E\left(\mathbf{b_i}\cdot\mathbf{x}_i + \epsilon_i|Y^{i-1}\right) = E\left(\mathbf{b_i}\cdot\mathbf{x}_i|Y^{i-1}\right) + E\left(\epsilon_i|Y^{i-1}\right) = \mathbf{b_i}\cdot E\left(\mathbf{x}_i|Y^{i-1}\right)$ because ϵ_i is a zero-mean noise. Concerning the covariance matrix $\mathbf{P}_{y_i}$, it is the covariance of the following measurement error: $y_i - \hat{y}_i = \mathbf{b_i}\cdot(\mathbf{x}_i - \hat{\mathbf{x}}_i) + \epsilon_i$. Therefore, $\mathbf{P}_{y_i}$ is equal to:

$$\mathbf{P}_{y_i} = \mathbf{b_i}\cdot\mathbf{P}_i\cdot\mathbf{b_i}^{\mathrm{T}} + \mathbf{R}. \tag{7.23}$$

With this information, how can this KF be improved to account for non-Gaussian noise? At this point of the appendix of [Plataniotis et al. (1997)], an analogy is performed with the KF to build an estimator close to the KF. Instead of having a zero-mean Gaussian noise, [Plataniotis et al. (1997)] expressed the noise PDF as a Gaussian mixture:

$$p(\epsilon_i) = \sum_{g=1}^{f} p_g \mathcal{N}(\mu_g, \sigma_g^2), \tag{7.24}$$

with p_g the weight of the Gaussian $\mathcal{N}(\mu_g, \sigma_g^2)$ of mean μ_g and variance σ_g^2. Similarly as before, [Plataniotis et al. (1997)] deduced that the PDF $p\left(y_i|\mathbf{x}_i, Y^{i-1}\right)$ is also a Gaussian mixture, such that:

$$p\left(y_i|\mathbf{x}_i, Y^{i-1}\right) = \sum_{g=1}^{f} w_g \mathcal{N}(\hat{y}_g, P_{y_g}), \tag{7.25}$$

with w_g the weight of the Gaussian $\mathcal{N}(\hat{y}_g, P_{y_g})$ of mean $\hat{y}_g$ and covariance matrix P_{y_g}. The Gaussian $\mathcal{N}(\hat{y}_g, P_{y_g})$ describes the contribution of the g^{th} elemental Gaussian term to the density $p\left(y_i|\mathbf{x}_i, Y^{i-1}\right)$ and we write it as $p\left(y_i|\mathbf{x}_i, Y^{i-1}, b_g\right)$ (b_g being the indicator of the g^{th} member). From the Bayes' rule, w_g is expressed such that: $w_g = p\left(b_g|y_i, \mathbf{x}_i, Y^{i-1}\right)$

At this point, Plataniotis *et al.* applied the following **Theorem 2.**, in which $p\left(b_g|y_i, \mathbf{x}_i, Y^{i-1}\right)$ simplifies to $p\left(b_g|i\right)$ [Laniotis (1976)]:

Theorem 2. *Let us consider the following equation:*

$$\mathbf{y}(i) = \mathbf{A}(i, b_g) \cdot \mathbf{x}(i) + \boldsymbol{\epsilon}(i),$$

where $\mathbf{x}(i)$ *and* $\mathbf{y}(i)$ *are the n- and m-dimensional state and measurement processes, respectively, at the* i^{th} *iteration.* $\boldsymbol{\epsilon}(i)$ *is the measurement noise random process, which conditioned on* b_g $(\forall g \in [\![1; f]\!])$ *is independent, zero-mean and white Gaussian, with covariance* $\mathbf{R}(i)$*. The initial state vector* $\mathbf{x}(0)$ *is independent of* $\boldsymbol{\epsilon}(i)$ *when conditioned on* b_g*, and has a* b_g*-conditional Gaussian density with mean and variance* $\mathbf{x}(0|b_g)$ *and* $\mathbf{P}(0|b_g)$*, respectively.*

The unknown parameters b_g*, which if known would completely specify the model, are assumed to be time-invariant random parameters with known or assumed a priori PDF* $p(b_g)$*.*

The MMSE estimate $\hat{\mathbf{x}}(i)$ *of* $\mathbf{x}(i)$ *and the corresponding error covariance matrix* $\mathbf{P}(i)$ *are given by:*

$$\hat{\mathbf{x}}(i|i) = \sum_{g=1}^{f} \hat{\mathbf{x}}(i|i, b_g) \cdot p(b_g|i), \tag{7.26}$$

$$\mathbf{P}(i|i) = \sum_{g=1}^{f} \left(\mathbf{P}(i|i, b_g) + [\hat{\mathbf{x}}(i|i, b_g) - \hat{\mathbf{x}}(i|i)][\hat{\mathbf{x}}(i|i, b_g) - \hat{\mathbf{x}}(i|i)]^{\mathrm{T}}\right) \cdot p(b_g|i), \tag{7.27}$$

where $\hat{\mathbf{x}}(i|i, b_g)$ *and* $\mathbf{P}(i|i, b_g)$ *are the* b_g*-conditional MMSE state vector estimate and the corresponding* b_g*-conditional error covariance matrix. The a posteriori PDF* $p(b_g|i)$ *is expressed such that:*

$$p(b_g|i) = \frac{L(i|b_g)p(b_g|i-1)}{\sum_{h=1}^{f} L(i|b_h)p(b_h|i-1)}, \tag{7.28}$$

where

$$L(i|b_g) = \frac{1}{\sqrt{2\pi}}|\mathbf{P}_{\mathbf{y}}(i|i-1, b_g)|^{-1/2} \cdot \exp\left(-\frac{\|\tilde{\mathbf{y}}(i|i-1, b_g)\|^2}{2\mathbf{P}_{\mathbf{y}}(i|i-1, b_g)}\right), \tag{7.29}$$

where the model-conditional innovation process $\tilde{\mathbf{y}}(i|i-1, b_g)$*, defined as* $\tilde{\mathbf{y}}(i|i-1, b_g) = \mathbf{y}(i) - \mathbf{A}(i, b_g)\hat{\mathbf{x}}(i|i-1, b_g)$*, is a white noise process conditioned on* b_g*, with* b_g*-conditional covariance matrix:*

$$\mathbf{P}_{\mathbf{y}}(i|i-1, b_g) = \mathbf{A}(i, b_g)\mathbf{P}(i|i-1, b_g)\mathbf{A}^{\mathrm{T}}(i, b_g) + \mathbf{R}(i). \tag{7.30}$$

Following the general solution to multiple-model estimation problem given by the latter theorem (eq. 7.26 to eq. 7.30), [Plataniotis et al. (1997)] derived the PKF equations regrouped in fig. 7.26. They replaced the term $p(b_g|i-1)$ by w_g, which is the coefficient of the corresponding Gaussian PDF in the Gaussian mixture in eq. 7.25. After the application of the **Theorem 2.**, [Plataniotis et al. (1997)] replaced the final a posteriori Bayesian estimate of the density $p\left(y_i|\mathbf{x}_i, Y^{i-1}\right)$ by a Gaussian distribution to inject the latter in a simple KF. This approximation by a collapsed Gaussian density is valid if the

Gaussian members of the original mixture are close. Using the **Theorem 2.**, the final collapsed Gaussian density has its mean and covariance at iteration i equal to:

$$\hat{y} = \sum_{g=1}^{f} w_g \hat{y}_g, \tag{7.31}$$

$$P_y = \sum_{g=1}^{f} w_g \left[P_{y_g} + (\hat{y} - \hat{y}_g)^2 \right]. \tag{7.32}$$

Part IV

Real-time tractography application

Chapter 8

Inference of the connectivity in RT

Along the thesis manuscript, we dealt with online dMRI processing possibilities. We focused on diffusion local models which enable to reconstruct the diffusion or fiber orientation distribution function inside the whole brain. These models, close to the DW data, constitute the first step to infer the distribution of the white matter (WM) fiber bundles inside the brain. The next step is performed using tractography, as seen in chapter 3, section 3.3. In this chapter, we address the feasibility of tracking the digital fiber tracts in RT and see them being refined at each new diffusion measurement during the exam. Such an approach is delicate, as the tractography algorithms are computationally demanding. What are the motivations behind this innovative idea? What is our first approach and what are the future prospects concerning RT tractography? In this chapter, we detail the prototype we tested as a first step towards an inference of the connectivity in RT. We show its results and highlight the effect on the tracts of our RT noise correction based on the LMMSE & KF framework.

8.1 Introduction

Tractography offers an information about the organization of the fiber tracts in the brain. Although a fiber tract is a digital pathway and does not necessarily correspond to an anatomical white matter (WM) fiber, it gives the most likely pathway with respect to the diffusion process that could use a WM fiber to connect two areas of the brain. Having access to the structural connectivity is of importance in clinical applications either to detect any atrophy of WM fiber bundles and better understand the physiopathology of brain diseases, or to avoid damaging WM pathways when performing surgical interventions. Several navigators have been developed to make the visualization of the tracts more flexible for neurosurgical planning, like for example [Golby et al. (2011); Vaillancourt et al. (2010); Mittmann et al. (2011); Chamberland et al. (2012)].

Because most tractography techniques are computationally intensive, no RT algorithm has been introduced up to now. This processing has only be performed offline yet. We wanted to address its feasibility in RT. Our motivations are twofold. First, a RT trac-

tography technique would improve a medical treatment or intervention if the tracts could be inferred during the dMRI exam. Indeed, it would accelerate the medical care of the patient, improving its chances of success. Some acceleration of certain tractography algorithms have been proposed in the literature with the aim of facilitating on-site diagnosis and making the acquisition of supplementary scans possible, while the patient is still in the MRI scanner [Kwatra et al. (2006); Singh et al. (2006)]. What we propose here, is different. Our objective is to perform tractography truly in RT, meaning that after each new diffusion measurement during the dMRI sequence, a map of tracts is incrementally refined. This process would allow to go further into the analysis. This brings us to our second motivation behind RT tractography which is to improve the methodology itself since the intermediate results can be exploited to improve the analysis. A RT workflow for tractography would give the possibility to adapt the parameters of the dMRI sequence and/or of the tractography algorithm during its own execution.

8.2 RT tractography workflow

8.2.1 Objectives

To study the feasibility of performing tractography in RT, we divided this ambitious project into several steps to reach, listed below:

1. The first step consists of running a tractography algorithm at each new iteration. This first evaluation would permit to judge about the quality and the possible exploitation of the intermediate tracts. In a first attempt, the RT constraint may not be satisfied, but would allow to queue the incoming DW data to process any new intermediate tractogram as soon as possible. In order to take into account *a minima* the time constraint, we decided to implement it using a streamline fiber tracking algorithm (deterministic or probabilistic) as it is known to be the most computationally efficient.

2. The second step would consist of distributing the code on a cluster of 80 CPUs to meet the RT requirement. Because we used a streamlining technique, this parallelization step is straightforward and is actually a work in progress.

3. The third step would investigate the possibility to transform the actual streamlining process into an incremental process in order to gain in efficiency to be able to make streamlines infered at iteration $(i-1)$ evolve from the knowledge of the DW data acquired at iteration i rather than reprocessing the entire streamline from the updated ODF field.

4. The fourth step would ultimately investigate the feasibility of global tractography in an incremental way to provide more accurate tracts than blind streamline tractography algorithms. In opposition to streamlining, this global solution would have to

account for the interactions between the tracts and would no longer consider each tract independently of the others.

In the frame of this thesis, we achieved the first step and we are currently distributing it on an RT environment.

8.2.2 RT results and discussion

RT results on simulated data & discussion.

We applied the RT streamline deterministic tractography workflow, using the aQBI model RT results, with a maximum spherical harmonic (SH) order set to 8 yielding 45 components of the coefficients' vector $\mathbf{C}^{\mathrm{DW}}$ of the decomposition of the normalized DW signal on the modified SH basis. The aQBI model was performed without any Laplace-Beltrami regularization on a noise-free simulated DW field, and with the Laplace-Beltrami regularization factor λ set to 0.006 on a noisy simulated DW field. The simulated data were generated with $b = 4500s \cdot mm^{-2}$, for which a nc-χ noise addition considering $n = 4$ channels and a noise standard deviation of $\sigma = 16$ was performed, assuming no noise correlation between the channels. More details on the generation of these simulated data can be found in appendix C, at the end of the manuscript. The streamline deterministic tractography was executed with 27 seeds per voxels to create a dense tractogram, with a streamlining forward step of 1.25 mm (corresponding to $0.25\times\mathrm{res}_v$, with $\mathrm{res}_v = 5$mm the voxel resolution), with a minimum and a maximum fiber lengths set to 5 mm and 191 mm, respectively. The aperture angle, in which the tracts are allowed to grow, was set to 60°. We studied the connectivity between ROI 1 and ROI 3 and the connectivity between ROI 2 and ROI 4 (see fig. 8.1). When performing an additional RT noise correction, we used the "LMMSE & KF" RT denoising method with a $5\times5\times5$ neighborhood and with the Laplace-Beltrami regularization factor λ set to 0.006. The values of the α and β parameters were 2 and 0.05, respectively, similar to what was chosen in the previous chapter, in subsection 7.1.3.

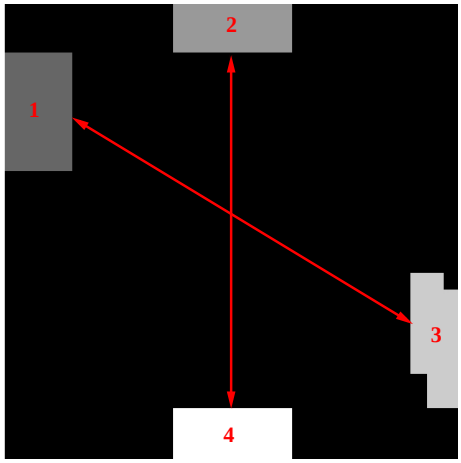


Figure 8.1: Description of the connectivity studied in our simulated data.

During the iterative process, the local dODFs remain quite smooth until the number of acquisitions reach the number of SH coefficient components corresponding to the minimum number of data required to correctly solve the underlying inverse problem of dODF (fig. 8.2). During these iterations, the intermediate tractograms depict almost random fibers (or no fiber at all) as the corresponding dODFs do not provide a coherent angular profile. In the following iterations of the process, the dODFs become sharper and the branches of the phantom become visible with a coherence of the fibers increasing at the same time of the sharpening of the dODFs (fig. 8.3). When noise is not corrected, the intermediate tractograms depict a fanning phenomenon due to the loss of accuracy in the main directions of the dODFs. One can notice that this fanning effect becomes less severe after each new acquisition because the Laplace-Beltrami regularization helps dealing a bit with noise even if it does not sharpen the dODFs.

When an LMMSE & KF denoising method is inserted in the RT pipeline between the acquisition and the dODF computation, the quality of the intermediate tractograms is significantly improved. The dODFs are sharpened and the main directions of the dODFs become less corrupted after few iterations providing quite reasonable results after only a fifth of the acquisitions. The final result does not present any spurious fiber like in the uncorrected RT tractography and looks quite similar to the ground truth tractogram.

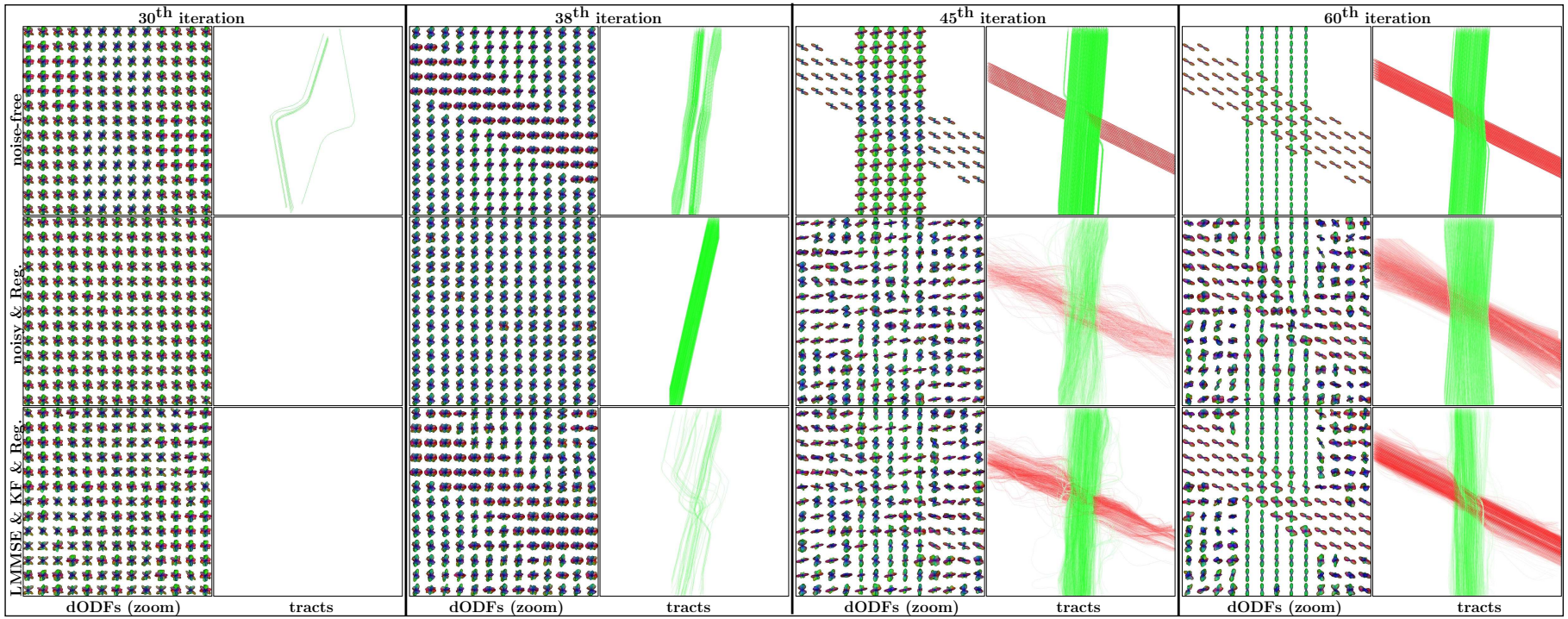


Figure 8.2: Tractography and dODF RT results on our simulated data along the RT process at $b = 4500s \cdot mm^{-2}$. The streamline deterministic tractography algorithm was performed using the dODF results obtained with the aQBI model with a maximum SH order N set to 8. When "Reg." is indicated, it means that the Laplace-Beltrami regularization λ was set to 0.006 (else, it was set to 0). All maps of tracts were displayed with the same opacity.

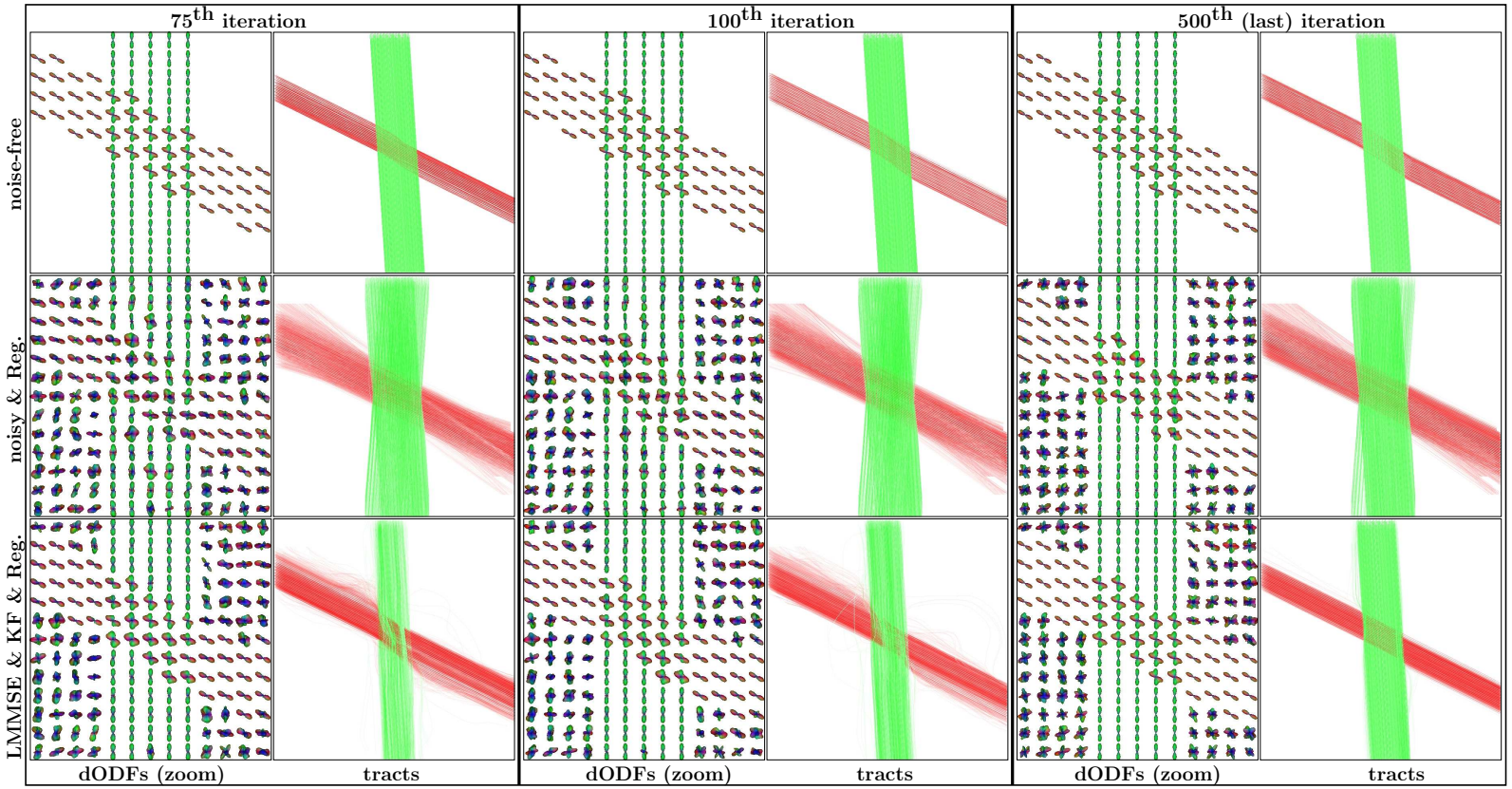


Figure 8.3: Tractography and dODF RT results on our simulated data along the RT process at $b = 4500s \cdot mm^{-2}$. The streamline deterministic tractography algorithm was performed using the dODF results obtained with the aQBI model with a maximum SH order N set to 8. When "Reg." is indicated, it means that the Laplace-Beltrami regularization λ was set to 0.006 (else, it was set to 0). All maps of tracts were displayed with the same opacity.

RT results on real data & discussion.

We applied the RT streamline deterministic tractography workflow, using the aQBI model RT results obtained with a maximum SH order first of 4 and then of 6 and with the regularization Laplace-Beltrami factor λ set to 0.006, on the real data of the "Archi database" acquired at $b = 4500s \cdot mm^{-2}$. For more details on the acquisition parameters of these real data, we refer the reader to the appendix A at the end of the manuscript. The streamline deterministic tractography was executed with 27 seeds per voxels, with a streamlining forward step of 0.5 mm, with a minimum and a maximum fiber lengths set to 5 mm and 200 mm, respectively. The aperture angle, in which the tracts are allowed to grow, was set to 60°. As for the simulated data, we additionally performed the "LMMSE & KF" RT denoising method using a 5×5×5 neighborhood. The value of the α parameter was 2, similarly as before. Concerning the β parameter, we set it to 0.11, as in chapter 7, subsection 7.1.4. The n_{eff} parameter was set to 1.8.

Fig. 8.4 depicts the results obtained at the 10^{th}, 19^{th}, 29^{th}, 39^{th} and at the final 60^{th} iterations for a maximum SH order N set to 4. With this latter parameter choice, we remark that the estimation of the tracts is efficient only starting from the 19^{th} iteration. This is consistent with the previous results shown in fig. 7.14 in the previous chapter, page 192, where we can see that the dODF map was presenting some structural information only from the 19^{th} iteration. This is due to the fact that the number of coefficients is equal to 15 with a maximum SH order set to 4. As long as there are less measurements than the number of unknowns, the estimation is not accurate. Concerning the next iterations in fig. 8.4, it is visible that the tracts gain in density along the iterations. When comparing the raw tracts with the tracts obtained after the application of the RT LMMSE & KF noise correction technique, it is visible that the tracts obtained with the correction present a higher coherence and the maps are globally clearer. Some tracts, which are absent in the raw configuration at the early iterations (19^{th} and 29^{th}), are better detected when using the noise removal.

Fig. 8.5 shows the same results for a maximum SH order N set to 6. For this higher parameter, the noise is a bit higher as more components of the coefficient of the DW signal are kept. Here, we see that the estimation of the tracts is efficient only from the 29^{th} iteration. This is again due to the fact that the number of coefficients is equal to 28 with a maximum SH order set to 6. As long as there are less measurements than the number of unknowns, the estimation is not accurate. This is again consistent with the previous results shown in fig. 7.15 in the previous chapter, page 193, where we can see that the dODF map was presenting some structural information only from the 29^{th} iteration. For the other iterations, the results are very similar to the ones obtained in fig. 8.4, with $N = 4$. Both figures highlight the improvement achieved with the LMMSE & KF on the tracts.

Finally, the computation time of the streamline deterministic tractography algorithm for a whole brain, with here 27 seeds per voxel, was of around 20 min for each iteration. The algorithm version tested here contained a parallelization on four CPUs. The high compu-

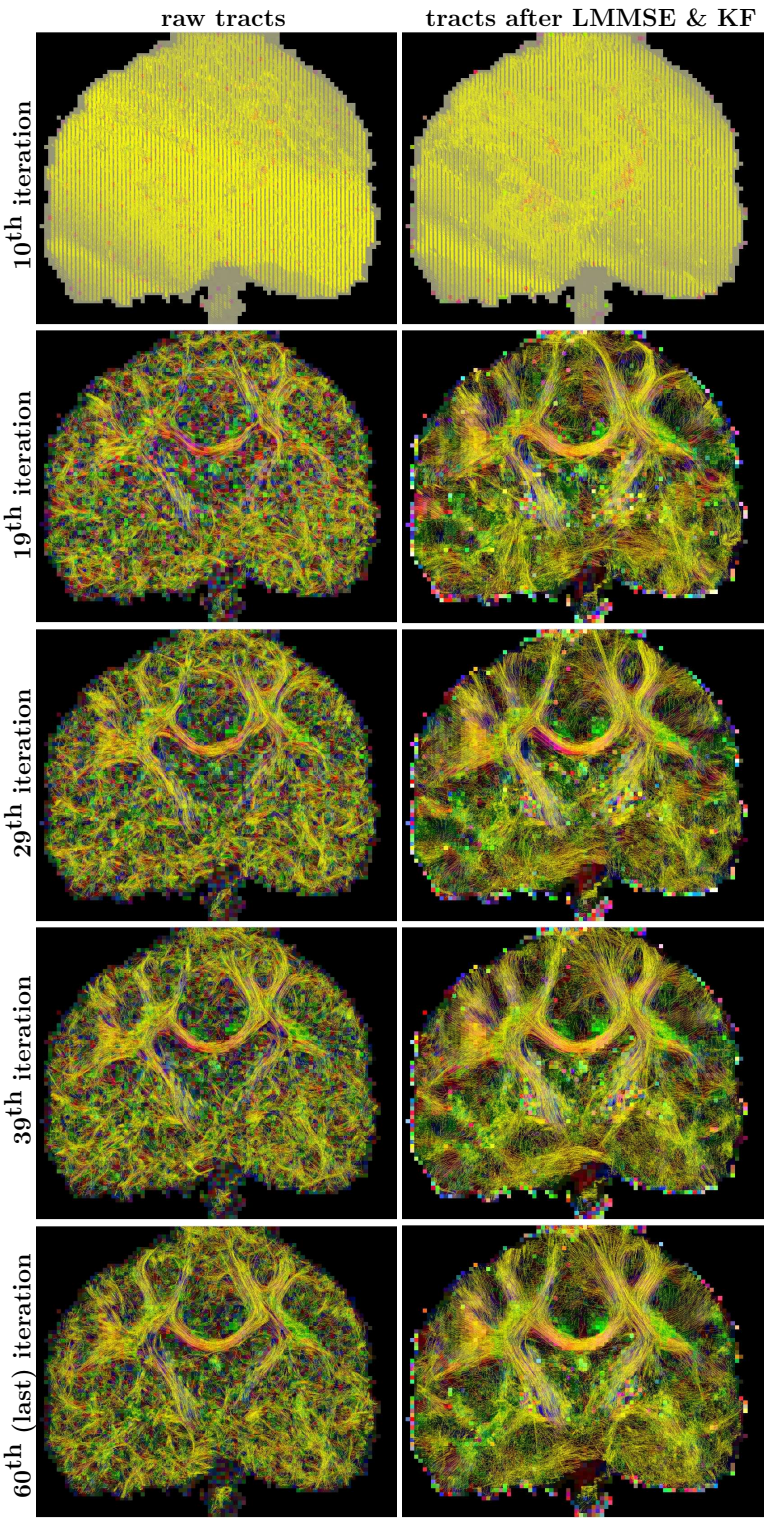


Figure 8.4: Tractography RT results on our real data along the RT process at $b = 4500s \cdot mm^{-2}$. The streamline deterministic tractography algorithm was performed using the dODF results obtained with the aQBI model with a maximum SH order N set to 4. The Laplace-Beltrami regularization λ was set to 0.006.

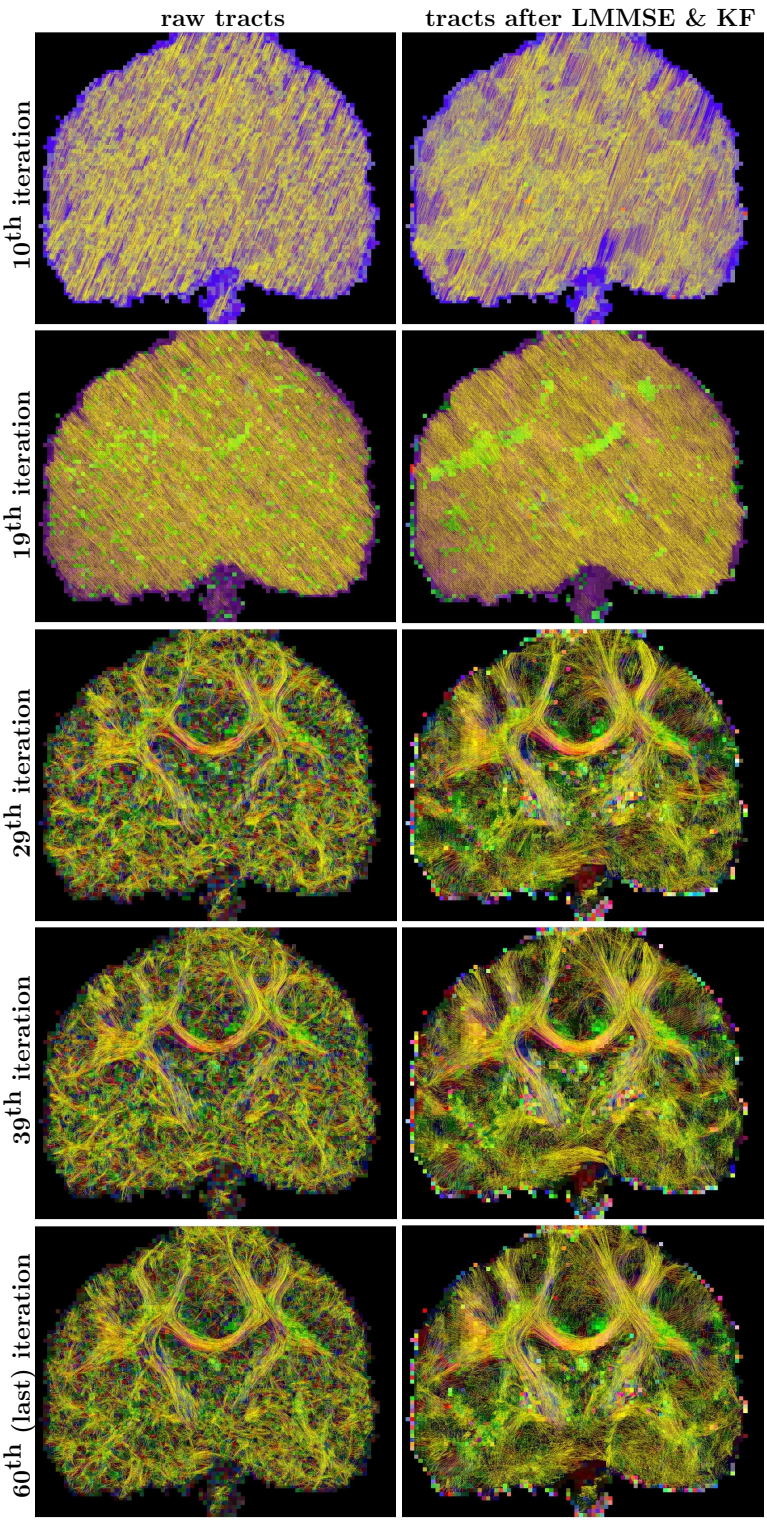


Figure 8.5: Tractography RT results on our real data along the RT process at $b = 4500s \cdot mm^{-2}$. The streamline deterministic tractography algorithm was performed using the dODF results obtained with the aQBI model with a maximum SH order N set to 6. The Laplace-Beltrami regularization λ was set to 0.006.

tation time points out the requirement for a more rapid algorithme through an incremental version of the tractography technique, as well as a distribution and parallelization on our cluster of 80 CPUs.

8.3 Conclusion

This last RT tractography application showed that it was possible to obtain accurate tractography results from a certain iteration, depending on the maximum SH order. We filled our objective, which was to investigate a prototype for running a tractography algorithm at each iteration of the dMRI sequence. This is a preliminary work and it would be very interesting to further investigate the development of a true RT workflow for deterministic tractography, as well as for other classes of tractography, like the global one. This development would necessarily require a modification of the algorithm to make it incremental, so that the computation cost of the algorithm decreases, as it would rely on the previous iteration and it would only add the modification brought by the change in the dODF map due to the current iteration.

Finally, this application strengthened the usefulness of our RT noise correction method relying on the combination of the LMMSE and the KF embedded with a feedback loop. Indeed, our results on simulated and real data highlighted the improvement gained when the noise removal was performed.

8.3.1 Contributions of this chapter

- The investigation of an interest for a RT workflow dedicated to tractography results, using a simple prototype.
- The application of the LMMSE & KF noise correction RT scheme for tractography results. Similar results were presented in the submitted paper (under revision) [Brion et al.].

Part V

Conclusion

Chapter 9

Conclusion

In this thesis, we proposed innovative methodological tools dedicated to the recently introduced real-time (RT) dMRI workflow. Our methods were mainly focused on RT dMRI noise estimation and correction, with a final opening towards the feasibility of performing tractography during the ongoing scan. This thesis manuscript consisted of a first part entitled *State-of-the-art report on real-time diffusion magnetic resonance imaging (rtdMRI)* containing all the background necessary to understand the motivations behind the outcome of the RT dMRI workflow, as well as the mathematical framework on which this RT system relies. Then, we presented a second part entitled *Real-time noise correction for real-time dMRI (rtdMRI)* in which we first tackled the MRI noise field with a large state-of-the-art concerning the MRI noise modeling. Then, we explained our problematic and analyzed the literature about the offline denoising techniques used in MRI with regard to our RT constraints. We focused on a solution based on an LMMSE which we extended to deal with noncentral χ (nc-χ) noise correction. Furthermore, we developed three RT noise correction techniques accounting for our previous noise modeling analysis. Our first method consists of an embedding of the extended LMMSE with a Kalman filter and a feedback loop. In opposition to this first method applied on the magnitude data, our second RT noise removal technique is located before the sum of squares reconstruction, in the **k**-space, and applies on the **k**-space complex data, corrupted by an uncorrelated zero-mean Gaussian noise. This technique requires to modify the reconstruction pipeline of the manufacturer. Finally, we proposed a third RT solution applied on the magnitude data. This last technique relies on a parallel Kalman filter accounting for a nc-χ noise. We compared all RT noise correction techniques together on diffusion-weighted (DW) simulated and real data. Our last part *Real-time tractography application* focused on a further step to infer human brain connectivity in RT through a prototype for performing tractography online during the exam of the patient. This last application enabled to reinforce the usefulness of one of our RT noise correction technique on tractography maps. We hope that this novel possibility of performing tractography in RT may contribute to the improvement of clinical diagnosis of WM connectivity in situations where decisions have to be taken in emergency.

Contributions

In this thesis, we listed all our contributions at the end of the chapters presented in parts III and IV. We detail them again hereinafter, dividing them into minor and major contributions.

Minor contributions.

We investigated **a new noise estimation method** [Brion et al. (2011b)], extended from the technique introduced by [Rajan et al. (2010)] and **dedicated to data corrupted by nc-χ noise and which do not contain any background**. This contribution is new for the community, as there is, to our knowledge, no other noise variance estimation technique accounting for a nc-χ noise and which does not rely on the background knowledge. Our results presented on a simulated T_1-weighted image showed that this method was less robust than others relying on the background knowledge. This technique also assumes a stationary noise which is in practice not realistic when using multiple coils for the acquisition of the MRI signal. However, it can serve as a first approximation when handling data, for which no background is present.

We proposed **a new image quality index dedicated to dMRI images, for which neither a noise-free reference, nor a noise estimation is necessary** [Brion et al.]. Our index was shown to be accurate in comparison with other commonly used indices, when tested on simulated data in chapter 5. The advantage of this index over other classical quality indices, is that is can be used on raw or filtered images, only relying on the intensity of voxels in two regions of a dMRI map, like the GFA map. The measurement of this index, called the GFA ratio as we used it on GFA maps, requires to choose two regions of different mean intensity. Our experiments on real data throughout this thesis were quantified thanks to this index, which was shown to be close to our visual perception of the data, although it is a local indicator of the quality and is therefore dependent on the two regions chosen.

Major contributions.

We proposed **an extension of the LMMSE method introduced by** [Aja-Fernández et al. (2008b)] **to produce a nc-χ noise correction** [Brion et al. (2011a,b,c)]. This new filter which we called nc-χ LMMSE was shown to outperform the original LMMSE dedicated to Rician noise on simulated and real data corrupted by nc-χ noise. The nc-χ LMMSE presents the same advantages of rapidity and robustness than the original LMMSE and is therefore of interest for anyone working with noisy MRI data, acquired with multiple coils and required to be filtered with a fast technique.

For a better accuracy of this nc-χ LMMSE, we proposed **an empirical new solution to account for eventual noise correlations known to appear with parallel MRI acquisitions** [Brion et al.]. Our solution requires to test the nc-χ LMMSE with different values of the effective number of channels n_{eff}. Consequently, our technique is based

on a trial and can only be used in RT when handling a database with constant sequence parameters, after having tuned the n_{eff} parameter on a first subject of the database. Once the trial has been performed, the nc-χ LMMSE runs as fast as before, with an effective number of channels n_{eff} and an effective noise standard deviation σ_{eff} accounting for correlations. This technique was shown to produce accurate denoised results.

We developed **the first RT noise correction technique dedicated to the RT dMRI workflow** introduced by [Poupon et al. (2008b)]. This new technique [Brion et al. (2010, 2011d,e,f); Brion et al.], relying on the association of the extended LMMSE and a Kalman filter (KF), fully enables to correct each DW volume acquired along the iterations of a dMRI sequence, while its computation time remains below the repetition time when using a parallelization and a distribution on a cluster of 80 CPUs. Our new method was elaborated to exploit the RT dMRI results obtained at an iteration in injecting them through a feedback loop in the estimation of the noise-free DW volume at the next iteration. Therefore, our solution makes full use of its RT workflow. Our results demonstrated that our method was accurate on both simulated and real DW data on a panel of *b-values* yielding low to high signal levels. A clear improvement was depicted on typical aQBI maps like the dODF maps, and further on deterministic tractography results. Furthermore, the technique fully respected the RT constraints imposed by our problematic.

We developed a correction scheme based on **a Gaussian low-pass filter which is applied on the k space real and imaginary data** for which noise follows a zero-mean Gaussian distribution with a stationary noise variance across each image. This noise configuration allows a simple correction method, but requires its insertion in the manufacturer reconstruction pipeline. We inserted our method in the Siemens reconstruction pipeline and therefore it is specific to sequences performed with a Siemens MRI magnet. Our results showed accurate results, with a slightly less controlled smoothing effect, compared to the previous RT noise correction.

Finally, we developed a third RT noise correction relying on **a parallel Kalman filter** being applied on the magnitude data for which a nc-χ noise has to be accounted [Brion et al. (2012a,b,c)]. This last technique is based on a global Bayesian framework performing the nc-χ noise removal in approximating each non-Gaussian noise distribution by a Gaussian mixture. This last denoising technique was shown to be efficient on simulated data and to have a comparable computation time with our first LMMSE & KF technique.

Last, we investigated **the interest for performing a tractography RT workflow** in showing for the first time intermediate results, given by the streamline deterministic tractography algorithm, along iterations of a dMRI sequence. Our approach did not include any change of the algorithm, but assessed for the interesting new field of RT tractography executed at each new iteration of a dMRI sequence.

To conclude on these major contributions dedicated to rtdMRI, we believe that they are useful to improve the image quality for the dMRI RT workflow. We believe that this RT workflow could improve a lot the efficiency of dMRI exams and that this thesis work contributes to further encourage the enthusiasm to develop rtdMRI as a clinical tool for

more accurate results, faster treatment of the patient and less waste of data, as the quality of the data is assessed online.

Future prospects

We believe that the tools developed in this thesis and dedicated to the RT dMRI workflow will improve the quality of the images processed in RT during the dMRI sequence. We want to highlight that the LMMSE & KF correction technique will be integrated into the Connectomist software to be used by the community.

As future prospects, we would like to elaborate further methodological tools for several applications. A patent [Poupon et al. (2008a)] has been filed in 2008 with the objective to use rtdMRI for the development of an optimal set of diffusion gradients to make the diffusion measurement more informative concerning the underlying structure. A typical example for which the common uniform set of diffusion gradients is not optimal is the study of the spinal cord, for which the fibers are mostly centered around one main direction. A method for optimizing the sampling of the gradients has been proposed in [Caruyer and Deriche (2009)]. Such methodological improvements would be useful in clinical practice to obtain more accurate images with higher details on the microarchitecture of the WM fibers. The advantage brought by the RT workflow would be the possibility to tune the next gradient orientation online, with respect to the previous result.

As we already mentioned it in this manuscript, the RT workflow proposed for dMRI has a huge potential for clinical applications. It could be used when there is a need of an emergency diagnosis or for clinical non-emergency routine in pediatric studies. Besides, an ANR (Agence Nationale de la Recherche) research project called PEDIART, with Professor Chiron, Dr. Hertz-Pannier and Dr. Sévin, has been validated and funded to integrate and further develop the rtdMRI environment for pediatric imaging studies. This project contains a focus on epileptic children, for which the disease is drug resitant. The dMRI RT workflow would permit in such cases to better handle uncontrolled movements of the patients in assessing the image quality in RT. This would improve the quality of the exam and therefore accelerate the medical understanding of these epileptic drug resistant cases.

Further methodological investigations would be very interesting to perform, like extending automatic segmentation tools [Poupon (1999b); Marrakchi-Kacem (2011)] to be able to segment anatomical brain structures in RT. Such a workflow would accelerate any study of a particular anatomical structure in the brain. Concerning the study of WM fiber bundles, a technique enabling the clustering of bundles [Guevara Alvez (2011)] dedicated to an RT investigation during the dMRI sequence would accelerate a study on specific groups of bundles for example. Finally, recent techniques have been developed to propose an axonal calibration of WM fibers [Assaf et al. (2008)]. An RT extension of such a technique would enable to rapidly infer distributions of axons in the brain and also to improve specific measurements in using feedback of the RT results on the sequence parameters.

Part VI

Appendices

Appendix A

Acquisition of real human brain data

The algorithms developed in our work were applied on real human brain data acquired at NeuroSpin in the "Archi database" project directed by Dr. Cyril Poupon ([Schmitt et al. (2013)]). This project consisted of acquiring a database of healthy human brain MRI data with the best resolution achievable at 3T. Among the sequences used, many were diffusion weighted. Therefore, this database is very convenient for diffusion studies and was used in the frame of this work.

The subjects were mostly right-handed and between 20 to 40 years old. The exams consisted of a single-shot echo-planar DW spin echo pulse sequence. They were performed on a Magnetom Tim Trio $\boxed{3T}$ MRI system (Siemens Healthcare, Erlangen, Germany). Each individual dataset includes three T_2-weighted volumes acquired at $b = 0s \cdot mm^{-2}$ and 60 uniformly distributed orientations at $b = 1500s \cdot mm^{-2}$ and $b = 3000s \cdot mm^{-2}$ using the strategy of [Dubois et al. (2006)]: each acquisition was divided into 3 consecutive scans of 5 minutes each, to improve the comfort of the subject. Three uniform gradient orientation subsets were used, as explained in chapter 4, subsection 4.4.1. They consisted of three blocks of 19, 20 and 21 orientations each. Other acquisition parameters are listed in the table A.1. Finally, the acquisition was performed with the 12-element head coil available on the Tim Trio, for which the 12 coil elements are combined into 4 groups of 3 coil elments each. These groups are connected to 4 distinct receiver channels, yielding $\boxed{n = 4}$. To combine together the data received by the different channels, the SoS algorithm was used.

Among the several acquisitions for the "Archi database" project, one sequence was especially performed for our study with the same parameters as previously, but for four *b-values* instead of two, and namely: $\boxed{b = 1500s \cdot mm^{-2}}$, $\boxed{b = 3000s \cdot mm^{-2}}$, $\boxed{b = 4500s \cdot mm^{-2}}$ and $\boxed{b = 6000s \cdot mm^{-2}}$. The images obtained with this acquisition were used extensively for the results described in the manuscript. Indeed, as noise has a higher impact at high *b-values*, it is more interesting to show the results on this dataset.

Table A.1: Some "Archi database" acquisition parameters.

$T_E = 93ms$ for $b = 1500s \cdot mm^{-2}$ and $T_E = 117ms$ for $b = 3000s \cdot mm^{-2}$
$T_E = 135ms$ for $b = 4500s \cdot mm^{-2}$ and $T_E = 149ms$ for $b = 6000s \cdot mm^{-2}$
$T_R = 14s$
$FOV = 220mm$
matrix of 128×128
70 slices
resolution of $1.7 \times 1.7 \times 1.7mm^3$
phase partial Fourier 6/8
GRAPPA acceleration factor of 2
read bandwidth (RBW) of 1628 $Hz \cdot pixel^{-1}$
Fat saturation ON
Apodization filter ON

An apodization filter was used prior to the Fourier transformation. This filter forces the signal to zero at the end of the data collection period [Brown and Semelka (2010)]: it results that the central **k**-space data are more weighted than the peripheral **k**-space data. Such a process increases the SNR. The noise remains normally distributed before the SoS combination, but with a lower standard deviation [Dietrich et al. (2008a)]. Because of the zero-padding process, a voxel-to-voxel correlation appears: the real and imaginary noise components of the different channels are no longer uncorrelated [Brown and Semelka (2010); Liang and Lauterbur (1999)]. This may slightly change the noise PDF definition.

Concerning the small processing done immediately after the exam, some was done to correct the effects due to some artifacts using the Connectomist software [Duclap et al. (2012)]. The final T_2-weighted volume was averaged over the three volumes acquired, to improve its SNR [Haacke et al. (1999)]. Let us here highlight that this average changes the noise nature in the final T_2-weighted volume: the nc-χ distribution approximation may no longer be valid. That is why in our algorithms, we prefered to apply neither noise correction nor noise variance estimation algorithms on this final T_2-weighted volume. Moreover, as the quality of the T_2-weighted images of this average volume was visually very good from a noise point of view, it was justified to not perform any noise removal. This T_2-weighted volume was however corrected from the susceptibility artifact, which is due to the difference of susceptibility between side by side regions (for example at an interface between bone and tissue: at the interface, an intrisic magnetic field is created which induces a dephasing of the spins, creating the artifact). The DW volumes were also corrected from the susceptibility artifact, as well as from outliers (due to spikes or vibration effects). Furthermore, as the subjects were all healthy, they could voluntarily stay unmoving, and there was no big motion artifact. Nevertheless, to correct from small motion artifact, the Connectomist motion correction algorithm was applied. The correction of the

gradients' non-linearities were done directly during the reconstruction on the Siemens console and the distortions due to susceptibility effects were corrected using a fieldmap and the Connectomist software. The T_2-weighted image was threshold and improved using morphological operations to provide a rough mask of the brain. No noise correction or noise modification —like averaging— was performed on the DW data. With all this information we considered that all the noise analysis explained in chapters 4 and 5 is valid for these data.

Appendix B

Some mathematical functions used in this thesis

In this appendix, we list some mathematical functions used in this thesis.

• **The modified Bessel functions of the first kind**:

The modified Bessel function of the first kind of order n (always a postive integer in our case) is defined by:

$$I_n(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+n+1)} \left(\frac{x}{2}\right)^{2k+n}. \tag{B.1}$$

In the C++ code, we implemented this function using the following approximation:

$$\begin{cases} \text{If } 0 < x << \sqrt{n+1}, \text{ then} & I_n(x) \approx \frac{1}{\Gamma(n+1)} \left(\frac{x}{2}\right)^n, \\ \text{If } x >> n, \text{ then} & I_n(x) \approx \frac{e^x}{\sqrt{2\pi x}}, \end{cases} \tag{B.2}$$

with Γ the Gamma function.

• **The Gamma function**:

The Gamma function is defined by the following integral for a real x:

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt. \tag{B.3}$$

For a positive integer n, the Gamma function is defined by:

$$\Gamma(n) = (n-1)!, \tag{B.4}$$

with the following property: $\Gamma(n+\frac{1}{2}) = \frac{(2n)!}{4^n \cdot n!}\sqrt{\pi}$.

• **The Pochhammer rising factorial symbol**:

The Pochhammer rising factorial symbol is defined for a real x and a real a such that:

$$(x)^{(a)} = \frac{\Gamma(x+a)}{\Gamma(x)}, \tag{B.5}$$

with Γ the Gamma function.

• **The confluent hypergeometric function of the first kind**:

The confluent hypergeometric function of the first kind is defined, for a, b and z all reals, by:

$$\begin{aligned} {}_1F_1(a;b;z) =& 1 + \frac{a}{b}z + \frac{a(a+1)}{b(b+1)}\frac{z^2}{2!} + ... \\ =& \sum_{k=0}^{\infty} \frac{(a)^{(k)}}{(b)^{(k)}}\frac{z^k}{k!}, \end{aligned} \tag{B.6}$$

where $(a)^{(k)}$ and $(b)^{(k)}$ are Pochhammer rising factorial symbols, with k a positive integer. For high values of $|z|$ ($|z| > 10^5$), we can use the Kummer's transformation to compute this function: ${}_1F_1(a;b;z) = e^z {}_1F_1(b-a;b;-z)$.

Appendix C

Simulation of DW data

To estimate the efficiency of our noise estimation and correction methods, we simulated DW data. We first created noise-free DW data to serve as reference (ground truth). Then, we added nc-χ noise to these data to obtain the noisy DW data, on which we applied our algorithms.

C.1 Noise-free DW data simulation

To generate a DW signal we used, as in [Descoteaux (2008)], the hypothesis that the diffusion propagator for a bundle of fibers is Gaussian. It is the same hypothesis as the one used in the DTI model. We saw in chapter 3 that the MRI signal is then given by:

$$S(\mathbf{q}, \tau) = S_0 e^{-\tau \mathbf{q}^{\mathrm{T}} \mathbf{D} \mathbf{q}}. \tag{C.1}$$

As we have the following relationship between the *b-value* and the norm of the $\mathbf{q}$ vector, $b = \tau \|\mathbf{q}\|^2$, the MRI signal can be expressed such that:

$$S(b, \mathbf{o}_i) = S_0 e^{-b \mathbf{o}_i^T \mathbf{D} \mathbf{o}_i}, \tag{C.2}$$

with $\mathbf{o}_i$ being the unit orientation vector of the diffusion gradients. The set of the orientations designated by $\mathbf{o}_i$ for $i \in [\![1; K]\!]$ is simulated using an electrostatic repulsion model, such as the ones of [Jones et al. (1999); Papadakis et al. (2000)] described in chapter 4. In our simulations, we set $K = 500$.

Eq. C.2 can then be extended for a crossing of two (or more) fiber bundles. The MRI signal of such a crossing is then modeled by the addition of the MRI signals due to the separate bundles:

$$S(b, \mathbf{o}_i) = S_0 \sum_{k=1}^{2} f_k e^{-b \mathbf{o}_i^T \mathbf{D}_k \mathbf{o}_i}, \tag{C.3}$$

where f_k is the volumic fraction ($\sum_{k=1}^{2} f_k = 1$) and D_k is the diffusion tensor, both associated to the k^{th} fiber bundle. We simulated a crossing of two bundles of equivalent size, thus $f_1 = f_2 = 0.5$. This choice of a Gaussian mixture model means that we assume that

there is no water molecules' exchange between the two compartments. This assumption is acceptable as the exchange time between the two compartments is much longer than the diffusion time in the typical dMRI sequences.

We simulated three signal configurations: a signal due to isotropic diffusion, a signal from a fiber bundle and a signal from two bundles crossing at 60°. We set the eigenvalues of the diffusion tensor such that: $[\lambda_1, \lambda_2, \lambda_3] = [0.3 \times 10^{-9}, 0.3 \times 10^{-9}, 0.3 \times 10^{-9}]$ (in $m^2 \cdot s^{-1}$) for the isotropic tensor. For the anisotropic tensor, we set: $[\lambda_1, \lambda_2, \lambda_3] = [1.7 \times 10^{-9}, 0.3 \times 10^{-9}, 0.3 \times 10^{-9}]$ (in $m^2 \cdot s^{-1}$). These values are the same as in [Tuch (2002); Descoteaux (2008)]. Finally, we used these simulations with different *b-values*, such as: $b = 1500s \cdot mm^{-2}$, $b = 3000s \cdot mm^{-2}$, $b = 4500s \cdot mm^{-2}$ and $b = 6000s \cdot mm^{-2}$. The noise-free T_2-weighted signal S_0 was set to 200, assuming the same proton density and T_2 relaxation time through all the voxels.

We used this DW signal simulation method to generate a DW data field containing several voxels. We chose to represent the crossing of two fiber bundles across a 27×31 volume of voxels containing 27 slices. Outside of this crossing we chose to simulate isotropic diffusion. We also added a background to the simulated field. This background is useful to test our variance estimation methods described in chapter 5, section 5.2. The result of this noise-free DW simulated field is shown with the dODF representation in fig. C.1 at $b = 4500s \cdot mm^{-2}$. On this figure, the background was removed for rendering purpose (the background is visible at the bottom of the image in fig. C.2).

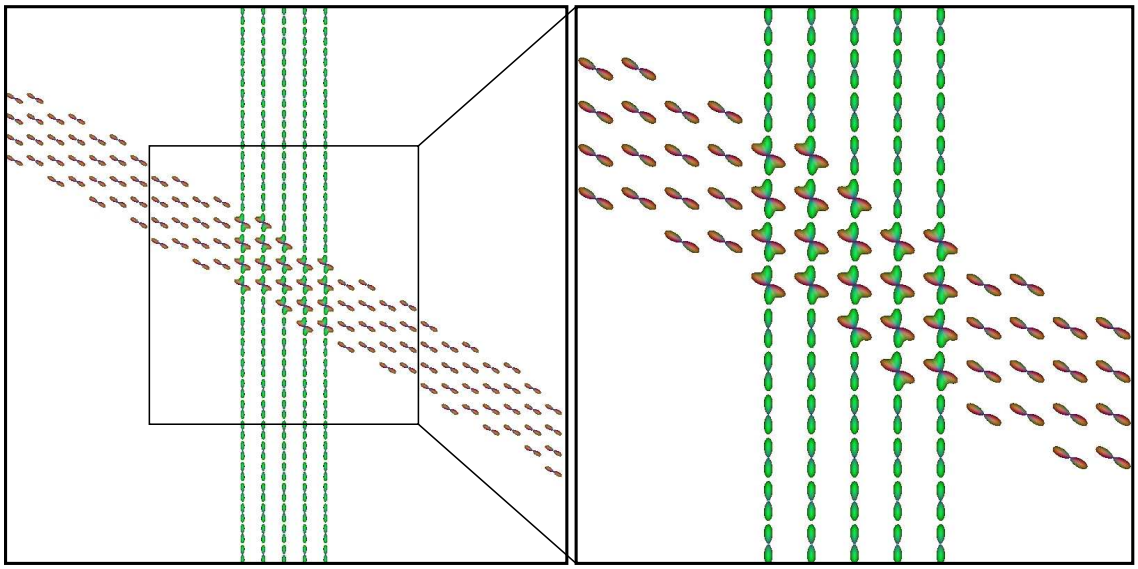


Figure C.1: Simulation of a noise-free DW field. Here the dODF map result is shown, with a zoom. Because of a normalization of each dODF by its maximal amplitude, the region outside the bundles with the isotropic tensors appears blank. Here, the simulated background is not shown.

C.2 Addition of nc-χ noise on the DW data

To test our noise estimation and filtering methods, we added a nc-χ noise using $n = 4$ channels corresponding to the standard case of the 12-element head coil antenna available on the Magnetom Tim Trio 3T MRI system (Siemens Healthcare, Erlangen, Germany) for which the 12 coil elements are combined into 4 groups of 3 coil elements each. To add the nc-χ noise with $n = 4$, we first calculated the noise-free signals $S(c)$ that each channel would receive, assuming that all channels receive the same signal: $\forall c \in [\![1; n]\!], S(c) = S/\sqrt{n}$, with S the noise-free signal given by eq. C.3. Then, the SoS recombined noisy signal M is given by the nonlinear transform:

$$M = \left(\sum_{c=1}^{n} \left(S(c) + \epsilon_r(c) \right)^2 + \epsilon_i(c)^2 \right)^{1/2}, \tag{C.4}$$

where $\epsilon_r(c)$ and $\epsilon_i(c)$ are the real and imaginary noises, respectively, on the channel c. They are generated using a Gaussian noise distribution of standard deviation σ.

Fig. C.2 shows the addition of nc-χ noise on the previous noise-free simulation at $b = 4500s \cdot mm^{-2}$. Here, we set σ to 16, yielding an SNR on the T_2-weighted image of $S_0/\sigma = 12.5$. For other values of quality indexes on this simulated field, see chapter 5, section 5.3.

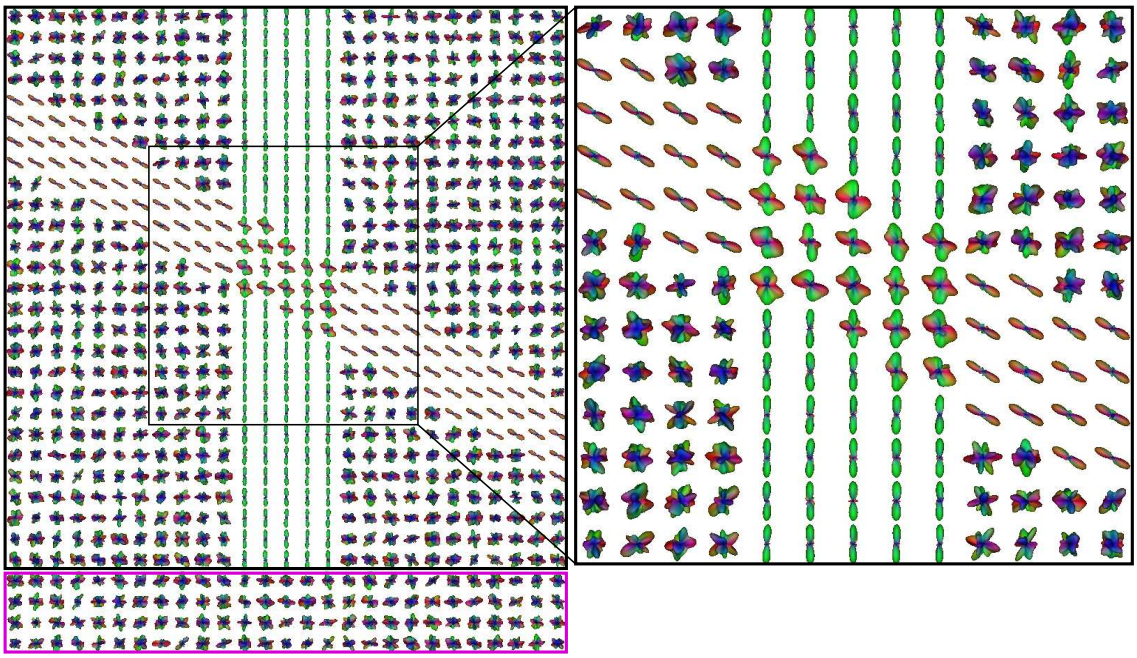


Figure C.2: Addition of nc-χ noise, with $\sigma = 16$ and $n = 4$, on the previous noise-free simulated DW field. The same zoom as previously is shown. Here, the simulated background is shown at the bottom in the magenta square.

Appendix D

Publications of the author arising from this work

Journal papers

1. V. Brion, C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. Le Bihan, F. Poupon. *Noise correction for HARDI and HYDI data obtained with multi-channel coils and Sum of Squares reconstruction: an anisotropic extension of the LMMSE*, Magnetic Resonance Imaging, In press.

2. S. Aja-Fernández, V. Brion, A. Tristán-Vega. *Effective Noise Estimation and Filtering from Correlated Multiple-Coil MR data*, Magnetic Resonance Imaging, Volume 31, Pages 272-285, 2013.

3. M. Iima, A. Yamamoto, V. Brion, T. Okada, M. Kanagaki, K. Togashi, D. Le Bihan. *Reduced-Distortion Diffusion MRI of the Craniovertebral Junction*, American Journal of Neuroradiology, Volume 33, Pages 1321-1325, 2012.

Conference papers

1. A. Tristán-Vega, V. Brion, G. Vegas-Sánchez-Ferrero, S. Aja-Fernández. *Merging squared-magnitude approaches to DWI denoising: An adaptive Wiener filter tuned to the anatomical contents of the image*, 35^{th} annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013.

2. V. Brion, O. Riff, M. Descoteaux, J.-F. Mangin, D. Le Bihan, C. Poupon, F. Poupon. *The Parallel Kalman Filter: an efficient tool to deal with real-time non central χ noise correction*, 9th IEEE International Symposium on Biomedical Imaging (ISBI), 2012.

3. V. Brion, C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. Le Bihan, F. Poupon. *Parallel MRI noise correction: an extension of the LMMSE to noncentral χ distributions*, Medical Image Computing and Computer Assisted Intervention (MICCAI), 2011.

4. V. Brion, I. Kezele, O. Riff, M. Descoteaux, J.-F. Mangin, C. Poupon, F. Poupon. *Real-time Rician noise correction applied to real-time HARDI and HYDI*, In MICCAI 2010, Workshop on Computational Diffusion MRI (CDMRI), 2010.

Conference abstracts

1. D. Duclap, A. Lebois, B. Schmitt, O. Riff, P. Guevara, L. Marrakchi-Kacem, V. Brion, F. Poupon, J.-F. Mangin, C. Poupon. *Connectomist-2.0: a novel diffusion analysis toolbox for BrainVISA*, 29^{th} annual meeting of the European Society for Magnetic Resonance in medicine and Biology (ESMRMB) 2012.

2. V. Brion, O. Riff, M. Descoteaux, J.-F. Mangin, D. Le Bihan, C. Poupon, F. Poupon. *The Parallel Kalman Filter : an efficient tool to deal with real-time non central χ noise correction*, 20^{th} annual meeting of the International Society for Magnetic Resonance in Medicine (ISMRM) 2012.

3. V. Brion, O. Riff, M. Descoteaux, J.-F. Mangin, D. Le Bihan, C. Poupon, F. Poupon. *The Parallel Kalman Filter : an efficient tool to deal with real-time non central χ noise correction*, annual meeting of the Société Française de résonance magnétique en biologie et médecine (SFRMBM) 2012.

4. V. Brion, O. Riff, J.-F. Mangin, D. Le Bihan, C. Poupon, F. Poupon. *Correction en temps-réel du bruit χ non centré pour l'IRM parallèle de diffusion en temps réel*, Journée thématique du Groupement de Recherche (GDR) transversal Imageries in vivo (imagiv) 2011.

5. V. Brion, C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. Le Bihan, F. Poupon. *Correction du bruit pour l'IRM parallèle: une extension du LMMSE aux distributions χ non centrées*, GDR imagiv 2011.

6. V. Brion, O. Riff, J.-F. Mangin, D. Le Bihan, C. Poupon, F. Poupon. *Real-time non central Chi noise correction adapted to parallel real-time diffusion-weighted MRI: impact on real-time HARDI maps*, 28^{th} annual meeting of the ESMRMB 2011. Certificate of Merit for the electronic poster.

7. V. Brion, C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. Le Bihan, F. Poupon. *Parallel MRI noise correction: an extension of the LMMSE to non central Chi distributions*, 28^{th} annual meeting of the ESMRMB 2011.

8. V. Brion, O. Riff, I. Kezele, M. Descoteaux, D. Le Bihan, J.-F. Mangin, C. Poupon, F. Poupon. *Real-time Rician noise correction applied to real-time HARDI and HYDI*, 19^{th} annual meeting of the ISMRM 2011.

9. C. Poupon, J. Dubois, L. Marrakchi, V. Brion, J.-F. Mangin, F. Poupon. *Real-Time EPI T1, T2 and T2* Mapping at 3T*, 18^{th} annual meeting of the ISMRM 2010.

10. V. Brion, I. Kezele, M. Descoteaux, J.-F. Mangin, C. Poupon. *Rician denoising dedicated to single-shell diffusion-weighted MR data using spherical harmonics: impact on fibre orientation distribution maps*, 26^{th} annual meeting of the ESMRMB 2009.

Bibliography

In this bibliography, the red numbers correspond to the pages of the manuscript where the reference is cited.

Aganj, I., C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel: 2010, 'Reconstruction of the Orientation Distribution Function in Single- and Multiple-Shell q-Ball Imaging Within Constant Solid Angle'. *Magnetic Resonance in Medicine* **64**, 554–566. xxxiv, xxxvi, 39, 51, 52

Ahn, C. B., Y. C. Song, and D. J. Park: 1999, 'Adaptive template filtering for signal-to-noise ratio enhancement in magnetic resonance imaging'. *IEEE Transactions on Medical Imaging* **18**, 549–556. 140

Aja-Fernández, S., C. Alberola-López, and C.-F. Westin: 2008a, 'Noise and Signal Estimation in Magnitude MRI and Rician Distributed Images : A LMMSE Approach'. *IEEE transactions on image processing* **17**(8), 1383–1398. lii, 141, 145, 146, 147, 148, 150

Aja-Fernández, S., V. Brion, and A. Tristán-Vega: 2013, 'Effective noise estimation and filtering from correlated multiple-coil MR data'. *Magnetic Resonance Imaging* **31**, 272–285. xvii, xlix, liv, 125, 126, 130, 131, 141, 144, 146, 150, 158, 163, 165, 166, 174, 208, 224

Aja-Fernández, S., R. S. J. Estépar, C. Alberola-López, and C.-F. Westin: 2006, 'Image Quality Assessment based on Local Variance'. In: *IEEE EMBS Annual International Conference*. 131

Aja-Fernández, S., M. Niethammer, M. Kubicki, M. E. Shenton, and C.-F. Westin: 2008b, 'Restoration of DWI Data Using a Rician LMMSE Estimator'. *IEEE Transactions on Medical Imaging* **27**, 1389–1403. lii, 141, 145, 146, 148, 149, 150, 152, 176, 197, 244

Aja-Fernández, S. and A. Tristán-Vega: 2012, 'Influence of Noise Correlation in Multiple-Coil Statistical Models With Sum of Squares Reconstruction'. *Magnetic Resonance in Medicine* **67**, 580–585. xlvii, 111, 117, 125, 151, 158, 224

Aja-Fernández, S., A. Tristán-Vega, and C. Alberola-López: 2009, 'Noise estimation in single- and multiple-coil magnetic resonance data based on statistical models'. *Magnetic Resonance Imaging* **27**, 1397–1409. xlix, 123, 128, 129

Aja-Fernández, S., A. Tristán-Vega, and W. S. Hoge: 2011, 'Statistical Noise Analysis in GRAPPA Using a Parametrized Noncentral Chi Approximation Model'. *Magnetic Resonance in Medicine* **65**, 1195–1206. xlvii, 111, 117, 118, 144, 151, 158, 224

Aksoy, M., C. Forman, M. Straka, S. Skare, and S. Holdsworth: 2011, 'Real-time optical motion correction for diffusion tensor imaging'. *Magnetic Resonance in Medicine* **66**, 366–378. 63

Alexander, A. L., Y. C. Wu, and P. C. Venkat: 2006, 'Hybrid Diffusion Imaging (HYDI).'. In: *Proc of IEEE Eng Med Biol Soc*. 42

Alspach, D. L. and H. Sorenson: 1972, 'Nonlinear Bayesian estimation using Gaussian sum approximations'. *IEEE Transactions on Automatic Control* **17**, 439–448. 65, 204

Alvarez, L.: 1996, 'Images and PDE's'. *Lecture Notes in Control and Information Sciences* **219**, 3–14. 139

Anderson, A. W. and Z. Ding: 2002, 'Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging'. In: *ISMRM*. xxxvi, 53

Anderson, B. D. O. and J. B. Moore: 1979, *Optimal Filtering*. Prentice-Hall. 64

Andersson, J. L.: 2008, 'Maximum a posteriori estimation of diffusion tensor parameters using a Rician noise model: why, how and but'. *NeuroImage* **42**, 1340–1356. 142

Ashburner, J. and K. J. Friston: 2000, 'Voxel-base morphometry - the methods'. *NeuroImage* **11**, 805–821. 139

Assaf, Y. and P. J. Basser: 2005, 'Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain'. *NeuroImage* **27**, 48–58. 41

Assaf, Y., T. Blumenfeld-Katzir, Y. Yovel, and P. J. Basser: 2008, 'AxCaliber: a method for measuring axon diameter distribution from diffusion MRI'. *Magnetic Resonance in Medicine* **59**, 1347–1354. lxvii, 246

Assemlal, H.-E., D. Tschumperlé, and L. Brun: 2007, 'Fiber Tracking on HARDI Data Using Robust ODF Fields'. In: *ICIP*. pp. 133–136. lii, 140, 145

Assemlal, H.-E., D. Tschumperlé, and L. Brun: 2009, 'Efficient and robust computation of PDF features from diffusion MR signal'. *Medical Image Analysis* **13**, 715–729. 42

Assemlal, H.-E., D. Tschumperlé, L. Brun, and K. Siddiqi: 2011, 'Recent Advances in Diffusion MRI Modeling: Angular and Radial Reconstruction'. *Medical Image Analysis* **15**, 369–396. 41

Awate, S. P. and R. T. Whitaker: 2005, 'Nonparametric Neighborhood Statistics for MRI Denoising'. *Information Processing in Medical Imaging* **19**, 677–688. 143, 144

Awate, S. P. and R. T. Whitaker: 2006, 'Unsupervised , Information-Theoretic, Adaptive Image Filtering for Image Restoration'. *IEEE Transactions on pattern analysis and machine intelligence* **28**, 364–376. 140, 143

Awate, S. P. and R. T. Whitaker: 2007, 'Feature-Preserving MRI Denoising: A Nonparametric Empirical Bayes Approach'. *IEEE Transactions on Medical Imaging* **26**, 1242–1255. 139, 142, 144

Barbier, E.: 2003-2004, 'Traitement du signal en RMN'. cours du Master 2 Recherche "Physique pour l'instrumentation". 105

Basser, P. J.: 1995, 'Inferring microstructural features and the physiological state of tissues from diffusion-weighted images'. *NMR in Biomedicine* **8**, 333–344. 43

Basser, P. J., J. Mattiello, and D. LeBihan: 1994, 'Estimation of the effective self-diffusion tensor from the NMR spin echo'. *Journal of Magnetic Resonance B* **103**, 247–254. xxxv, 40, 43

Basser, P. J., S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi: 2000, 'In vivo fiber tractography using DT-MRI data'. *Magnetic Resonance in Medicine* **44**, 625–632. 54

Basu, S., T. Fletcher, and R. Whitaker: 2006, 'Rician Noise Removal in Diffusion Tensor MRI'. In: *MICCAI*. pp. 117–125. 139, 142

Behrens, T., H. J. Berg, S. Jbabdi, M. Rushworth, and M. Woolrich: 2007, 'Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?'. *NeuroImage* **34**, 144–155. 57

Behrens, T., M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith: 2003, 'Characterization and propagation of uncertainty in diffusion-weighted MR imaging'. *Magnetic Resonance in Medicine* **50**, 1077–1088. 41, 57

Bergmann, O., G. Kindlmann, S. Peled, and C.-F. Westin: 2007, 'Two-tensor fiber tractography'. In: *ISBI*. 55

Berman, J. I., S. Chung, P. Mukherjee, C. P. Hess, E. T. Han, and R. G. Henry: 2008, 'Probabilistic streamline q-ball tractography using the residual bootstrap'. *NeuroImage* **39**, 215–222. 56

Bernstein, M. A., D. M. Thomasson, and W. H. Perman: 1989, 'Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction'. *Medical Physics* **16**, 813–817. xxxix, xlvi, 110

Birbaumer, N., N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor: 1999, 'A spelling device for the paralyzed'. *Nature* **398**, 297–298. xxiii, 3

Black, M. J., G. Sapiro, D. H. Marimont, and D. Heeger: 1998, 'Robust anisotropic diffusion'. *IEEE Transaction on Image Processing* **7**, 421–432. 139

Blaimer, M., F. Breuer, M. Mueller, R. M. Heidemann, M. A. Griswold, and P. M. Jakob: 2004, 'SMASH, SENSE, PILS, GRAPPA: How to Choose the Optimal Method'. *Top Magnetic Resonance Imaging* **15**, 223–236. 32

Bloch, F.: 1946, 'Nuclear Induction'. *Physical Review* **70**, 460–474. xxvi, 16

Breuer, F. A., S. A. R. Kannengiesser, M. Blaimer, N. Seiberlich, P. M. Jakob, and M. A. Griswold: 2009, 'General Formulation for Quantitative G-factor Calculation in GRAPPA Reconstructions'. *Magnetic Resonance in Medicine* **62**, 739–746. 118

Brion, V., I. Kezele, M. Descoteaux, J.-F. Mangin, and C. Poupon: 2009, 'Rician denoising dedicated to single-shell diffusion-weighted MR data using spherical harmonics: impact on fibre orientation distribution maps'. In: *ESMRMB*. lii, 142, 145

Brion, V., I. Kezele, O. Riff, M. Descoteaux, J.-F. Mangin, C. Poupon, and F. Poupon: 2010, 'Real-time Rician noise correction applied to real-time HARDI and HYDI'. In: *Workshop CDMRI of MICCAI*. pp. 2–13. lxvi, 145, 173, 224, 245

Brion, V., C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. LeBihan, and F. Poupon, 'Parallel MRI noise correction for HARDI and HYDI MRI data obtained with a Sum of Squares reconstruction: an anisotropic extension of the LMMSE'. *Magnetic Resonance Imaging* **under revision**. lxv, lxvi, 240, 244, 245

Brion, V., C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. LeBihan, and F. Poupon: 2011a, 'Correction du bruit pour l'IRM parallèle: une extension du LMMSE aux distributions Chi non centrées.'. In: *GDR Imagiv*. lii, lxvi, 149, 165, 244

Brion, V., C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. LeBihan, and F. Poupon: 2011b, 'Parallel MRI noise correction: an extension of the LMMSE to non central Chi distributions'. In: LNCS (ed.): *MICCAI*, Vol. 14(pt2). pp. 226–233. l, lii, lxv, lxvi, 127, 135, 141, 149, 165, 244

Brion, V., C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J.-F. Mangin, D. LeBihan, and F. Poupon: 2011c, 'Parallel MRI noise correction: an extension of the LMMSE to non central Chi distributions'. In: *ESMRMB*. lii, lxvi, 149, 165, 244

Brion, V., O. Riff, M. Descoteaux, J.-F. Mangin, D. LeBihan, C. Poupon, and F. Poupon: 2012a, 'The Parallel Kalman Filter: an efficient tool to deal with real-time non central Chi noise correction'. In: *ISBI*. lxvi, 145, 204, 224, 245

Brion, V., O. Riff, M. Descoteaux, J.-F. Mangin, D. LeBihan, C. Poupon, and F. Poupon: 2012b, 'The Parallel Kalman Filter: an efficient tool to deal with real-time non central Chi noise correction'. In: *ISMRM*. lxvi, 204, 224, 245

Brion, V., O. Riff, M. Descoteaux, J.-F. Mangin, D. LeBihan, C. Poupon, and F. Poupon: 2012c, 'The Parallel Kalman Filter: an efficient tool to deal with real-time non central Chi noise correction'. In: *SFRMBM*. lxvi, 204, 224, 245

Brion, V., O. Riff, I. Kezele, M. Descoteaux, D. LeBihan, J.-F. Mangin, C. Poupon, and F. Poupon: 2011d, 'Real-time Rician noise correction applied to real-time HARDI and HYDI'. In: *ISMRM*. lxvi, 145, 173, 224, 245

Brion, V., O. Riff, J.-F. Mangin, D. LeBihan, C. Poupon, and F. Poupon: 2011e, 'Correction en temps réel du bruit Chi non centré pour l'IRM parallèle de diffusion en temps réel.'. In: *GDR Imagiv*. lxvi, 173, 224, 245

Brion, V., O. Riff, J.-F. Mangin, D. LeBihan, C. Poupon, and F. Poupon: 2011f, 'Real-time non central Chi noise correction adapted to parallel real-time diffusion-weighted MRI: impact on real-time HARDI maps'. In: *ESMRMB*. lxvi, 141, 145, 173, 224, 245

Brown, M. A. and R. C. Semelka: 2010, *MRI: Basic Principles and Applications*. Wiley-Blackwell. 250

Brown, R.: 1828, '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'. *Philosophical Magazine* **4**, 161–173. xxxii, 34

Brown, R., Y. Wang, P. Spincemaille, and R. F. Lee: 2007, 'On the Noise Correlation Matrix for Multiple Radio Frequency Coils'. *Magnetic Resonance in Medicine* **58**, 218–224. xlvi, 108

Buades, A., B. Coll, and J. M. Morel: 2005, 'A review of image denoising algorithms, with a new one'. *SIAM Journal on Multiscale Modeling and Simulation* **4**, 490–530. 140

Bucy, R. S.: 1969, 'Bayes theorem and digital realization for nonlinear filters'. *Journal of Astronautical Sciences* **17**, 80–94. 64

Bucy, R. S. and K. D. Senne: 1971, 'Digital synthesis of non-linear filters'. *Automatica* **7**, 287–298. 64

Caan, M. W. A., G. Khedoe, D. Poot, A. den Dekker, S. Olabarriaga, K. Grimbergen, L. van Vliet, and F. Vos: 2010, 'Adaptive Noise Filtering for Accurate and Precise Diffusion Estimation in Fiber Crossings'. In: T. Jiang, N. Navab, J. P. W. Pluim, and M. A. Viergever (eds.): *MICCAI 2010, Part I, LNCS*, Vol. 6361. pp. 167–174, Springer, Heidelberg. lii, 141, 145

Callaghan, P. T.: 1991, *Principles of nuclear magnetic resonance microscopy*. Oxford University Press. xxxv, 40

Callaghan, P. T., C. D. Eccles, and Y. Xia: 1988, 'NMR microscopy of dynamic displacements: k-space and q-space imaging'. *Journal of Physics E: Scientific Instruments* **21**, 820–822. 41

Carlson, J. W. and T. Minemura: 1993, 'Imaging time reduction through multiple receiver coil data acquisition and image reconstruction'. *Magnetic Resonance in Medicine* **29**, 681–687. 30

Carr, H. Y. and E. M. Purcell: 1954, 'Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments'. *Physical Review* **94**, 630–638. xxx, 23

Caruyer, E., I. Aganj, C. Lenglet, G. Sapiro, and R. Deriche: 2011, 'Online motion detection in high angular resolution diffusion imaging'. In: *ISBI*. 63, 77, 78

Caruyer, E., I. Aganj, R. L. Muetzel, C. Lenglet, G. Sapiro, and R. Deriche: 2010, 'Online Orientation Distribution Function Reconstruction in Constant Solid Angle and its Application to Motion Detection in HARDI'. In: *ISBI*. xli, 63, 76

Caruyer, E. and R. Deriche: 2009, 'Adaptive Design of Sampling Directions in Diffusion Tensor MRI and Validation on Human Brain Images'. In: *Workshop on Diffusion Modelling and the Fiber Cup of MICCAI*. 246

Casaseca-de-la-Higuera, P., A. Tristán-Vega, S. Aja-Fernández, C. Alberola-López, C.-F. Westin, and R. S. J. Estépar: 2012, 'Optimal real-time estimation in diffusion tensor imaging'. *Magnetic Resonance Imaging* **30**, 506–517. 72, 73, 141, 145, 176, 177

Catté, F., L. P. Lions, J. M. Morel, and T. Coll: 1992, 'Image Selective Smoothing and Edge Detection by Nonlinear Diffusion I'. *SIAM Journal on Numerical Analysis* **29**, 182–193. 139

Chamberland, M., D. Fortin, and M. Descoteaux: 2012, 'Real-time fiber tractography: interactive parameter tuning for neurosurgical interventions'. In: *HBM*. 231

Chang, L.-C., D. K. Jones, and C. Pierpaoli: 2005, 'RESTORE: Robust Estimation of Tensors by Outlier Rejection'. *Magnetic Resonance in Medicine* **53**, 1088–1095. lii, 139, 145

Chao, Y.-P., J.-H. Chen, K.-H. Cho, C.-H. Yeh, K.-H. Chou, and C.-P. Lin: 2008, 'A multiple streamline approach to high angular resolution diffusion tractography'. *Medical Engineering and Physics* **30**, 989–996. 55, 56

Chefd'Hotel, C., D. Tschumperlé, R. Deriche, and O. Faugeras: 2002, 'Constrained Flows of Matrix-Valued Functions: Application to Diffusion Tensor Regularization'. In: L. N. in Computer Science (ed.): *ECCV*, Vol. 2350. pp. 251–265. lii, 140, 145

Chen, B. and E. W. Hsu: 2005, 'Noise removal in magnetic resonance diffusion tensor imaging'. *Magnetic Resonance in Medicine* **54**, 393–407. 139

Chen, R. and J. S. Liu: 2000, 'Mixture Kalman filters'. *Journal of the Royal Statistical Society B* **62**, 493–508. 65

Chen, Z.: 2003, 'Bayesian filtering: From Kalman Filters to Particle Filters, and Beyond'. Technical report, Technical Report McMasters University, Hamilton, ON. 65

Chui, C. K. and G. Chen: 1987, *Kalman Filtering with Real-Time Applications*. Springer-Verlag. 61, 67, 97

Clarke, R. A., P. Scifo, G. Rizzo, F. D. Acqua, G. Scotti, and F. Fazio: 2008, 'Noise Correction on Rician Distributed Data for Fibre Orientation Estimators'. *IEEE Transactions on Medical Imaging* **27**, 1242–1251. lii, 120, 142, 145

Cointepas, Y., C. Poupon, and J.-F. Mangin: 2002, 'A spin glass based framework to reconstruct brain fiber bundles from images of the water diffusion process'. *Information Processes* **2**, 30–36. 57

Collins, D., A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans: 1998, 'Design and Construction of a Realistic Digital Brain Phantom'. *IEEE Transactions on Medical Imaging* **17**, 463–468. 128

Constantinides, C. D., E. Atalar, and E. R. McVeigh: 1997, 'Signal-to-Noise Measurements in Magnitude Images from NMR Phased Arrays'. *Magnetic Resonance in Medicine* **38**, 852–857. xxxix, xlvi, xlvii, 69, 108, 111, 114, 131, 144

Conturo, T. E., N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle: 1999, 'Tracking neuronal fiber pathways in the living human brain'. *Proceedings of the National Academy of Sciences* **96**, 10422–10427. 54

Cook, P. A., M. Symms, P. A. Boulby, and D. C. Alexander: 2007, 'Optimal acquisition orders of diffusion-weighted MRI measurements'. *Journal of Magnetic Resonance Imaging* **25**, 1051–1058. 79

Coupé, P., P. Hellier, S. Prima, C. Kervrann, and C. Barillot: 2008a, '3D Wavelet Subbands Mixing for Image Denoising'. *International Journal of Biomedical Imaging* **2008**, 1–11. 141

Coupé, P., J. V. Manjón, E. Gedamu, D. Arnold, M. Robles, and D. L. Collins: 2010, 'Robust Rician noise estimation for MR images'. *Medical Image Analysis* **14**, 483–493. l, 127

Coupé, P., P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot: 2008b, 'An Optimized Blockwise Nonlocal Means Denoising Filter for 3-D Magnetic Resonance Images'. *IEEE Transactions on Medical Imaging* **27**, 425–441. 140, 146, 164

Cox, R. W., A. Jesmanowicz, and J. S. Hyde: 1995, 'Real-time functional magnetic resonance imaging'. *Magnetic Resonance in Medicine* **33**, 230–236. xxiii, 3

deCharms, R. C.: 2007, 'Reading and controlling human brain activation using real-time functional magnetic resonance imaging'. *TRENDS in Cognitive Sciences* **11**(11), 473–481. xxiii, 3

deCharms, R. C.: 2008, 'Applications of real-time fMRI'. *Nature Reviews Neuroscience* **9**, 720–729. xxiii, 3

Delakis, I., O. Hammad, and R. I. Kitney: 2007, 'Wavelet-based de-noising algorithm for images acquired with parallel magnetic resonance imaging (MRI)'. *Physics in Medicine and Biology* **52**, 3741–3751. 138, 139

Deriche, R., J. Calder, and M. Descoteaux: 2009, 'Optimal Real-Time Q-Ball Imaging using Regularized Kalman Filtering with Incremental Orientation Sets'. *Medical Image Analysis* **13**, 564–579. xl, 61, 67, 75, 76, 78, 79, 89, 98, 100, 101, 210

Descoteaux, M.: 2008, 'High Angular Resolution Diffusion MRI: from Local Estimation to Segmentation and Tractography'. Ph.D. thesis, Université de Nice-Sophia Antipolis. 33, 40, 41, 43, 48, 53, 54, 59, 60, 152, 175, 255, 256

Descoteaux, M., E. Angelino, S. Fitzgibbons, and R. Deriche: 2007, 'Regularized, Fast, and Robust Analytical Q-Ball Imaging'. *Magnetic Resonance in Medicine* **58**, 497–510. xxxvi, 45, 47, 48

Descoteaux, M., R. Deriche, T. R. Knösche, and A. Anwander: 2009, 'Deterministic and probabilistic tractography based on complex fibre orientation distributions'. *IEEE Transactions on Medical Imaging* **28**, 269–286. 55, 56

Descoteaux, M., R. Deriche, D. LeBihan, J.-F. Mangin, and C. Poupon: 2010, 'Multiple q-shell diffusion propagator imaging'. *Medical Image Analysis* **15**, 603–621. 42

Descoteaux, M., N. Wiest-Daesslé, S. Prima, C. Barillot, and R. Deriche: 2008, 'Impact of Rician Adapted Non-Local Means Filtering on HARDI'. In: L. N. in Computer Science (ed.): *MICCAI*. pp. 122–130. 140

Dietrich, O., J. G. Raya, S. B. Reeder, M. Ingrisch, M. F. Reiser, and S. O. Schoenberg: 2008a, 'Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics.'. *Magnetic Resonance Imaging* **26**, 754–762. xlvii, 111, 116, 117, 118, 250

Dietrich, O., J. G. Raya, S. B. Reeder, M. F. Reiser, and S. O. Schoenberg: 2007, 'Measurement of Signal-to-Noise Ratios in MR Images: Influence of Multichannel Coils, Parallel Imaging, and Reconstruction Filters'. *Journal of Magnetic Resonance Imaging* **26**, 375–385. 131, 132

Dietrich, O., J. G. Raya, and M. F. Reiser: 2008b, 'Magnetic Resonance Noise Measurements and Signal-Quantization Effects at Very Low Noise Levels'. *Magnetic Resonance in Medicine* **60**, 1477–1487. 106

Donoho, D. L. and I. M. Johnstone: 1994, 'Ideal Spatial Adaptation by Wavelet Shrinkage'. *Biometrika* **81**, 425–455. 138

Doucet, A., S. J. Godsill, and C. Andrieu: 2000, 'On sequential Monte Carlo sampling methods for Bayesian filtering'. *Statistics and Computing* **10**, 197–208. 65

Dubois, J., C. Poupon, F. Lethimonnier, and D. LeBihan: 2006, 'Optimized diffusion gradient orientation schemes for corrupted clinical DTI data sets'. *Magnetic Resonance Materials in Physics, Biology and Medicine* **19**, 134–143. xli, 79, 80, 81, 84, 249

Duclap, D., A. Lebois, B. Schmitt, O. Riff, P. Guevara, L. Marrakchi-Kacem, V. Brion, F. Poupon, J.-F. Mangin, and C. Poupon: 2012, 'Connectomist-2.0: a novel diffusion analysis toolbox for BrainVISA'. In: *ESMRMB*. 250

Edelstein, W. A., P. A. Bottomley, and L. M. Pfeifer: 1984, 'A signal-to-noise calibration procedure for NMR imaging systems'. *Medical Physics* **11**, 180–185. 109

Edelstein, W. A., G. H. Glover, C. J. Hardy, and R. W. Redington: 1986, 'The Intrinsic Signal-to-Noise Ratio in NMR Imaging'. *Magnetic Resonance in Medicine* **3**, 604–618. 107

Einstein, A.: 1905, 'Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen'. *Annalen der Physik* **322**, 549–560. 35

Fick, A.: 1855, 'On liquid diffusion'. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **10**, 30–39. 34

Fillard, P., X. Pennec, V. Arsigny, and N. Ayache: 2007, 'Clinical DT-MRI Estimation, Smoothing and Fiber Tracking with Log-Euclidean Metrics'. *IEEE Transactions on Medical Imaging* **26**, 1472–1482. lii, 142, 145

Fillard, P., C. Poupon, and J.-F. Mangin: 2009, 'A novel global tractography algorithm based on an adptive spin glass model'. In: *MICCAI*. 57

Frank, L. R.: 2002, 'Characterization of Anisotropy in High Angular Resolution Diffusion-Weighted MRI'. *Magnetic Resonance in Medicine* **47**, 1083–1099. xxxvi, 45, 47

Friman, O., G. Farnebäck, and C.-F. Westin: 2006, 'A Bayesian approach for stochastic white matter tractography'. *IEEE Transactions on Medical Imaging* **25**, 965–978. 57

Gauss, K. F.: 1809, *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium*. Boston: Little, Brown and company, Boston, MA, USA. English translation: Theory of the motion of the heavenly bodies moving about the sun in conic sections. 63

Gelb, A., J. F. Kasper, R. A. Nash, C. F. Price, and A. A. Sutherland: 1974, *Applied Optimal Estimation*. Technical Staff - The Analytic Sciences Corporation. 64

Gerig, G., O. Kubler, R. Kikinis, and F. A. Jolesz: 1992, 'Nonlinear anisotropic filtering of MRI data'. *IEEE Transactions on Medical Imaging* **11**, 221–232. 139

Goerner, F. L. and G. D. Clarke: 2011, 'Measuring signal-to-noise ratio in partially parallel imaging MRI'. *Medical Physics* **38**, 5049–5057. 131

Golby, A. J., G. Kindlmann, I. Norton, A. Yarmarkovich, S. Pieper, and R. Kikinis: 2011, 'Interactive Diffusion Tensor Tractography Vizualization for Neurosurgical Planning'. *Neurosurgery* **68**, 496–505. 231

Gordon, N. J., D. J. Salmond, and A. F. M. Smith: 1993, 'A novel approach to nonlinear/non-Gaussian Bayesian state estimation'. *IEE Proceedings - Part F: Radar and Sognal Processing* **140**, 107–113. 64

Greer, J. B. and A. L. Bertozzi: 2004, 'Traveling wave solutions of fourth order PDEs for image processing'. *SIAM Journal on Mathematical Analysis* **36**, 38–68. 140

Griswold, M. A., P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase: 2002, 'Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)'. *Magnetic Resonance in Medicine* **47**, 1202–1210. 30

Gudbjartsson, H. and S. Patz: 1995, 'The Rician Distribution of Noisy MRI Data'. *Magnetic Resonance Imaging* **34**, 910–914. 108, 112, 141

Guevara, P., D. Duclap, L. Marrakchi-Kacem, D. Rivière, Y. Cointepas, C. Poupon, and J.-F. Mangin: 2011, 'Accurate tractography propagation mask using T1-weighted data rather than FA'. In: *ISMRM*. 55

Guevara Alvez, P. B.: 2011, 'Inference of a human brain fiber bundle atlas from high angular resiolution diffusion imaging'. Ph.D. thesis, Université Paris-Sud XI. lxvii, 246

Guleryuz, O. G.: 2003, 'Weighted overcomplete denoising'. In: *Proceedings of the Asilomar Conference on Signals and Systems*. 139

Guleryuz, O. G.: 2007, 'Weighted averaging for denoising with overcomplete dictionaries'. *IEEE Transactions on Image Processing* **16**, 3020–3034. 139

Guo, L., Y. Wu, X. Liu, Y. Li, G. Xu, and W. Yan: 2006, 'Threshold Optimization of Adaptive Template Filtering for MRI Based on Intelligent Optimization Algorithm'. In: *Proceedings of the 28th IEEE EMBS Annual International Conference*. 140

Haacke, E. M., R. W. Brown, M. R. Thompson, and R. Venkatesan: 1999, *Magnetic resonance imaging: Physical Principles and Sequence Design*. Wiley-Liss. xxx, 16, 26, 107, 132, 250

Hahn, E. L.: 1950, 'Spin Echoes'. *Physical Review* **80**, 580–594. xxvii, 11, 18, 33

Haldar, J. P., V. J. Wedeen, M. Nezamzadeh, G. Dai, M. W. Weiner, N. Schuff, and Z.-P. Liang: 2012, 'Improved Diffusion Imaging Through SNR-Enhancing Joint Reconstruction'. *Magnetic Resonance in Medicine* **In Press**. 142, 144

Harpen, M. D.: 1992, 'Noise Correlations Exist for Independent RF Coils'. *Magnetic Resonance in Medicine* **23**, 394–397. xlvi, 108

Hayes, C. E. and P. B. Roemer: 1990, 'Noise Correlations in Data Simultaneously Acquired from Multiple Surface Coil Arrays'. *Magnetic Resonance in Medicine* **16**, 181–191. xlvi, 108, 111

He, L. and I. R. Greenshields: 2009, 'A nonlocal maximum likelihood estimation method for rician noise reduction in MR images'. *IEEE Transactions on Medical Imaging* **28**, 165–172. 141, 142

Healy, D. M. and J. B. Weaver: 1992, 'Two Applications of Wavelet Transforms in Magnetic Resonance Imaging'. *IEEE Transactions on Information Theory* **38**, 840–860. 138

Henkelman, R. M.: 1985, 'Measurement of signal intensities in the presence of noise in MR images'. *Medical Physics* **12**, 232–233. xxxix, xlvi, 69, 70, 108, 109, 112, 114, 141, 144, 147

Hess, C. P., P. Mukherjee, E. T. Han, D. Xu, and D. B. Vigneron: 2006, 'Q-Ball Reconstruction of Multimodal Fiber Orientations Using The Spherical Harmonic Basis'. *Magnetic Resonance in Medicine* **56**, 104–117. xxxvi, 45

Hilton, M., T. Ogden, D. Hattery, G. Eden, and B. Jawerth: 1996, *Wavelets in Biology and Medicine.*, Chapt. Wavelet Denoising of Functional MRI Data, pp. 93–112. CRC Press. 138

Hong, J. O., F. Luisier, and P. J. Wolfe: 2012, 'Magnitude MR image denoising via CURE-optimized Non-Local Means'. In: *ISBI*. 141

Hosey, T., G. Williams, and R. Ansorge: 2005, 'Inference of multiple fiber orientations in high angular resolution diffusion imaging'. *Magnetic Resonance in Medicine* **54**, 1480–1489. 41

Hoult, D. I. and R. E. Richards: 1976, 'The signal-to-noise ratio of the nuclear magnetic resonance experiment'. *Journal of Magnetic Resonance (1969)* **24**, 71–85. 106

Ito, K. and K. Xiong: 2000, 'Gaussian filters for nonlinear filtering problems'. *IEEE Transactions on Automatic Control* **45**, 910–927. 65

Jansons, K. M. and D. C. Alexander: 2003, 'Persistent angular structure: new insights from diffusion magnetic resonance imaging data'. *Inverse Problems* **19**, 1031–1046. 41

Jazwinski, A. H.: 1970, *Stochastic processes and filtering theory*. Academic Press, New York. 64

Jbabdi, S., P. Bellec, G. Marrelec, V. Perlbarg, and H. Bénali: 2004, 'A level set method for building anatomical connectivity paths between brain areas using DTI'. In: *ISBI*. 56

Jbabdi, S., P. Bellec, R. Toro, J. Daunizeau, M. Pélégrini-Issac, and H. Bénali: 2008, 'Accurate Anisotropic Fast Marching for Diffusion-Based Geodesic Tractography'. *International Journal of Biomedical Imaging* **2008**, 320195. 56

Jbabdi, S., M. Woolrich, J. Andersson, and T. Behrens: 2007, 'A Bayesian framework for global tractography'. *NeuroImage* **37**, 116–129. 57

Jezzard, P. and R. S. Balaban: 1995, 'Correction for geometric distortion in echo planar images from B0 field variations'. *Magnetic Resonance in Medicine* **34**, 65–73. 69

Jian, B. and B. C. Vemuri: 2007, 'A Unified Computational Framework for Deconvolution to Reconstruct Multiple Fibers From Diffusion Weighted MRI'. *IEEE Transactions on Medical Imaging* **26**, 1464–1471. xxxvi, 53

Jones, D. K. and P. J. Basser: 2004, 'Squashing Peanuts and Smashing Pumpkins: How Noise Distorts Diffusion-Weighted MR Data'. *Magnetic Resonance in Medicine* **52**, 979–993. 120, 123

Jones, D. K., M. A. Horsfield, and A. Simmons: 1999, 'Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging'. *Magnetic Resonance in Medicine* **42**, 515–525. 79, 255

Julier, S. J. and J. K. Uhlmann: 1997, 'A New Extension of the Kalman Filter to Nonlinear Systems'. In: *The 11th International Symposium on Aerospace / Defence Sensing, Simulation and Controls*. 64

Julier, S. J., J. K. Uhlmann, and H. F. Durrant-Whyte: 1995, 'A new approach for filtering nonlinear systems'. In: *American Control Conference*. pp. 1628–1632. 64

Kalman, R. E.: 1960, 'A New Approach to Linear Filtering and Prediction Problems'. *Transactions of the ASME-Journal of Basic Engineering* **82**, 35–45. xxxviii, 64, 65

Kalman, R. E. and R. S. Bucy: 1961, 'New Results in Linear Filtering and Prediction Theory'. *Transactions of the ASME-Journal of Basic Engineering* **83**, 95–107. 64

Kastler, B., D. Vetter, Z. Patay, and P. Germain: 2006, *Comprendre l'IRM Manuel d'autoapprentissage*. xxvi, xxviii, xxxi, 11, 13, 19, 24, 28

Kay, S. M.: 1993, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall PTR. 61, 66, 68, 146, 226

Kezele, I., M. Descoteaux, C. Poupon, F. Poupon, and J.-F. Mangin: 2010, 'Spherical wavelet transform for ODF sharpening'. *Medical Image Analysis* **14**, 332–342. 42

Kim, S. and H. Lim: 2009, 'Fourth-order partial differential equations for effective image denoising'. *Electronic Journal of Differential Equations* **Conf. 17**, 107–121. 140

Koay, C. G. and P. J. Basser: 2006, 'Analytically exact correction scheme for signal extraction from noisy magnitude MR signals'. *Journal of Magnetic Resonance* **179**, 317–322. lii, 142, 145

Koay, C. G., L.-C. Chang, J. D. Carew, C. Pierpaoli, and P. J. Basser: 2006, 'A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging'. *Journal of Magnetic Resonance* **182**, 115–125. lii, 139, 145

Koay, C. G., E. Özarslan, and P. J. Basser: 2009a, 'A signal transformational framework for breaking the noise floor and its applications in MRI'. *Journal of Magnetic Resonance* **197**, 108–119. lii, 142, 144, 145

Koay, C. G., E. Özarslan, and C. Pierpaoli: 2009b, 'Probabilistic Identification and Estimation of Noise (PIESNO): A self-consistent approach and its applications in MRI'. *Journal of Magnetic Resonance* **199**, 94–103. 123

Kolmogorov, A. N.: 1941a, 'Interpolation and extrapolation of stationary random sequences'. *Bulletin of the Academy of Sciences of the USSR: Division of Mathematics* **5**, 3–14. 64

Kolmogorov, A. N.: 1941b, 'Stationary sequences in Hilbert spaces'. *Moscow University Mathematics Bulletin* **2**, 40. 64

Kong, A., J. S. Liu, and W. H. Wong: 1994, 'Sequential imputations and Bayesian missing data problems'. *Journal of the American Statistical Association* **89**, 278–288. 64

Kreher, B. W., I. Mader, and V. G. Kiselev: 2008, 'Gibbs tracking: a novel approach for the reconstruction of neuronal pathways'. *Magnetic Resonance in Medicine* **60**, 953–963. 57

Krissian, K. and S. Aja-Fernández: 2009, 'Noise-Driven Anisotropic Diffusion Filtering of MRI'. *IEEE Transactions on Image Processing* **18**, 2265–2274. 139

Kumar, A., D. Welti, and R. R. Ernst: 1975, 'NMR Fourier zeugmatography'. *Journal of Magnetic Resonance* **18**, 69–83. 27

Kwatra, A., V. Prasanna, and M. Singh: 2006, 'Accelerating DTI Tractography using FPGAs'. In: *IPDPS*. 232

Lam, F., S. D. Babacan, J. P. Haldar, N. Schuff, and Z.-P. Liang: 2012, 'Denoising Diffusion-Weighted MR Image Sequences using Low Rank and Edge Constraints'. In: *ISBI*. 142, 144

Landman, B. A., P.-L. Bazin, and J. L. Prince: 2009, 'Estimation and application of spatially variable noise fields in diffusion tensor imaging'. *Magnetic Resonance Imaging* **27**, 741–751. 124

Laniotis, D. G.: 1976, 'Partitioning: A Unifying Framework for Adaptive Systems, I: Estimation'. *Proceedings of the IEEE* **64**, 1126–1143. 226

Lauterbur, P. C.: 1973, 'Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance'. *Nature* **242**, 190–191. xxv, xxx, 11, 23

Lazar, M. and A. L. Alexander: 2003, 'An error analysis of white matter tractography methods: synthetic diffusion tensor field simulations'. *NeuroImage* **20**, 1140–1153. 54

LeBihan, D.: 1991, 'Molecular diffusion nuclear magnetic resonance imaging.'. *Magnetic Resonance Quarterly* **7**, 1–30. xxxiii, 37

LeBihan, D. and E. Breton: 1985, 'Imagerie de diffusion in vivo par résonance magnétique nucléaire'. *Comptes rendus de l'Académie des sciences. Paris* **301 Série II**, 1109–1112. xxiii, xxxii, 3, 34

LeBihan, D., R. Turner, P. Douek, and N. Patronas: 1992, 'Diffusion MR imaging: clinical applications'. *American Journal of Roentgenology* **Vol 159**, 591–599. xxxii, 34

Legendre, A. M.: 1805, *Nouvelles méthodes pour la détermination des orbites des comètes*. F. Didot. 63

Lenglet, C., R. Deriche, and O. Faugeras: 2004, 'Inferring white matter geometry from diffusion tensor MRI: application to connectivity'. In: *ECCV*. 56

Liang, Z.-P. and P. C. Lauterbur: 1999, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. 250

Liévin, M., F. Luthon, and E. Keeve: 2002, 'Entropic Estimation of Noise for Medical Volume Restoration'. In: *International Conference on Pattern Recognition*. 142, 144

Lindenbaum, M., M. Fischer, and A. M. Bruckstein: 1994, 'On Gabor's contribution to image enhancement'. *Pattern Recognition* **27**, 1–8. 139

Ling, J. and A. C. Bovik: 2002, 'Smoothing low-SNR molecular images via anisotropic median-diffusion'. *IEEE Transactions on Medical Imaging* **21**, 377–384. 139

Liu, J. S. and R. Chen: 1995, 'Blind deconvolution via sequential imputations'. *Journal of the American Statistical Association* **90**, 567–576. 64

Liu, J. S. and R. Chen: 1998, 'Sequential Monte Carlo methods for dynamic systems'. *Journal of the American Statistical Association* **93**, 1032–1044. 64

Liu, X., L. Huang, and Z. Guo: 2011, 'Adaptive fourth-order partial differential equation filter for image denoising'. *Applied Mathematics Letters* **24**, 1282–1288. 140

Ljunggren, S.: 1983, 'A Simple Graphical Representation of Fourier-Based Imaging Methods'. *Journal of Magnetic Resonance* **54**, 338–343. 39

Lysaker, M., A. Lundervold, and X.-C. Tai: 2003, 'Noise Removal Using Fourth-Order Partial Differential Equation With Applications to Medical Magnetic Resonance Images in Space and Time'. *IEEE Transactions on Image Processing* **12**, 1579–1590. 140

Maclaren, J., B. S. Armstrong, R. T. Barrows, K. A. Danishad, T. Ernst, C. L. Foster, K. Gumus, M. Herbst, I. Y. Kadashevish, T. P. Kusik, Q. Li, C. Lovell-Smith, T. Prieto, P. Schulze, O. Speck, D. Stucht, and M. Zaitsev: 2012, 'Measurement and correction of microscopic head motion during magnetic resonance imaging of the brain'. *PLOS one* **7**, In press. 63

Manjón, J. V., J. Carbonell-Caballero, J. J. Lull, and G. García-Martí: 2008, 'MRI denoising using Non-Local Means'. *Medical Image Analysis* **12**, 514–523. 139, 140

Manjón, J. V., P. Coupé, A. Buades, D. L. Collins, and M. Robles: 2012, 'New methods for MRI denoising based on sparseness and self-similarity'. *Medical Image Analysis* **16**, 18–27. 139, 141

Manjón, J. V., P. Coupé, L. Martí-Bonmatí, D. L. Collins, and M. Robles: 2010, 'Adaptive Non-Local Means Denoising of MR Images With Spatially Varying Noise Levels'. *Journal of Magnetic Resonance Imaging* **31**, 192–203. 140

Mansfield, P.: 1977, 'Multi-Planar Image Formation using NMR Spin Echoes'. *Journal of Physics C* **10**, L55–L58. xxv, 11, 39

Marrakchi-Kacem, L.: 2011, 'Connectivité anatomique des noyaux gris centraux: développements méthodologiques et application aux troubles moteurs'. Ph.D. thesis, Université Paris-Sud XI. lxvii, 11, 33, 40, 246

Martin-Fernández, M., E. Munoz-Moreno, L. Cammoun, J.-P. Thiran, C.-F. Westin, and C. Alberola-López: 2009, 'Sequential anisotropic multichannel Wiener filtering with Rician bias correction applied to 3D regularization of DWI data'. *Medical Image Analysis* **13**, 19–35. lii, 141, 145

Martin-Fernández, M., C.-F. Westin, and C. Alberola-López: 2004, '3D Bayesian Regularization of Diffusion Tensor MRI Using Multivariate Gaussian Markov Random Fields'. In: L. N. in Computer Science (ed.): *MICCAI*, Vol. 3216. pp. 351–359. 143

Marzetta, T. L.: 1995, 'EM algorithm for estimating the parameters of multivariate complex Rician density for polarimetric SAR'. In: *ICASSP*, Vol. 5. pp. 3651–3654. 142

Masreliez, C. J.: 1975, 'Approximate Non-Gaussian Filtering with Linear State and Observation Relations'. *IEEE Transactions on Automatic Control* **20**, 107–110. 65, 204

Maybeck, P. S.: 1979, *Stochastic models, estimation, and control*. Academic Press, Inc. (London) LTD. 66

McGibney, G. and M. R. Smith: 1993, 'An unbiased signal-to-noise ratio measure for magnetic resonance images'. *Medical Physics* **20**, 1077–1078. 141, 148

McGraw, T., B. C. Vemuri, Y. Chen, M. Rao, and T. Mareci: 2004, 'DT-MRI denoising and neuronal fiber tracking'. *Medical Image Analysis* **8**, 95–111. 140

Melie-García, L., E. J. Canales-Rodrígez, Y. Alemán-Gómez, C.-P. Lin, Y. Iturria-Medina, and P. A. Valdés-Hernández: 2008, 'A Bayesian framework to identify principal intravoxel diffusion profiles based on diffusion-weighted MR imaging'. *NeuroImage* **42**, 750–770. 57

Merboldt, K.-D., W. Hanicke, and J. Frahm: 1985, 'Self-diffusion NMR imaging using stimulated echoes'. *Journal of Magnetic Resonance (1969)* **3**, 479–486. xxiii, xxxii, 3, 34

Miller, A. and P. M. Joseph: 1993, 'The use of power images to perform quantitative analysis on low SNR MR images.'. *Magnetic Resonance Imaging* **11**, 1051–1056. 141, 148

Mittmann, A., T. H. C. Nobrega, E. Comunello, J. P. O. Pinto, P. R. Dellani, P. Stoeter, and A. von Wangenheim: 2011, 'Performing Real-Time Interactive Fiber Tracking'. *Journal of Digital Imaging* **24**, 339–351. 231

Mori, S., B. J. Crain, V. P. Chacko, and P. C. van Zijl: 1999, 'Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging'. *Annals of Neurology* **45**, 265–269. 54

Morris, D. M., K. V. Embleton, and G. J. M. Parker: 2008, 'Probabilistic fibre tracking: differentiation of connections from chance events'. *NeuroImage* **42**, 1329–1339. 57

Moseley, M. E., J. Mintorovitch, Y. Cohen, H. S. Asgari, N. Derugin, D. Norman, and J. Kucharczyk: 1990, 'Early detection of ischemic injury: comparison of spectroscopy, diffusion-, T2-, and magnetic susceptibility-weighted MRI in cats'. *Magnetic Resonance in Medicine* **51**, 207–209. xxxii, 34, 38

Muresan, D. D. and T. W. Parks: 2003, 'Adaptive principal components and image denoising'. In: *IEEE International Conference on Image Processing*, Vol. 1. pp. 101–104. 139

Nowak, R. D.: 1999, 'Wavelet-Based Rician Noise Removal for Magnetic Resonance Imaging'. *IEEE Transactions on Image Processing* **8**, 1408–1419. xlvi, 108, 138, 139, 144

Ooi, M. B., S. Krueger, W. J. Thomas, S. V. Swaminathan, and T. R. Brown: 2009, 'Prospective real-time correction for arbitrary head motion using active markers'. *Magnetic Resonance in Medicine* **62**, 943–954. 63

Özarslan, E., C. G. Koay, and P. J. Basser: 2009, 'Remarks on q-space MR propagator in partially restricted, axially-symmetric, and isotropic environments'. *Magnetic Resonance Imaging* **27**, 834–844. 42

Papadakis, N. G., C. D. Murrills, L. D. Hall, C. L.-H. Huang, and T. A. Carpenter: 2000, 'Minimal gradient encoding for robust estimation of diffusion anisotropy'. *Magnetic Resonance Imaging* **18**, 671–679. 79, 255

Parker, G. J. and D. C. Alexander: 2003, 'Probabilistic Monte Carlo Based Mapping of Cerebral Connections Utilizing Whole-Brain Crossing Fibre Information'. *Information Processing in Medical Imaging* **18**, 684–695. 56

Parker, G. J. M., J. A. Schnabel, M. R. Symms, D. J. Werring, and G. J. Barker: 2000, 'Nonlinear Smoothing for Reduction of Systematic and Random Errors in Diffusion Tensor Imaging'. *Journal of Magnetic Resonance Imaging* **11**, 702–710. 139

Parker, G. J. M., C. A. M. Wheeler-Kingshott, and G. J. Barker: 2002, 'Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging'. *IEEE Transactions on Medical Imaging* **21**, 505–512. 56

Perona, P. and J. Malik: 1990, 'Scale-Space and Edge Detection Using Anisotropic Diffusion'. *IEEE Transactions on pattern analysis and machine intelligence* **12**, 629–639. 139

Perrin, M., Y. Cointepas, A. Cachia, C. Poupon, B. Thirion, D. Rivière, P. Cathier, V. E. Kouby, A. Constantinesco, D. LeBihan, and J.-F. Mangin: 2008, 'Connectivity-based parcellation of the cortical mantle using q-ball diffusion imaging'. *International Journal of Biomedical Imaging*. 55

Perrin, M., C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Rivière, A. Constantinesco, D. LeBihan, and J.-F. Mangin: 2005, 'Fiber tracking in q-ball fields using regularized particle trajectories'. In: *IPMI*. 54, 56

Pierpaoli, C.: 1997, 'Oh no! One more method for color mapping of fiber tract direction using diffusion MR imaging data'. In: *ISMRM*. 43

Pitt, M. K. and N. Shephard: 1999, 'Filtering via simulation: auxiliary particle filters'. *Journal of the American Statistical Association* **94**, 590–599. 64

Pižurica, A., W. Philips, I. Lemahieu, and M. Acheroy: 2003, 'A Versatile Wavelet Domain Noise Filtration Technique for Medical Imaging'. *IEEE Transactions on Medical Imaging* **22**, 323–331. 138

Plataniotis, K. N., D. Androutsos, and A. N. Venetsanopoulos: 1997, 'Nonlinear Filtering of Non-Gaussian Noise'. *Journal of Intelligent and Robotic Systems* **19**, 207–231. lix, 65, 204, 207, 210, 225, 226, 227

Poupon, C.: 1999a, 'Détection des faisceaux de fibres de la substance blanche pour l'étude de la connectivité anatomique cérébrale'. Ph.D. thesis, Ecole Nationale Supérieure des Télécommunications. xxxiii, 11, 33, 35, 44, 54, 57

Poupon, C.: 2010, 'Le processus de diffusion: un puissant outil pour sonder la structure du cerveau in vivo'. Technical report, CEA, NeuroSpin. xxxiii, 33, 35, 42, 45

Poupon, C., C. A. Clark, V. Frouin, J. Régis, I. Bloch, D. LeBihan, and J.-F. Mangin: 2000, 'Regularization of Diffusion-Based Direction Maps for the Tracking of Brain White Matter Fascicles'. *NeuroImage* **12**, 184–195. 143

Poupon, C., F. Poupon, A. Roche, and J.-F. Mangin: 2008a, 'Real-time magnetic resonance diffusion imaging'. Patent. lxvi, 246

Poupon, C., A. Roche, J. Dubois, J.-F. Mangin, and F. Poupon: 2008b, 'Real-time MR diffusion tensor and Q-ball imaging using Kalman filtering'. *Medical Image Analysis* **12**, 527–534. xxiii, xxxvii, xxxviii, xl, xlii, 3, 61, 62, 65, 71, 72, 73, 74, 75, 76, 78, 79, 80, 81, 245

Poupon, F.: 1999b, '"Parcellisation" systématique du cerveau en volumes d'intérêt. Le cas des structures profondes.'. Ph.D. thesis, Institut National des Sciences Appliquées de Lyon. lxvii, 246

Poupon, F.: 2008, 'Techniques de reconstruction en imagerie parallèle'. In: *GRAMM*. p. 33. 31

Poupon, F. and O. Riff: 2009, 'Projet Imagerie temps-réel'. Technical report, CEA, NeuroSpin. xliii, 82, 83

Pruessmann, K. P., M. Weiger, M. B. Scheidegger, and P. Boesiger: 1999, 'SENSE: Sensitivity Encoding for Fast MRI'. *Magnetic Resonance in Medicine* **42**, 952–962. 30, 32

Purcell, E. M., H. C. Torrey, and R. V. Pound: 1946, 'Resonance Absorption by Nuclear Magnetic Moments'. *Physical Review* **69**, 37–38. xxvi, 16

Quin, L., P. van Gelderen, J. A. Derbyshire, F. Jin, and J. Lee: 2009, 'Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system'. *Magnetic Resonance in Medicine* **62**, 924–934. 63

Rajan, J., B. Jeurissen, M. Verhoye, J. Van Audekerke, and J. Sijbers: 2011, 'Maximum likelihood estimation-based denoising of magnetic resonance images using restricted local neighborhoods'. *Physics in Medicine and Biology* **56**, 5221–5234. 141, 142

Rajan, J., D. Poot, J. Juntu, and J. Sijbers: 2010, 'Noise measurement from magnitude MRI using local estimates of variance and skewness'. *Physics in Medicine and Biology* **55**, N441–N449. l, lxv, 127, 128, 129, 244

Rajan, J., J. Van Audekerke, A. Van der Linden, M. Verhoye, and J. Sijbers: 2012a, 'An adaptive non local maximum likelihood estimation method for denoising magnetic resonance images'. In: *ISBI*. 141, 142

Rajan, J., J. Veraart, J. V. Audekerke, M. Verhoye, and J. Sijbers: 2012b, 'Nonlocal maximum likelihood estimation method for denoising multiple-coil magnetic resonance images'. *Magnetic Resonance Imaging* **in press**. xlvii, 118, 141, 142

Redpath, T. W.: 1992, 'Noise Correlation in Multicoil Receiver Systems'. *Magnetic Resonance in Medicine* **24**, 85–89. xlvi, 108, 111

Redpath, T. W.: 1998, 'Signal-to-noise ratio in MRI'. *The British Journal of Radiology* **71**, 704–707. 106, 107

Reisert, M., I. Mader, C. Anastasopoulos, M. Weigel, S. Schnell, and V. Kiselev: 2011, 'Global fiber reconstruction becomes practical'. *NeuroImage* **54**, 955–962. xxxvii, 57, 58

Rice, S. O.: 1952, 'Mathematical analysis of random noise- and Appendixes'. Technical report, Bell Telephone Labs Inc. New York. xlvi, 110

Roemer, P. B., W. A. Edelstein, C. E. Hayes, S. P. Souza, and O. M. Mueller: 1990, 'The NMR phased array'. *Magnetic Resonance in Medicine* **16**, 192–225. xxxi, xlvi, 29, 108

Rudin, L., S. Osher, and E. Fatemi: 1992, 'Nonlinear total variation based noise removal algorithms'. *Physica D* **60**, 259–268. 140

Salvador, R., A. Peña, D. K. Menon, T. A. Carpenter, J. D. Pickard, and E. T. Bullmore: 2005, 'Formal Characterization and Extension of the Linearized Diffusion Tensor Model'. *Human Brain Mapping* **24**, 144–155. lii, 71, 139, 145

Samsonov, A. A. and C. R. Johnson: 2004, 'Noise-Adaptive Nonlinear Diffusion Filtering of MR Images With Spatially Varying Noise Levels'. *Magnetic Resonance in Medicine* **52**, 798–806. 139, 140

Sapiro, G. and D. L. Ringach: 1996, 'Anisotropic diffusion of multivalued images with applications to color filtering'. *IEEE Transactions on Image Processing* **5**, 1582–1586. 139

Schmitt, B., D. Duclap, A. Lebois, P. Guevara, H. Zhang, Y. Assaf, F. Poupon, D. Rivière, T. Cointepas, D. LeBihan, J.-F. Mangin, and C. Poupon: 2013, 'A novel probabilistic connectivity atlas for the human connectome: the CONNECT/ARCHI atlas'. In: *ISMRM*. 249

Sijbers, J., A. J. D. Dekker, P. Scheunders, and D. V. Dyck: 1998, 'Maximum Likelihood estimation of Rician distribution parameters'. *IEEE Transactions on Medical Imaging* **17**, 357–361. 70, 112, 113, 142

Sijbers, J. and A. J. den Dekker: 2004, 'Maximum Likelihood estimation of signal amplitude and noise variance from MR data'. *Magnetic Resonance in Medicine* **51**, 586–594. 142, 144

Sijbers, J., A. J. den Dekker, A. V. D. Linden, M. Verhoye, and D. V. Dyck: 1999a, 'Adaptive anisotropic noise filtering for magnitude MR data'. *Magnetic Resonance Imaging* **17**, 1533–1539. 139

Sijbers, J., A. J. den Dekker, E. Raman, and D. V. Dyck: 1999b, 'Parameter estimation from magnitude MR images'. *International Journal of Imaging Science and Technology* **10**, 109–114. 142

Singh, M., A. Kwatra, C.-W. Wong, and V. Prasanna: 2006, 'Acceleration of fiber tracking in DTI tractography by reconfigurable computer hardware'. In: *IEEE EMBS*. 232

Sodickson, D. K. and W. J. Manning: 1997, 'Simultaneous Acquisition of Spatial Harmonics (SMASH): Fast Imaging with Radiofrequency Coil Arrays'. *Magnetic Resonance in Medicine* **38**, 591–603. 30

Sorenson, H. and D. L. Alspach: 1971, 'Recursive Bayesian estimation using Gaussian sum'. *Automatica* **7**, 465–479. 65, 204

Stejskal, E. and J. Tanner: 1965, 'Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient'. *The Journal of Chemical Physics* **42**, 288–292. xxxiii, 34, 36, 40, 41

Strong, D. M. and T. F. Chan: 1996, 'Spatially and Scale Adaptive Total Variation Based Regularization and Anisotropic Diffusion in Image Processing'. Technical report, UCLA CAM Report. 140

Swerling, P.: 1959, 'First Order Error Propagation in a Stagewise Smoothing Procedure for Satellite Observations'. *Journal of the Astronautical Sciences* **6**, 46–52. 64

Taylor, D. G. and M. C. Bushell: 1985, 'The spatial mapping of translational diffusion coefficients by the NMR imaging technique'. *Phyics in Medicine and Biology* **30**, 345–349. xxiii, xxxii, 3, 34

Thiele, T. N.: 1880, 'Om Anvendelse af mindste Kvadraters Methode i nogle Tilfaelde, hvor en Komplikation af visse Slags uensartede tilfaeldige Fejlkilder giver Fejlene en systematisk Karakter. English version: On the application of the method of least squares to some cases, in which a combination of certain types of inhomogeneous random sources of errors gives these a systematic character'. *naturvidenskabelig og mathematisk Afdeling* **12**, 381–408. 64

Thunberg, P. and P. Zetterberg: 2007, 'Noise distribution in SENSE- and GRAPPA-reconstructed images: a computer simulation study'. *Magnetic Resonance Imaging* **25**, 1089–1094. 111, 118

Torrey, H. C.: 1956, 'Bloch Equations with Diffusion Terms'. *Physical Review* **104**, 563–565. 34

Tournier, J.-D., F. Calamante, and A. Connelly: 2007, 'Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution'. *NeuroImage* **35**, 1459–1472. xxxvi, 53

Tournier, J.-D., F. Calamante, D. G. Gadian, and A. Connelly: 2004, 'Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution'. *NeuroImage* **23**, 1176–1185. xxxv, xxxvi, 39, 53

Tristán-Vega, A. and S. Aja-Fernández: 2010, 'DWI filtering using joint information for DTI and HARDI'. *Medical Image Analysis* **14**, 205–218. lii, 141, 145, 148, 150

Tristán-Vega, A., S. Aja-Fernández, and C.-F. Westin: 2012, 'Least squares for diffusion tensor estimation revisited: Propagation of uncertainty with Rician and non-Rician signals'. *NeuroImage* **59**, 4032–4043. lii, 139, 145

Tristán-Vega, A., V. García-Pérez, S. Aja-Fernández, and C.-F. Westin: 2012, 'Efficient and robust nonlocal means denoising of MR data based on salient features matching'. *Computer Methods and Programs in Biomedicine* **105**, 131–144. 140, 146, 164

Tristán-Vega, A., C.-F. Westin, and S. Aja-Fernández: 2009, 'Estimation of fiber Orientation Probability Density Functions in High Angular Resolution Diffusion Imaging'. *NeuroImage* **47**, 638–650. xxxiv, xxxvi, 39

Tschumperlé, D. and L. Brun: 2009, 'Lissage Non Local d'Images Multi-Valuées par Diffusion Anisotrope dans l'Espace des Patches'. In: *Colloque GRETSI*. 139

Tuch, D. S.: 2002, 'Diffusion MRI of complex tissue structure'. Ph.D. thesis, Harvard University–MIT Division of Health Sciences and Technology. 33, 41, 152, 256

Tuch, D. S.: 2004, 'Q-Ball imaging'. *Magnetic Resonance in Medicine* **52**, 1358–1372. xxxiv, xxxvi, 39, 46, 47, 51, 52, 59

Tuch, D. S., T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen: 2002, 'High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity'. *Magnetic Resonance in Medicine* **48**, 577–582. 54

Šimandl, M., J. Královec, and T. Söderström: 2006, 'Advanced point-mass method for nonlinear state estimation'. *Automatica* **42**, 1133–1145. 64

Vaillancourt, O., A. Boré, G. Girard, and M. Descoteaux: 2010, 'A Fiber Navigator for Neurosurgical Planning (NeuroPlanningNavigator)'. In: *IEEE Visualization*. 231

van der Merwe, R., A. Doucet, N. de Freitas, and E. Wan: 2000, 'The Unscented Particle Filter'. In: *Conference on Neural Information Processing Systems*. pp. 584–590. 64

Vegas-Sánchez-Ferrero, G., A. Tristán-Vega, S. Aja-Fernández, M. Martín-Fernández, C. Palencia, and R. Deriche: 2012, 'Anisotropic LMMSE denoising of MRI based on statistical tissue models'. In: *ISBI*. 141

Vlaardingerbroek, M. T. and J. A. den Boer: 2004, *Magnetic Resonance Imaging: Theory and Practice*. 106, 107

Wan, E. A. and R. van der Merwe: 2000, 'The Unscented Kalman Filter for Nonlinear Estimation'. In: *Symposium on Adaptive Systems for Signal Processing, Communications, and Control*. pp. 153–158. 64

Wang, J., T. Kluge, M. Nittka, V. Jellus, B. Kuehn, and B. Kiefer: 2001, 'Parallel Acquisition Techniques with Modified SENSE Reconstruction: mSENSE'. In: *Proceedings of the First Würzburg Workshop on Parallel Imaging: Basics and Clinical Applications*. 118

Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli: 2004a, 'Image Quality Assessment: From Error Visibility to Structural Similarity'. *IEEE Transactions on Image Processing* **13**, 600–612. 131

Wang, Z., B. C. Vemuri, Y. Chen, and T. H. Mareci: 2004b, 'A Constrained Variational Principle for Direct Estimation and Smoothing of the Diffusion Tensor Field From Complex DWI'. *IEEE Transactions on Medical Imaging* **23**, 930–939. lii, 140, 145

Wedeen, V., T. Reese, D. S. Tuch, M. Wiegel, J.-G. Dou, R. Weiskoff, and D. Chessler: 2000, 'Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI'. In: *ISMRM*. 41

Wedeen, V. J., P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff: 2005, 'Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging'. *Magnetic Resonance in Medicine* **54**, 1377–1386. 41

Weiger, M., K. P. Pruessmann, C. Leussler, P. Röschmann, and P. Boesiger: 2001, 'Specific Coil Design for SENSE: A Six-Element Cardiac Array'. *Magnetic Resonance in Medicine* **45**, 495–504. 118

Weiskopf, N., K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer: 2004, 'Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)'. *IEEE transactions on biomedical engineering* **51**(6), 966–970. xxiii, 3

Westin, C.-F., S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis: 2002, 'Processing and visualization for diffusion tensor MRI'. *Medical Image Analysis* **6**, 93–108. 54

White, N., C. Roddey, A. Shankaranarayanan, E. Han, D. Rettmann, J. Santos, J. Kuperman, and A. Dale: 2010, 'PROMO: Real-Time Prospective Motion Correction in MRI Using Image-Based Tracking'. *Magnetic Resonance in Medicine* **63**, 91–105. 63

Wiener, N.: 1949, *Extrapolation, Interpolation and Smoothing of Time Series, with Engineering Applications*. Wiley. Originally apperas in 1942 as a classified National Defense Research Council Report. 64

Wiener, N. and E. Hopf: 1931, 'On a class of singular integral equations'. In: *Proceedings of the Prussian Acad. Math. - Phys. Ser.* p. 696. 64

Wiest-Daesslé, N., S. Prima, P. Coupé, M. S. Patrick, and C. Barillot: 2008, 'Rician noise removal by non-Local Means filtering for low signal-to-noise ratio MRI: applications to DT-MRI'. In: *Medical Image Computing and Computer-Assisted Intervention*, Vol. 5242. pp. 171–179. 140

Wood, J. C. and K. M. Johnson: 1999, 'Wavelet Packet Denoising of Magnetic Resonance Images: Importance of Rician Noise at Low SNR'. *Magnetic Resonance in Medicine* **41**, 631–635. 138, 139

Yang, G. Z., P. Burger, D. N. Firmin, and S. R. Underwood: 1995, 'Structure adaptive anisotropic filtering for magnetic resonance images'. *Lecture Notes in Computer Science* **970**, 384–391. 139

Yaroslavsky, L. P., K. O. Egiazarian, and J. T. Astola: 2001, 'Transform domain image restoration methods: review, comparison, and interpretation'. In: *SPIE Proceedings, Nonlinear Image Processing and Pattern Analysis XII*. 139

Zaitsev, M., C. Dold, G. Sakas, J. Hennig, and O. Speck: 2006, 'Magnetic resonance imaging of freely moving objects: Prospective real-time motion correction using an external optical motion tracking system'. *NeuroImage* **31**, 1038–1050. 63

Zhang, F., C. Goodlett, E. Hancock, and G. Gerig: 2007, 'Probabilistic Fiber Tracking Using Particle Filtering'. In: *MICCAI*. 57