Universidade de Lisboa
Faculdade de Ciência
Departamento de Física


# Diffusion Kurtosis Imaging of the Healthy Human Brain

**Rafael Neto Henriques**

Dissertação
Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Radiações em Diagnóstico e Terapia

**2012**

Universidade de Lisboa
Faculdade de Ciência
Departamento de Física


# Diffusion Kurtosis Imaging of the Healthy Human Brain

**Rafael Neto Henriques**

Dissertação
Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Radiações em Diagnóstico e Terapia

Orientadores:
Doutora Marta Morgado Correia
Professor Doutor Hugo Alexandre Ferreira

**2012**

## Resumo

Nas últimas décadas, a esperança média de vida tem vindo a aumentar. Desta forma, é cada vez mais importante perceber como se pode providenciar um envelhecimento saudável. Apesar de ser comum relacionar o envelhecimento com a perda de capacidades cognitivas, estudos recentes sugerem que certas capacidades cognitivas podem ser preservadas, ou mesmo melhoradas, uma vez que o cérebro pode manter-se flexível ao longo dos anos. O projecto '*Cambridge Centre for Ageing and Neuroscience*' (Cam-CAN) é um projecto colaborativo que tem como objectivo determinar o grau de flexibilidade neural ao longo da vida e o potencial de reorganização neuronal para a sustentação das capacidades cognitivas. Esse conhecimento será importante para compreender como podemos preservar os recursos cognitivos e aumentar a qualidade de vida. Deste modo, pretende-se analisar dados demográficos, psicológicos, físicos e neuronais de 700 participantes com idades compreendidas entre os 18 e os 90 anos. Relativamente aos dados neuronais, este projecto inclui dados obtidos por magnetoencefalografia e por ressonância magnética estrutural, funcional e por difusão.

Desde a década de oitenta, as técnicas de ressonância magnética por difusão têm mostrado grande utilidade em várias aplicações. Como os processos de difusão dependem das características microscópicas dos tecidos, a ressonância magnética por difusão pode revelar alterações microscópicas na resolução normal das imagens de ressonância magnética. Em 1994, Basser, Mattiello e Lebihan introduziram uma nova perspectiva nas medidas de difusão para tecidos anisotrópicos, como é o caso da matéria branca do cérebro. Conhecida por imagem por tensor de difusão ou *diffusion tensor imaging* (DTI), esta técnica caracteriza os processos de difusão por tensores em vez de um único escalar. Contudo, a DTI baseia-se na hipótese que a difusão de moléculas de água pode ser caracterizada por uma distribuição Gaussiana, o que foi demonstrado não ser completamente adequado. Em particular, foi observado que as barreiras das microestruturas tecidulares introduzem desvios simétricos relativamente à distribuição Gaussiana. Imagem por curtose de difusão ou *diffusion kurtosis imaging* (DKI) é uma expansão da DTI, onde o tensor de difusão é estimado em conjunto com o tensor de curtose que caracteriza o grau de não-gaussianidade de difusão em meios anisotrópicos. Esta técnica proporciona melhores estimativas dos parâmetros de difusão comparativamente à DTI, obtendo-se um índice de complexidade das barreiras dos

tecidos. Por exemplo, a partir dos tensores de difusão e curtoses podem ser obtidos: 1) os valores da difusão e curtoses na direcção paralela às fibras de matéria branca ($\lambda_{\parallel}$ e $k_{\parallel}$); 2) os valores médios da difusão e curtoses nas direcções perpendiculares às fibras de matéria branca ($\lambda_{\perp}$ e $k_{\perp}$); 3) os valores médios de difusão e curtoses em todas as orientações espaciais (MD e MK); e 4) um índice de anisotropia do tensor de difusão (FA). Tal como qualquer técnica de ressonância magnética, a DKI está sujeita a ruído e a artefactos, o que pode levar a valores implausíveis de difusão e curtoses. Desta forma, é importante desenvolver métodos mais robustos para o processamento e análise da DKI.

O trabalho desta tese teve como objectivo melhorar os dados da DKI extraidos dos dados de difusão do Cam-CAN (que correspondem a 63 volumes de imagens para cada sujeito: 30 direcções para b-values = 1000, 2000 $s.mm^{-2}$ e 3 b-value=0 $s.mm^{-2}$). Para tal, foram testados diferentes algoritmos para a estimativa dos tensores de difusão e curtoses em dados de ressonância magnética por difusão simulados e corrompidos por ruído Riciano artificial. Os diferentes algoritmos incluíram: 1) uma solução simples obtida pela minimização do modelo da DKI pelos métodos dos mínimos quadrados (OLS); 2) uma solução mais geral obtida pela minimização dos mínimos quadrados onde uma estimativa da variância dos pontos é inserida para corrigir a não homocedasticidade dos dados (WLS); 3) um método iterativo não linear (NLS); 4) um método iterativo não linear em que o tensor de difusão é escrito pela decomposição de Choslesky para forçar estimativas plausíveis de difusão (chNLS); 5) um método iterativo linear em que restrições são introduzidas para impedir que valores de difusão e curtoses sejam estimados por valores implausíveis (CLS); e 6) um novo método proposto para a estimação directa dos valores de MD e MK (DLS). Dois tipos de simulações foram testados: simulações baseadas em tensores para um único voxel de matéria branca e baseados em tensores correspondentes a um cérebro inteiro. Para os dois tipos de simulação, os métodos WLS e NLS mostram ter uma eficiência semelhante ao método OLS. O método chNLS mostra ser capaz de evitar os valores implausíveis de FA, contudo é um método de baixa eficiência de computação. Apesar de ser uma das aproximações mais utilizadas nos estudos mais recentes da DKI, o método CLS resultou em dados contaminados por artefactos. Por outro lado, o novo método DLS mostrou estimar os valores de MD e MK menos sensíveis aos artefactos introduzidos pelo ruído artificial Riciano, no entanto este método não possibilita a estimação de outros parâmetros além de MD e MK.

Após selecção do método para a extracção dos tensores de difusão e curtoses, são estudados métodos de pré-tratamento das imagens de ressonância magnética ponderada

em difusão. Como nenhum dos métodos alternativos apresentou vantagens satisfatórias em extrair os valores de tensor difusão e curtoses, OLS vai ser utilizado para extrair os valores de difusão e curtoses de imagens de ressonância magnética ponderada em difusão filtradas com o filtro Gaussiano com diferentes valores de FWHM, com o objectivo de encontrar um valor ideal de FWHM. Como os valores de MK foram os dados que mostraram ser mais sensíveis ao ruído Riciano, a selecção do pré-processamento foi baseado em parâmetros de qualidade das imagens de MK. Como a sensibilidade às diferenças relacionadas com o envelhecimento dependem também dos diferentes métodos de pré-tratamento, a selecção dos métodos adequados para o projecto Cam-CAN também tem em conta os valores de significância de diferenças detectadas entre dois grupos de sujeitos com diferentes idades: jovens adultos (m 26.2, sd 3.9) e adultos de meia-idade (m 53.4, sd 2.0). Para analisar possíveis vantagens em utilizar voxels com pior resolução, valores diferentes de FWHM do filtro Gaussiano são estudados não só nos dados com a resolução original de $2x2x2mm^3$, mas também em dados com resolução de $2.7x2.7x2.7mm^3$ obtidos por uma aquisição com esta resolução e por um algoritmo de *downsampling* implementado em Matlab. A utilização do filtro Gaussiano nas imagens ponderadas em difusão mostram reduzir os impactos introduzidos pelo ruído e artefactos nas imagens, em particular são observadas reduções na subestimação dos valores de curtoses. No estudo de optimização e para imagens com resolução de $2x2x2mm^3$, o filtro Gaussiano de FWHM = 2.5 mm mostra um compromisso ideal entre a qualidade das imagens, a resolução e a sensibilidade em detectar diferenças entre os dois grupos de diferentes idades. Além de possibilitar acquisições mais rápidas, a utilização de dados com voxels de maior resolução ($2.7x2.7x2.7mm^3$) não mostrou fornecer vantagens para DKI, uma vez que não foram suficientes para a eliminação dos valores implausíveis de curtose.

Após seleccionados os métodos adequados para o seu processamento, DKI é aplicado em 103 sujeitos do Cam-CAN (de idades entre os 18 e os 89 anos) para um estudo preliminar do evelhecimento saudável. Este estudo teve como objectivo mostrar a potencialidade da DKI em fornecer informação única acerca das alterações das microestruturas do cérebro para estudos futuros do Cam-CAN. Desta forma, são efectuados três tipos de análise aos dados extraídos pela DKI: análises baseadas em valores médios de *voxels* seleccionados por regiões de interesse; análises baseadas nos histogramas de *voxels* seleccionados por regiões de interesse; e uma análise estatística espacial baseada em tractos (ou denominada em inglês, *tract based spatial statistics* -

TBSS). Relativamente aos valores extraídos pelo tensor de difusão (i.e. MD, FA, $\lambda_{\parallel}$ e $\lambda_{\perp}$), as alterações que se observam em função da idade estão de acordo com estudos anteriores da DTI e *post mortem*. Resumidamente, o aumento dos valores de MD, $\lambda_{\parallel}$ e $\lambda_{\perp}$ e decréscimo dos valores de FA na matéria branca reflectem processos de degeneração, como exemplo, a desmielinização. Uma vez que a DKI é uma técnica muito recente, os valores de curtose podem requerer testes adicionais para validar a sua precisão, como o índice de complexidade da barreira microestrutural, nomeadamente na intrepretação de diferenças observadas entre as análises baseadas em valores médios de MK e picos de histogramas das intensidades de MK na matéria cinzenta. A análise TBSS mostrou que, ao longo do envelhecimento, diferentes tractos apresentam perfis de alterações diferentes, sugerindo que as mudanças na capacidade cognitiva podem estar relacionadas com mudanças específicas na conectividade entre determinadas regiões do cérebro humano. Em geral, os valores de curtose mostram ter potencialidade em revelar informação única das alterações das barreiras microestruturais com a idade, tendo uma aplicação pontencial no estudo do envelhecimento saudável do cérebro pelo projecto Cam-CAN.

Em trabalhos futuros, o desenvolvimento dos procedimentos para DKI poderão ser continuados com a introdução de um algoritmo adequado para a correcção de artefactos de movimento, e com a introdução de algoritmos mais sofisticados para a remoção de ruído Riciano baseados na probabilística *bayeseana,* ou algoritmos baseados em médias não locais. A DKI poderá também ser aplicada nos restantes dados adquiridos pelo Cam-CAN. Outras regiões de interesse poderão ser estudadas, e o estudo TBSS pode ser continuado testando modelos lineares mais gerais. Finalmente, poderão ser estudadas as relações entre as alterações estruturais medidas pela DKI com os dados de ressonância magnética de difusão, e com dados demográficos e comportamentais.

## Abstract

Diffusion kurtosis imaging (DKI) is an extension of diffusion tensor imaging (DTI), which provides estimates of the kurtosis tensor in addition to the diffusion tensor. Values of kurtosis in biological tissues are of interest since they are believed to provide information of microstructural barrier complexity. The objective of this thesis is to optimize processing and analysis procedures of DKI to obtain reliable measures of brain microstructures age changes for a large collaborative project, the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) project. Additionally, it aims to show the potential of DKI on a preliminary ageing study.

In this project, the quality of DKI metrics will be improved by optimising the steps for pre-processing of diffusion-weighted images, as well as the methods for the estimation of diffusion and kurtosis tensors. Then DKI is applied to the Cam-CAN data recorded so far (103 subjects aged from 18 to 89) and the ageing changes are studied using an analysis based on regions of interest, histograms and tract-based spatial statistics.

The results of this thesis show novel findings on the development of DKI estimation framework. For example, constraints on a DKI linear solution to impose plausible values of diffusion and kurtosis shows also to produce results with large amounts of overestimations and underestimations. On the other hand, a newly proposed method, direct linear squares, provides the most robust values of mean diffusion and kurtosis. In addition, a Gaussian kernel with FWHM of 2.5 mm reduces the noise effects with an optimal compromise between accuracy of measures and resolution. Regarding the preliminary ageing study, the indexes of tissue complexity extracted from the kurtosis tensor showed different information relative to metrics extracted from the diffusion tensor, and thus, DKI is a potential technique to identify unique ageing changes on microstructures for the Cam-CAN project.

## Acknowledgements

It was a real pleasure to work at the MRC Cognition and Brain Sciences Unit, and a privilege to give my comtribution to The Cambridge Centre for Ageing and Neuroscience. Therefore I would like to express my gratitude to all of those who made this thesis possible.

First of all, I want to thank the Erasmus Programme for the funding provided and my parents for the all the support given. I would like to thank my supervisor, Dr Marta Correia for all the geniuses suggestions, comments and directions for the work reported on the thesis, and also for the sympaty and patiente. I am also thankful to my supervisor, Prof. Dr. Hugo Ferreira for the brilant ideias and for having always trusted on my skills.

Thanks to my colleagues and friends at the MRC CBU you make my time in Cambridge very enjoyable, especially Andrea Greve, Alex Walther, Alexandra Krugliak, Beth Parkin, Darius Gadeikis, Felix Triple, Francesco, Marieke Mur, Max Garagnani, Nick Furl, Ruud Berkers, Sasa Redzepoic, Shimmin Wang, Stanimira Georgieva, and Viktoria Havas.

I want to thank my housemates Joana Loureiro e Catarina Rua for the patience, and my friends Débora Salvado, Filipa Costa e Luís Lacerda for the great weekends on London. I am also gratefull for the support given by Débora during my masters.

Very special thanks to Marta Dias for being always present on the last years. I also want to thank my friend and colleagues Ana Figueira, André Ribeiro, Cláudia Lopes, Catarina Fernandes, Diana Rosa, Federico Severo, João Monteiro, Maria Rodriges, Nuno Silva.

I would like to thank my friends from home for always binging in touch: Ana Augosto, Ana Carolina Cordeiro, Luís Cordeiro, João Santos, João Pedro Cordeiro, Rui Fiel, Ricardo Augosto, e Tiago Cerejo.

Thanks to Professor Guiomar Evans for helping me with the bureaucratic issues related to the ERASMUS funding, and Prof. Pascal van Lieshout for always being in touch.

Finally, I want to thank my brothers Emanuel, Tiago, and Francisco, and my cousins Carlos e Carla Vicente.

Aos meu pais pelo apoio e paciência
E para aqueles que sempre tiveram motivação e gosto,
mas nunca a oportunidade

# Contents

# Chapter 1

# Introduction

## 1.1. Cambridge Centre for Ageing and Neuroscience

Life expectancy has been increasing over the last decades making it increasingly important to understand how to age healthily. Ageing affect cognitive abilities such as memory, language, attention, emotion, action, and motor learning. Although ageing is often seen as a process of cognitive decline and decay, new scientific discoveries suggest a different view. The brain seems to remain flexible and adaptable across the lifespan and therefore many cognitive abilities can be preserved or even improved to compensate other abilities that have been lost (Reuter-Lorenz & Park, 2010). For example, previous studies have shown that older adults have larger activations or activate more brain regions when compared with younger subjects while performing the same tasks (e.g. Cabeza, Anderson, Locantore, & McIntosh, 2002; Grady et al., 1994; Reuter-Lorenz & Cappell, 2008; Vallesi, McIntosh, & Stuss, 2011). These studies suggested that recruitment of extra neural resources compensates for the impaired brain regions for successfully overcoming proponent and inappropriate functional responses.

The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) is a large collaborative research project funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and aims to determine the extent of neural flexibility across the lifespan and the potential for neural re-organisation to sustain mental cognitive (Dixon, 2010). This knowledge will be important to understand how one can preserve cognitive capabilities and increase the quality of life. Moreover understanding the process of healthy ageing will provide background to detect and interprets neurodegenerative diseases beyond those of normal ageing. Cam-CAN will have the

contribution of over 30 research projects from groups of Cambridge, including the Department of Experimental Psychology, Public health and Primary Care, Psychiatry, Clinical Neurosciences, and Engineering in the University of Cambridge and the Medical Research Council Cognition and Brain Sciences Units (CBU).

For a clear understanding of the different stages of healthy ageing it is necessary to study a population-representative cohort with a large number of subjects. Therefore Cam-CAN is proposing to study a large group of 700 participants aged 18 to 90. Since only healthy subjects can be included in this study, the participants will be selected from 3000 members of the Cambridge City GP surgeries based on their health history and lifestyle and some objective measurements as reaction time and hearing tests. The selected subjects will carry on behavioural tests to quantify cognitive capacities related to memory, action and motor learning, attention, emotion, and language. Brain activity and structure will be measured using magnetoencephalography (MEG) and magnetic resonance imaging (MRI), including structural, functional and diffusion MRI.

## 1.2. The Impacts of Diffusion Magnetic Resonance Imaging

Since the 1980s diffusion magnetic resonance imaging (dMRI) had shown to provide useful techniques for several studies. Since diffusion is highly dependent on the intrinsic proprieties of the surrounding microstructures, dMRI is sensitive to unique structural information that is not resolved by conventional MRI structural images.

Measures of diffusion were pioneered in nuclear magnetic resonance (NMR) spectroscopy using NMR spin-echo (Carr & Purcell, 1954; Hahn, 1950) and NMR pulsed-gradient spin-echo sequences (Stejskal & Tanner, 1965) however it was only in the 1980s that the NMR diffusion sequences were applied to brain imaging (D Le Bihan et al., 1986) and shown promising preliminary results in studying brain abnormalities, as in the context of stroke, multiple sclerosis or Alzheirmer's disease, (e.g. Miller, Grossman, Reingold, & McFarland, 1998; Sandson, Felician, Edelman, & Warach, 1999; Sevick et al., 1990). In the early 1990s it was also observed that the diffusion of water molecules in biological tissues measured by conventional dMRI was dependent on the spatial direction (Moseley et al., 1990). Since tissues are composed by well oriented fibres, diffusion appears to be higher along the direction parallel to the fibres and lower in

perpendicular direction where water diffusion is more limited by barriers such as the myelin covering the axons on neural tissue.

In 1994, Basser, Mattiello, and LeBihan introduced a new approach in which the diffusion on each image voxel is modelled as an ellipsoid described with a tensor. This technique was latter named diffusion tensor imaging (DTI) and allows a direct examination of anisotropic aspects of tissue microstructure. The spatial invariant measures provided by DTI have been shown to provide important findings in several areas as on studies of cerebral ischemia, leukoariaosis, wallerian degeneration, diffuse axonal injury, epilepsy, alzheimer's disease, tumors, healthy maturation and ageing (Sundgren et al., 2004). As example, for healthy subjects across different range of ages, DTI is sensitive to increases of white matter anisotropy during brain maturation in children and adolescents and sensitive to declines in ageing (M. Moseley, 2002). This suggests that DTI can provide useful information about the brain ageing process.

Although DTI shows remarkable success, this technique has some drawbacks. Since it assumes that the distribution of water molecules as they diffuse across brain tissue is characterised by a Gaussian distribution, it fails to fully characterize diffusion measurements that are inherent to tissue microstructure. While such approximation is valid for free diffusion media that are unrestricted by barriers, this approximation was experimentally demonstrated to be invalid for biological tissues (Assaf & Cohen, 1998; T Niendorf, Dijkhuizen, Norris, van Lookeren Campagne, & Nicolay, 1996). The presence of complex underlying cellular components and microstructures confines the motion of the water molecules and therefore the distribution profile of water molecules displacement shows to have less bounded and more peaked profiles when compared to the Gaussian distribution profile.

Several general approaches have been proposed to characterize diffusion according to the non-Gaussian properties including the multicompartment model (Clark, Hedehus, & Moseley, 2002), statistical diffusion model (Yablonskiy, Bretthorst, & Ackerman, 2003), generalized diffusion tensor (Özarslan & Mareci, 2003), q-space imaging (Tuch, 2004) and diffusion kurtosis imaging (Jensen, Helpern, Ramani, Lu, & Kaczynski, 2005). Among them, q-space imaging is the one that shows better accuracy and ability to fully describe the diffusion profile however it requires long scan time which cannot be available in projects that uses multimodality scanning as the case of the Cam-CAN. Diffusion kurtosis imaging (DKI) overcame this limitation by providing a more practical approach. DKI is an expansion of DTI where the diffusion tensor is estimated

together with the kurtosis tensor which characterizes the degree of non-Gaussianity of diffusion in a 3D frame of reference. This technique provides more accurate values of diffusion parameters relative to DTI (Lu, Jensen, Ramani, & Helpern, 2006) and values of kurtosis have been shown to be useful to quantify the complexity of tissue barriers (Fieremans, Jensen, & Helpern, 2011). DKI has shown promising results in several fields including stroke and ischemia (J. S. Cheung, Wang, Lo, & Sun, 2012; Jensen et al., 2011), hepatocellular carcinoma (Rosenkrantz et al., 2012), attention-deficit hyperactivity disorder (Helpern et al., 2011), the staging of glioblastomas (Raab, Hattingen, Franz, Zanella, & Lanfermann, 2010), rodent brain maturation (M. M. Cheung et al., 2009), and normal ageing (Falangola et al., 2008).

As any dMRI technique, an accurate interpretability of the metrics extracted from the DKI can be challenging and subject to many pitfalls (Jones & Cercignani, 2010). The diffusion and kurtosis estimates can be influenced by noise, motion and image artefacts. In particularly it has been demonstrated that sufficiently large error can cause values that are physically and/or biologically implausible (Tabesh, Jensen, Ardekani, & Helpern, 2011). Therefore to increase the sensitivity and avoid erratic results, it is important to indentify and study the influence of such artefacts and develop more robust methods for data analysis.

## 1.3. Thesis Objectives and Project Plan

The objective of this thesis is to develop robust methods for pre-processing and analysing DKI data for the Cam-CAN project. Secondly it aims to show the potential of this technique in imaging the brain structural changes across lifespan in the Cam-CAN's data recorded so far.

Notably, to improve the robustness of DKI, one can also increase the number of diffusion-weight images (DWIs) required to extract the diffusion and kurtosis tensors. However, since the Cam-CAN project involves the measure of several MRI modalities, the scanning time available for dMRI was limited to 10 minutes. Therefore this thesis is not focus on alternative acquisition schemes. Instead, the quality of DKI metrics will be improved by optimising methods of pre-processing and improving the pipeline analysis of ageing differences across lifespan based on the Cam-CAN's pre-selected acquisition scheme of 63 DWIs (acquired on a 3T Siemens Trio with a twice-refocused-spin-echo, 30

diffusion gradient directions for each b-values 1000 and 2000s.mm$^{-2}$ and three images acquired using b-value 0, TR=9100ms, TE=104ms, voxel size=2x2x2mm$^3$, FOV=192x192mm$^2$, 66 axial slices, number of averages=1).

Figure 1.1 shows the project plan on which the work of this thesis is based. The improvement of DKI methods started by studying different fitting methods to extract diffusion and kurtosis tensor from DWIs (October to December of 2011). After studying the behaviour of this different methods on simulated data corrupted by noise, the more optimal methods is selected to estimate the diffusion and kurtosis metrics on real data transformed with different schemes of pre-processing steps (January to February of 2012), which included methods to correct artefacts, as induced by head motion and eddy currents, and methods aiming to remove MRI noise, as steps of downsampling and regularization. The built procedure for DKI is then applied on diffusion MRI data recorded so far by Cam-CAN, and different analysis methods for studying ageing changes on the diffusion and kurtosis data are studied (March to April). At last, the changes observed on the different methods for ageing changes are analysed according to recent brain hypothesis (May to June).

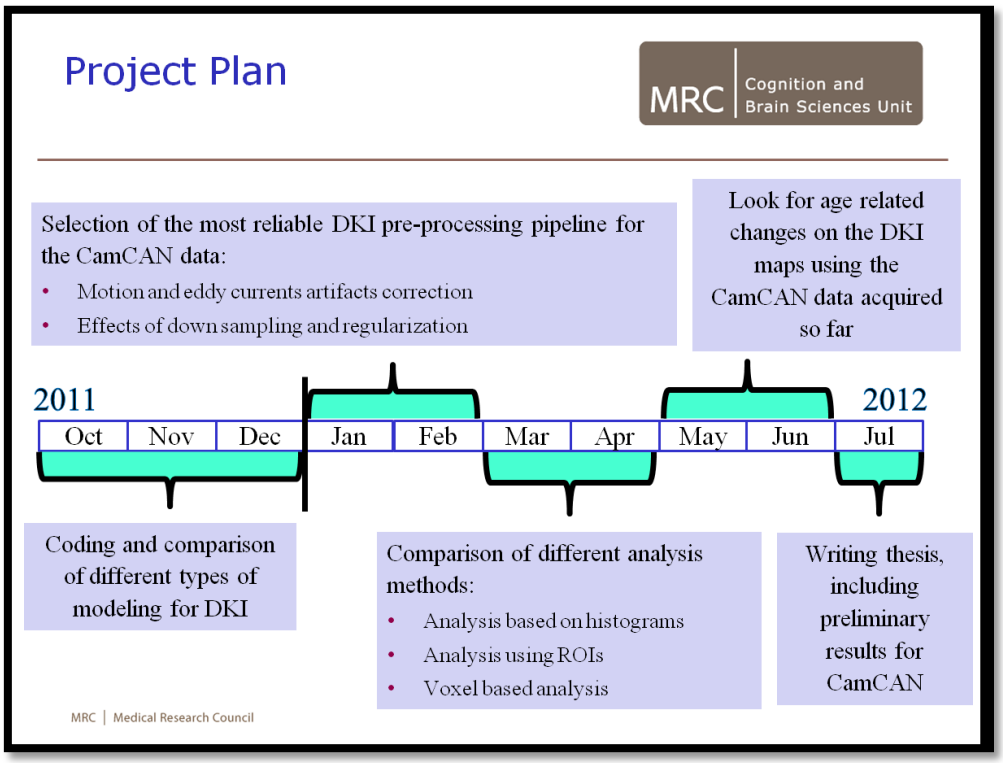



Figure 1.1 – Project plan on which this thesis is based. This figure was presented on the 16th of January CamCAN's groups meeting.

## 1.4. Thesis Outline

After this introduction, the background for this thesis is presented on Chapter 2. This will start with a brief review of the Physics of diffusion (section 2.1). Based on the theory described on the former, section 2.2 will describe how one can use MRI to measure diffusion. The details of diffusion tensor imaging (DTI) is given on section 2.3 which includes its formulation (2.3.1), metrics to extract information from the diffusion tensor (2.3.2), and the noise effects on the accuracy of this technique (2.3.3). The very recent dMRI technique diffusion kurtosis imaging (DKI) is discussed on section 2.4. As for DTI, the details of DKI formulation are given (2.4.1) as well as novel measures that one can use to extract information from the diffusion kurtosis tensor (2.4.2) and its noise influences (2.4.3).

The study of different types of DKI fitting is reported on chapter 3. In this chapter, advances on DKI estimation framework and novel observations of noise artefacts on DKI are described according to the findings of previous studies, as example the published by Tabesh et al (2011) relative to a DKI constrained method that is quickly becoming one of most used methods on recent DKI studies, and the study performed by Jones and Basser (2004) regarding noise influences on DTI.

On chapter 4 the study of reliable pre-processing steps is reported. This will include evidences that algorithms to correct motion artefacts based on registration methods as the ones available in many toolboxes are derisory for dMRI, and result that shows the advantages of using downsampling and regularization algorithms to decrease the impacts of MRI artefacts. In particularly, the optimal value of the size of a Gaussian kernel for noise removal is investigated.

The analysis procedures selected from studying the ageing changes by DKI as well as a description of the preliminary DKI results for the Cam-CAN project are reported in chapter 5. The findings on this chapter will be supported with previous studies, and the impacts of new insights are explored.

On the last chapter (chapter 6), the thesis will be summarized, where the objective accomplished, finding impacts, and future steps are discussed.

# Chapter 2

# Background

## 2.1. Physics of Diffusion

### 2.1.1. Characterization of random translation of particles

Diffusion is a mass process in liquid and gas states which is related to a translation of particles due to thermal energy. This process results in particle mixing without requiring bulk motion as convection and dispersion process (Mori & Barker, 1999). On a molecular level, the thermal energy that each particle carries will result in collisions, and particles will show random translation trajectories known as Brownian motion (Berg, 1993).

In 1905, Albert Einstein introduces the concept of displacement distribution to characterize the Brownian motion of a particle using a probabilistic framework (Einstein, 1956). The displacement distribution quantifies the fraction of particles that travel a specific distance $\Delta\boldsymbol{r}$ within a particular timeframe $t$. As example, for a free medium unrestricted by barriers, the probability that one particle have a displacement $\Delta\boldsymbol{r}$ after a time interval $t$ can be derived from Einstein's framework as a Gaussian distribution:

$$P(\boldsymbol{r}|\boldsymbol{r_0}, t) = (4\pi Dt\,)^{-3/2}\exp\left(-\frac{|\Delta\boldsymbol{r}|^2}{4Dt}\right), \qquad \text{(Eq. 2.1.1)}$$

where $D$ is the diffusion coefficient which depends on intrinsic propriety of the medium as the size and mass of the diffusing particles, the temperature, and the nature of the medium. Media with larger values of diffusion are related to faster dispersion of particles. For example gas states media have larger diffusion coefficient than liquids.

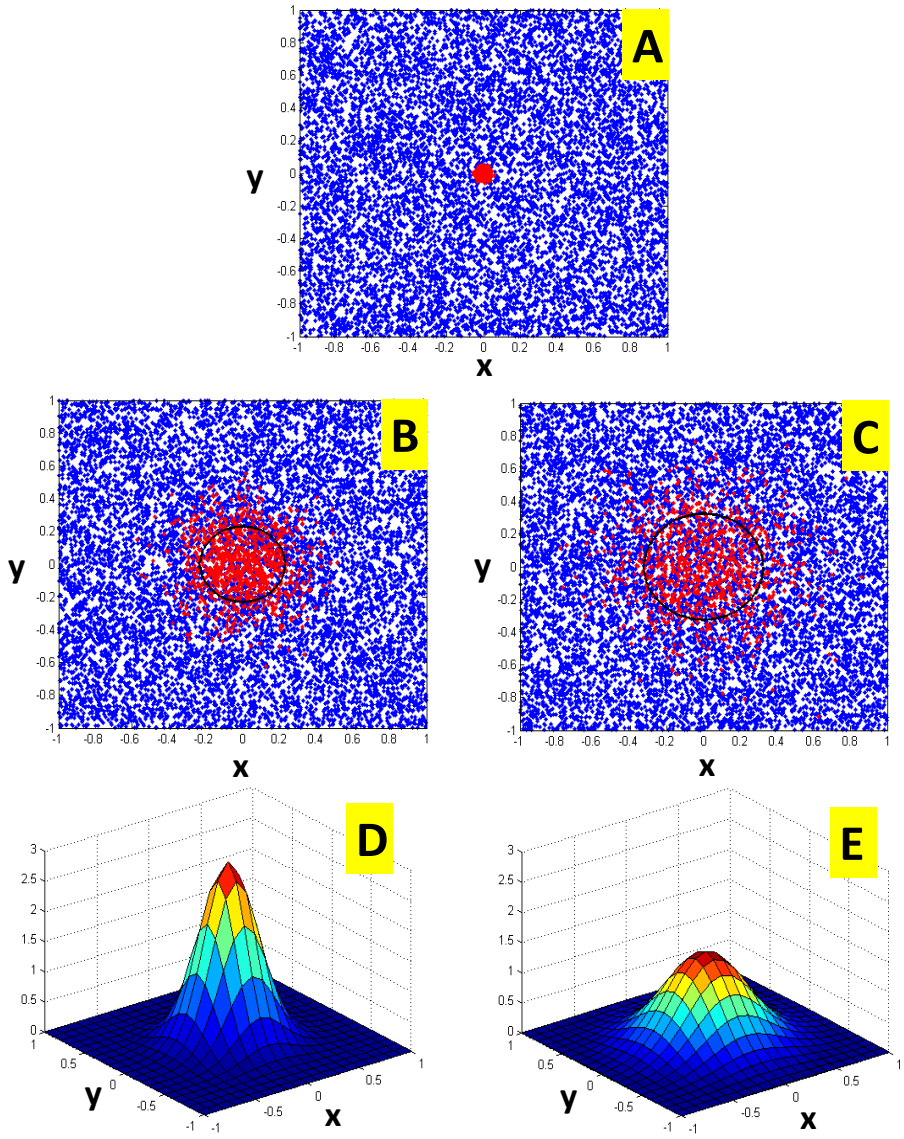



Figure 2.1 – Diffusion process of a two dimensional free medium unrestricted by barriers. Panel A corresponds to an initial state where some central particles were marked in red. After an interval of time *t*, the marked particles will randomly move in all directions (panel B). The displacement of particles during the instant of time between the positions shown in panel A and panel B can be described with a two dimensional Gaussian distribution of probability, where *x* and *y* are the Cartesian components of $\Delta\boldsymbol{r}$ (panel D). After larger intervals of time the particles will randomly translate over larger areas (panel C) and the Gaussian distribution will have broader shape (panel E). The back circles in panel B and C corresponds to diffusion circles of radius $\sqrt{2Dt}$.

Figure 2.1 shows the graphical representation of the displacement distribution for particles diffusing in a two dimensional unrestricted medium. These surfaces are concentric once points with the same probability are defined by diffusion circles (Basser, 1995). For example, setting the exponent of Eq. 2.1.1 by the constant $-1/2$ and assuming that *x* and *y* are the components of the two dimensional version of $\Delta\boldsymbol{r}$, one can deduce the equation of the diffusion circle with radius$\sqrt{2Dt}$ as:

$$x^2 + y^2 = \left(\sqrt{2Dt}\right)^2. \qquad \text{(Eq. 2.1.2)}$$

Eq. 2.1.2 is of importance since it corresponds to the standard deviation $\sigma$ of the displacement distribution. This corresponds to the well-known Einstein formula:

$$\sigma = \sqrt{2Dt}. \qquad \text{(Eq. 2.1.3)}$$

Although Figure 2.1 described diffusion on a two dimensional space, similar observations and derivations can be performed easily for three dimensional cases by assuming three Cartesian components *x*, *y*, and *z* for the displacement vector $\Delta \boldsymbol{r}$. In this case Eq. 2.1.1 will be described by a three dimensional Gaussian distribution, and the points with the sane probability will define surfaces with spherical geometries – the diffusion spheres (Basser, 1995). For example, defining the constant $-1/2$ on the exponent of Eq.2.1.1, the equation of a 3 dimensional diffusion sphere surfaces with radius $\sqrt{2Dt}$ can be written as:

$$x^2 + y^2 + z^2 = \left(\sqrt{2Dt}\right)^2. \qquad \text{(Eq. 2.1.4)}$$

The diffusion sphere described by Eq. 2.1.4 is graphically represented on Figure 2.2.

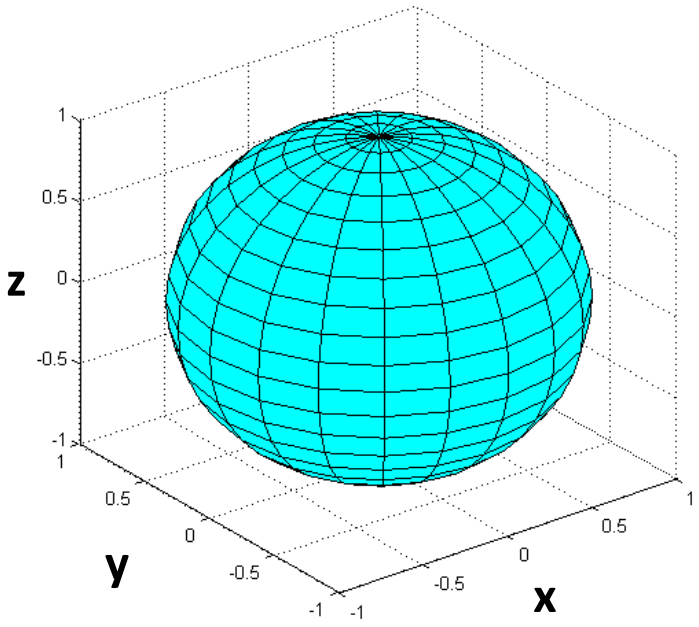



Figure 2.2 – Diffusion sphere showing the surface of a mean square three dimensional displacement for an isotropic medium.

### 2.1.2. Classical diffusions laws and relation with diffusion atomic view

Although, it was only in 1905 that diffusion was described from an atomic perspective, the concept of the diffusion coefficient have already been introduced earlier by Fick (Fick, 1855). Based on extensive investigations on the diffusion of salts in water, Fick stated that under a gradient of concentration $\boldsymbol{\nabla} C$ and assuming a medium with diffusion coefficient *D*, the flux of diffusing particle $\boldsymbol{J}$ can be described as:

$$\boldsymbol{J} = -D\boldsymbol{\nabla} C \qquad \text{(Eq. 2.1.5)}$$

From Eq. 2.1.5 one can deduce a null flux of particles for a null gradient of concentration. However, such situation does not mean that particles are not in influence of diffusion process. According to the law of large numbers of the probability theory (Jaynes, 2003), the number of particles crossing a section of area in one direction will be the same from the opposite (null net flux) once the number of particle are uniformly distributed. In the other hand, if a gradient of concentration is present, one side of the medium delimited by a section of area will have a larger number of particles, and therefore it will be more likely that a larger number of particles from that higher concentrated region will cross the section of area. For instance, in the case described on Figure 2.1, since initially the particles are high concentrated on the centre of the medium, they are more likely to spread in a radial symmetry, and the net flux $\boldsymbol{J}$ will be positive in direction from the centre to the periphery. The migration of the particles will results in a decrease of the concentration gradient, reaching a steady state where the concentration is uniform across the medium. By combining Eq. 2.1.5 with the conservation of mass equation:

$$\nabla . \boldsymbol{J}(\boldsymbol{r}, t) = -\frac{\partial C(\boldsymbol{r},t)}{\partial t} \qquad \text{(Eq. 2.1.6)}$$

one can obtain the expression that describes the variation of concentration of a medium, i.e. the Fick's second law (Wilkinson, 2000):

$$D\boldsymbol{\nabla}^2 \mathrm{C}(\boldsymbol{r}, t) = \frac{\partial C(\boldsymbol{r},t)}{\partial t} \qquad \text{(Eq. 2.1.7)}$$

As mentioned in section 2.1.1, from Einstein's framework the distribution of particle displacement in a free diffusion medium is derived as a Gaussian distribution. Such result is also supported by the second Fick's Law (Jaynes, 1989). As the probability *P* of finding a particle per unit of volume is proportional to the concentration, Eq. 2.1.7 can be rewriting in function of *P*:

$$D\boldsymbol{\nabla}^2\mathrm{P}(\boldsymbol{r},t) = \frac{\partial P(\boldsymbol{r},t)}{\partial t} \qquad \text{(Eq. 2.1.8)}$$

Knowing the initial position $\boldsymbol{r}_0$ of a particle, one can write the function of the distribution of probability in the time instant $t$=0 as an impulse function $\delta$:

$$P(\boldsymbol{r}|\boldsymbol{r_0},0) = \delta(\boldsymbol{r}-\boldsymbol{r_0}) \qquad \text{(Eq. 2.1.9)}$$

Using this initial condition, one can resolve Eq. 2.1.8 as the Gaussian distribution for a free medium unrestricted by barriers:

$$P(\boldsymbol{r}|\boldsymbol{r_0},t) = (4\pi Dt)^{-3/2}\exp\left(-\frac{|\boldsymbol{r}-\boldsymbol{r_0}|^2}{4Dt}\right) \qquad \text{(Eq. 2.1.10)}$$

where $\boldsymbol{r}-\boldsymbol{r_0}$ is the particle displacement $\Delta\boldsymbol{r}$.

### 2.1.3. Diffusion tensor for characterizing anisotropic media

In the previous sections diffusion is characterized by a single scalar coefficient. This approach is sufficient for media where diffusion does not depend in spatial direction, i.e. isotropic diffusion. However this situation is not always valid – for some media, particles can have more difficulty on translating over some specific directions. This is the case of media formed by oriented structures or microstructures that makes diffusion anisotropic (Basser, Mattiello, & LeBihan, 1994a).

An example of anisotropic media is shown in Figure 2.3.A in which particles are confined to a cylindrical shape tube. In that case, particles are free to randomly translate in *x* direction while along the *y* and *z* directions they are limited by the walls of the tube, and thus diffusion is higher along the *x* direction than those along *y* and *z*.

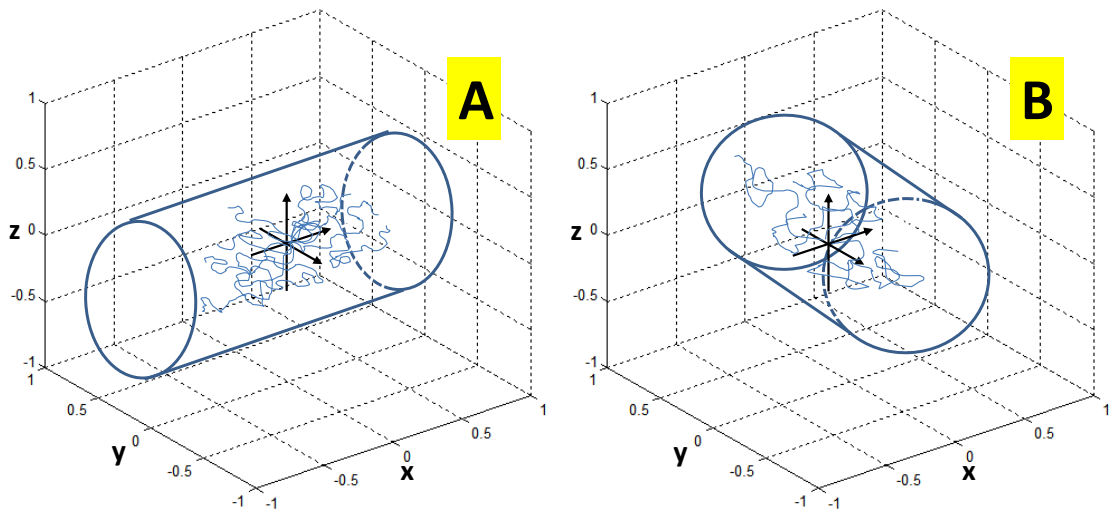



Figure 2.3 – Representation of two media restricted by cylindrical tubes. In Panel A particles are free to translate along direction x, while along x and z they are limited by the boundaries of the tube. In opposite to panel A, panel B is not oriented with the frame of reference x-y-z, and thus particles are free to translate only on an oblique direction.

Defining $D_x$, $D_y$ and $D_z$ as the diffusion coefficient along the axis *x*, *y* and *z*, the first Fick Law for the case described in Figure 2.3.A can be written as:

$$\begin{bmatrix} J_x \\ J_y \\ J_z \end{bmatrix} = -\begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_z \end{bmatrix} . \begin{bmatrix} \partial C/\partial x \\ \partial C/\partial y \\ \partial C/\partial z \end{bmatrix}. \qquad \text{(Eq. 2.1.11)}$$

For more general cases where the cylindrical geometry is not aligned with the frame of reference defined by the axis *x*, *y* and *z* (Figure 2.3.B), Fick's Law is only full described with a 3x3 diffusion tensor (Basser, 1995; Basser et al., 1994a):

$$\begin{bmatrix} J_x \\ J_y \\ J_z \end{bmatrix} = -\begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} . \begin{bmatrix} \partial C/\partial x \\ \partial C/\partial y \\ \partial C/\partial z \end{bmatrix}. \qquad \text{(Eq. 2.1.12)}$$

In a similar way to what was described in the previous section for the isotropic media case, the displacement of distribution of the cylindrical anisotropic media can be derived from Eq. 2.1.12 as:

$$P(\boldsymbol{r}|\boldsymbol{r_0}, t) = \frac{1}{\sqrt{(4\pi t)^3 \det(\boldsymbol{D})}} \exp\left(-\frac{(\boldsymbol{r}-\boldsymbol{r_0})^T \boldsymbol{D}^{-1} (\boldsymbol{r}-\boldsymbol{r_0})}{4t}\right) \qquad \text{(Eq. 2.1.13)}$$

where $\boldsymbol{D}$ is the diffusion tensor:

$$\boldsymbol{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}. \qquad \text{(Eq. 2.1.14)}$$

Similarly to the scalar diffusion coefficient $D$ for isotropic media which have to be defined by a positive number for being physically plausible, the diffusion tensor $\boldsymbol{D}$ have to be defined by a positive matrix (Koay, Carew, Alexander, Basser, & Meyerand, 2006), i.e: (1) $\boldsymbol{D}$ have to be symmetric ($D_{ij} = D_{ji}$), (2) all the diagonal elements of $\boldsymbol{D}$ have to be positive, and (3) all the determinants of all the leading submatrix have to be positive. Due to the symmetric propriety, $\boldsymbol{D}$ can be rewritten only with six independent elements $(D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{xy})$:

$$\boldsymbol{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}. \qquad \text{(Eq. 2.1.15)}$$

As the isotropic diffusion can be geometrically represented on a 3D space with a sphere, the anisotropic diffusion tensor can be represented geometrically by a 3D ellipsoid (Basser, 1995), where the six independent elements of $\boldsymbol{D}$ fully described its size, shape and orientation (see panel A and B of Figure 2.4 which represents the diffusion ellipsoid for the cases described on Figure 2.3). For example by defining the exponent of the displacement distribution on Eq. 2.1.13 to a constant we can obtain the expression (Basser, 1995):

$$\begin{aligned} &\left(D_{yy}D_{zz} - D_{yz}^2\right)x^2 + 2\left(D_{xz}D_{yz} + D_{xy}D_{zz}\right)xy \\ &\quad + (D_{xx}D_{zz} - D_{xz}^2)y^2 + 2\left(D_{xy}D_{yz} - D_{xz}D_{yy}\right)xz \\ &\quad + 2\left(D_{xy}D_{xz} - D_{xx}D_{yz}\right)yz + \left(D_{xx}D_{yy} + D_{xy}^2\right)z^2 = \det(\boldsymbol{D})\, t \end{aligned} \qquad \text{(Eq. 2.1.16)}$$

which can be related with the expression of an ellipsoid:

$$a'x^2 + 2b'xy + d'y^2 + 2c'xz + 2e'yz + f'z^2 =$$

$$= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a' & b' & c' \\ b' & d' & e' \\ c' & e' & f' \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \text{(Eq. 2.1.17)}$$

The length and direction of the three orthogonal principal axes of the diffusion ellipsoids (see Figure 2.4) can be decoupled by the following matrix decomposition of ***D*** (Basser, Mattiello, & LeBihan, 1994b):

$$\begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} . \begin{bmatrix} \varepsilon_{1x} & \varepsilon_{2x} & \varepsilon_{3x} \\ \varepsilon_{1y} & \varepsilon_{2y} & \varepsilon_{3y} \\ \varepsilon_{1z} & \varepsilon_{2z} & \varepsilon_{3z} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1x} & \varepsilon_{2x} & \varepsilon_{3x} \\ \varepsilon_{1y} & \varepsilon_{2y} & \varepsilon_{3y} \\ \varepsilon_{1z} & \varepsilon_{2z} & \varepsilon_{3z} \end{bmatrix} . \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \qquad \text{(Eq. 2.1.18)}$$

where $\lambda_1$, $\lambda_2$ and $\lambda_3$ are the three eigenvalues related to the values of diffusion along the three principal axes of the diffusion ellipsoid, while the triplets of number $(\varepsilon_{1x}, \varepsilon_{1y}, \varepsilon_{1z})$, $(\varepsilon_{2x}, \varepsilon_{2y}, \varepsilon_{2z})$, and $(\varepsilon_{3x}, \varepsilon_{3y}, \varepsilon_{3z})$ are the eigenvectors which gives the direction of each principal axis. In particularly, vector$(\varepsilon_{1x}, \varepsilon_{1y}, \varepsilon_{1z})$ and $\lambda_1$ corresponds to the values of the ellipsoid axis with larger length.

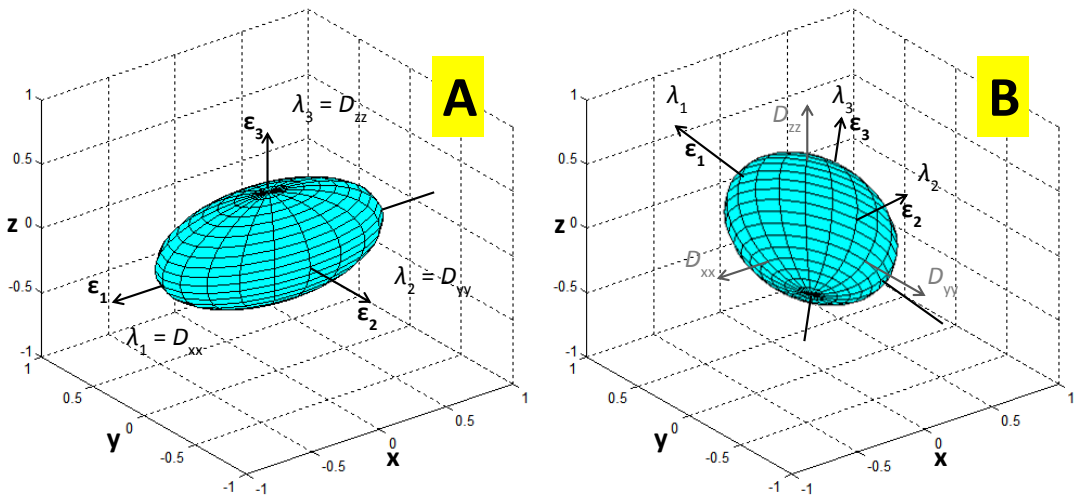



Figure 2.4 – Geometrical representation of the diffusion tensor for the two isotropic media represented in Figure 2.3. Since diffusion ellipsoid in panel A is aligned with the frame of reference x-y-z, the eigenvalues $\lambda_1$, $\lambda_2$, and $\lambda_3$ corresponds to the tensor elements $D_{xx}$, $D_{yy}$, and $D_{zz}$. On the other hand, for panel B, the eigenvalues can be computed by matrix decomposition shown by Eq. 2.1.18.

## 2.2. MRI and Diffusion

In MRI, subjects are exposed to a strong magnetic field and the magnetic moments of the water protons (also referred as spins) start to precess (i.e. rotate) around the direction of magnetic field (Conolly et al., 2000). The magnetic moments can then be described by two components: a longitudinal component aligned to the magnetic field and a transversal component perpendicular to the magnetic field. If within a voxel the spins are precessing in phase, the net transversal component can be measured by the MRI receive coils. To increase the transversal component and have a significant signal to measure, it is applied specific radio frequency (RF) pulses to insure that spins within a voxel are in phase (McRobbie, Moore, Graves, & Prince, 2003). As example, in the commonly used spin eco sequences (Figure 2.5), two RF pulse are applied: 1) a pulse of 90º is applied simultaneously with a gradient to selectivity stimulated the spins of a slice of the body; 2) after the phase and frequency gradient which encodes each voxel with a specific frequency and phase, a 180º pulse is applied to realign the spins within a voxel and reproduce a measurable signal referred as the echo (McRobbie et al., 2003).

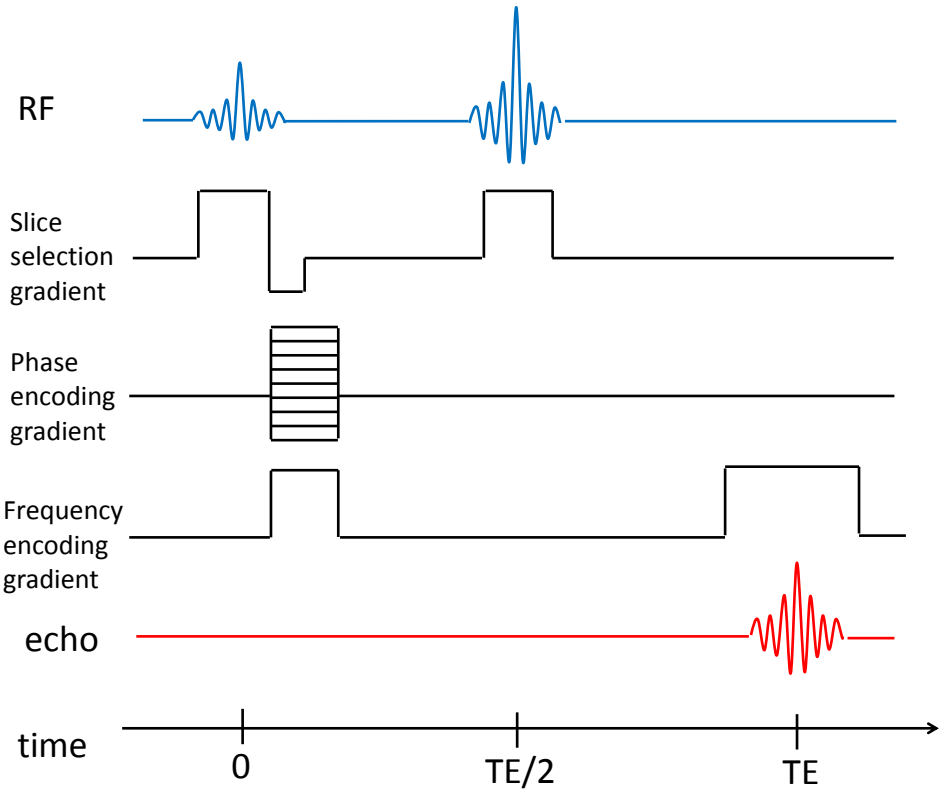



Figure 2.5 – Spin eco sequence

### 2.2.1. Pulsed gradient spin echo (PGSE)

MRI can be used to quantify the diffusion of the brain using proper sequences (Basser et al., 1994a). One basic and commonly diffusion sensitizing gradient is an adaption of the spin eco sequence called pulsed gradient spin echo (PGSE). Relative to the spin eco sequence, PGSE have two extra gradients, one after the 90º RF pulse and other after the 180º, Figure 2.6. The objective of the first additional diffusion gradient is to add a phase shift $d(\varphi)$ that will depend on the spin position *r* relative to the direction of the additional gradient field (Johansen-Berg & Behrens, 2009) as described by the following equation:

$$d(\varphi) = \gamma \int_0^{\delta} Gr\, dt = \gamma \delta G r, \qquad \text{(Eq. 2.2.1)}$$

where $\gamma$ is the gyromagnetic ration which depends on the type of the atomic nuclei (for protons is 42.56 MHz $T^{-1}$), $G$ the intensity of the diffusion gradient and $\delta$ its duration. The second diffusion gradient is applied after a time interval $\Delta$ inducing an opposite phase shift. Such shift has the same amplitude than the previous one since both gradient pulses have the same values of intensity $G$ and duration $\delta$; however it has an opposite effect once applied after the 180º RF pulse which inverted the protons phase for spin realignment. Therefore, and assuming that the duration $\delta$ of each diffusion pulse is much smaller than the time interval $\Delta$, the total shift induced on a spin can be expressed as:

$$d(\varphi) = \gamma G \delta\, dr, \qquad \text{(Eq. 2.2.2)}$$

where $dr$ is the displacement of the proton spin relative to the direction of the diffusion gradient on the time interval $\Delta$.

Eq. 2.2.2 suggest that voxels related to protons that have larger displacements $dr$ will present spins precessing in a ampler range of phase, while voxels that are less conditioned by the diffusion process will have proton with negligible phase shift. MRI receive coils will therefore measure smaller intensities for voxels related to media with larger diffusion coefficients (spins that are precessing out of phase) and larger intensities for voxels related to media with smaller diffusions coefficients (spins possessing in phase).

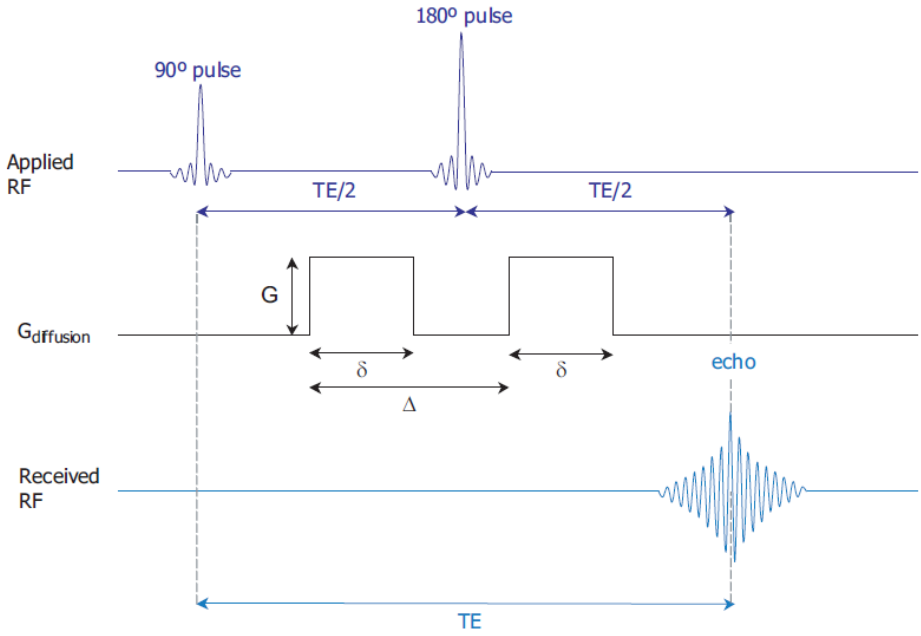



Figure 2.6 – Representation of the two additional diffusion gradients on the pulsed gradient spin eco. This figure was adapted from (Correia, 2009).

The phase shift $d(\varphi)$ described in Eq. 2.2.2 can mathematically be expressed in order to the diffusion coefficient *D*. Given that the total magnetization *M* within a voxel is a sum over all individual magnetic moments $M_0$:

$$M = M_0 \sum_j exp\{id(\varphi)_j\}, \qquad \text{(Eq. 2.2.3)}$$

and the spin population can be described with a displacement distribution of $P(z_f|z_i, \Delta)$, the amplitude of the signal attenuation can be calculated as:

$$S = S_0 \int exp\{id(\varphi)_j\} P(z_f|z_i, \Delta)\, dz_f, \qquad \text{(Eq. 2.2.4)}$$

where $S_0$ is the signal that should be detected if no diffusion gradient is applied. Assuming the Gaussian displacement distribution described by Eq. 2.1.1, Eq. 2.2.4 can be rewrite as:

$$S = S_0 exp\{-(\gamma G\delta)^2 \Delta . D\}. \qquad \text{(Eq. 2.2.5)}$$

It is notable that for Eq. 2.2.5 one can verify that voxels with larger diffusion coefficient are related to lower intensities measured by the MRI receiver coils. Moreover, one can see that larger signal attenuation can also be related to larger water particles displacements if diffusion-weighted signal is measured using larger diffusion time $\Delta$ values. Gradients with larger amplitude $G$ and duration $\delta$ will also highlight the diffusion contrast, once they induce larger phase shift on diffusing water molecules.

### 2.2.2. The Stejskal-Tunner equation

The more general relationship between the diffusion coefficient and the signal measured in diffusion MRI was established in 1965 by Stejskal and Tunner. Defining the applied diffusion gradient field $\boldsymbol{G}(t)$ as:

$$\boldsymbol{G}(t) = \left(G_x(t), G_y(t), G_z(t)\right)^T, \qquad \text{(Eq. 2.2.6)}$$

and describing an arbitrary function $\boldsymbol{F}(t)$ as:

$$\boldsymbol{G}(t) = \int_0^t \boldsymbol{F}(t')dt', \qquad \text{(Eq. 2.2.7)}$$

the MRI measured diffusion-weighted signal *S* can be expressed with the Stejskal-Tunner equation as:

$$\ln[S/S_0] = -\gamma^2 \int_0^{TE} \left(\boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right)^T \cdot D\left(\boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right)dt', \qquad \text{(Eq. 2.2.8)}$$

where $S_0$ is the MRI measured signal when the diffusion gradient is not applied, $H(t)$ the Heaviside function, *TE* the time interval between the instant that MRI sequence was applied (*t=0*) to the instant when the echo is recorded (see Figure 2.5), and *D* the scalar diffusion coefficient. By defining the parameters that are only dependent on the diffusion gradient by a single constant *b* (known also as b-value):

$$b = \gamma^2 \int_0^{TE} \left( \boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right)^T \cdot \left(\boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right) dt', \qquad \text{(Eq. 2.2.9)}$$

the notation of Eq .2.2.5 can be simplified as:

$$S = S_0 exp\{-b.D\}. \qquad \text{(Eq. 2.2.10)}$$

Remarkably, Eq. 2.2.10 is of importance because it shows how easily one can estimate a diffusion coefficient. *D* can be computed on each voxel by linearization of Eq. 2.2.10:

$$D = -\frac{1}{b}\ln\left(\frac{S}{S_0}\right). \qquad \text{(Eq. 2.2.11)}$$

where values of *S* can be acquired with a diffusion-weighted image (DWI) while values of $S_0$ acquired by setting the diffusion gradient to zero (i.e. b-value equal zero). Moreover, the b-value can easily be computed off-line with the known parameters of the diffusion gradient sequence using Eq. 2.2.9 (Mattiello, Basser, & LeBihan, 1994). For example, assuming the PGSE on Figure 2.6, the expression for the b-value can be derived from Eq. 2.2.9 as:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{1}{3}\delta\right). \qquad \text{(Eq. 2.2.12)}$$

For a PGSE acquisition where gradient duration $\delta$ is much smaller that the time interval Δ, Eq. 2.2.12 can be even simplified as:

$$b = \gamma^2 G^2 \delta^2 \Delta\,. \qquad \text{(Eq. 2.2.13)}$$

Notably, Eq. 2.2.13 corresponds to the case described in section 2.2.1, and therefore if we apply this equation to Eq. 2.2.10, Eq. 2.2.5 is rewritten.

### 2.2.3. Twice Refocused Spin Echo (TRSE)

Since an MRI acquisition requires time-varying gradients, non-desired electric currents (named as eddy currents) are induced on the nearby conductors. This will generate local magnetic field gradients that will be either added or subtracted from the phase and frequency encoding gradients, and therefore MRI images will show spatial distortions (Jones & Cercignani, 2010). This artefact is not problematic in most of MRI modalities since the rising and falling of the encoding gradients are normally close in time which induces a cancellation of eddy currents. However, due to limitations of the gradient amplitude $G$ on the current clinical scanners, dMRI requires diffusion gradients with larger duration $\delta$ to achieve b-values that provides robust estimations of diffusion coefficient (normally around 1000 s.mm$^{-2}$). Therefore for dMRI the rising and falling of the diffusion gradient are not close in time and eddy current distortions are significant.

The Twice Refocused Spin echo (TRSE) sequence was developed with the aim of reducing eddy distortions at no cost in scanning efficiency or effectiveness (Reese, Heid, Weisskoff, & Wedeen, 2003). TRSE uses two 180º pulses (refocusing pulses) and the effects of the eddy currents are cancelled with two bipolar diffusion gradients of length $\delta_1 + \delta_2$ and $\delta_3 + \delta_4$, Figure 2.7

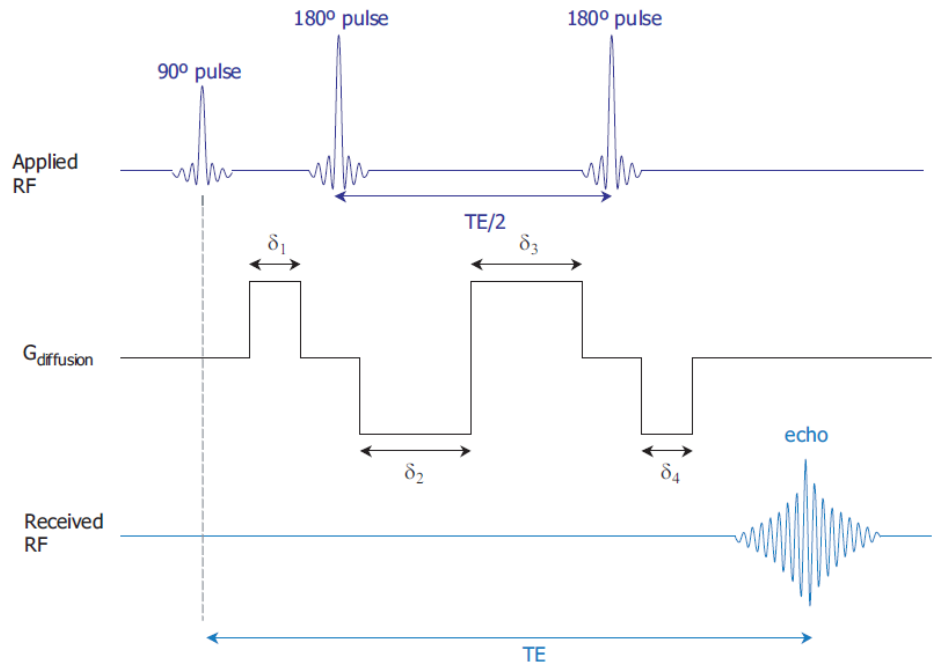



Figure 2.7 – Representation of the twice refocused spin echo sequence. This figure was adapted from (Correia, 2009).

## 2.3. Diffusion Tensor Imaging

In neural tissues, the water molecules and metabolites are under the influence of diffusion processes. They are constantly bouncing, crossing and interacting with many tissue components as cell membranes, cytoskeleton and macromolecules (Johansen-Berg & Behrens, 2009). Diffusion process depends on microstructure, being sensitive to the number, orientation, and permeability of barriers as myelin, and to the presence of various cell types of organelles as dendrites, axons, neurofilaments and microtubes (Denis Le Bihan et al., 2001). Furthermore, since the microstructures are highly oriented, diffusion depends on spatial direction (M. E. Moseley et al., 1990). For example, in brain white matter the water molecules will have higher mobility in the direction of the fibres tracks than across the myelinated barriers of the neural axons, i.e. diffusion in the direction of the fibres is larger than in perpendicular directions (Basser, 1995). In opposite, diffusion will show isotropic proprieties in grey matter once this type of tissue does not consist on oriented structures.

As mentioned in section 2.1.3, anisotropy in media with cylindrical geometry can be fully described using a diffusion tensor. Assuming that neural tissue consist of white matter tracts which can be modelled as cylindrical tubes, such approach can be useful in modelling the dependence of diffusion on biological tissues (Mori & Barker, 1999).

Proposed by Basser et al. 1994, diffusion tensor imaging (DTI) uses the information of diffusion-weighted images (DWIs), obtained with different diffusion gradients, to provide a non-invasive estimation of the diffusion tensor for all the voxels of a brain volume (Basser et al., 1994a). Over almost two decades since it was development, DTI had become an established technique which had shown to provide subtle information of neural structures for several clinical studies as cerebral ischemia, brain maturation, traumatic brain injury, epilepsy, multiple sclerosis and Alzheimer's disease (for a review of several clinical applications of DTI see Sundgren et al., 2004). Moreover, DTI can provide information for algorithms of fibre tracking which can be an important technique in studying brain connectivity (Fillard et al., 2011).

### 2.3.1. The formulation beyond DTI

DTI was based on the adaption of the Stejskal-Tanner equation for anisotropic mediums (Stejskal, 1965):

$$\ln[S/S_0] = -\gamma^2 \int_0^{TE} \left( \boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right)^T \cdot \boldsymbol{D}\left(\boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right) dt' , \quad \text{(Eq. 2.3.1)}$$

where $S$ and $S_0$ are the MRI measured signal with and without applying an diffusion gradient, $H(t)$ the Heaviside function, $F(t)$ the function defined by Eq. 2.2.7, *TE* the time interval between the instant that MRI sequence is applied (*t=0*) to the instant that the echo is recorded, and $\boldsymbol{D}$ the diffusion tensor. By analogy with the scalar b-value in section 2.2.2, it is possible to define a matrix $\boldsymbol{B}$ which only depends on the parameters of the diffusion gradient (Basser et al., 1994a):

$$\mathbf{B} = \gamma^2 \int_0^{TE} \left( \boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right) \cdot \left(\boldsymbol{F}(t') - 2H\left(t' - \frac{TE}{2}\right)\boldsymbol{F}\left(\frac{TE}{2}\right)\right)^T dt', \quad \text{(Eq. 2.3.2)}$$

therefore, Eq 2.3.1 can be rewritten as:

$$S = S_0 exp\{-\boldsymbol{B}:\boldsymbol{D}\}, \quad \text{(Eq. 2.3.3)}$$

where ':' is the generalized dot product. Matrix $\boldsymbol{B}$ is a symmetric 3x3 matrix:

$$\boldsymbol{B} = \begin{bmatrix} b_{xx} & b_{xy} & b_{xz} \\ b_{xy} & b_{yy} & b_{yz} \\ b_{xz} & b_{yz} & b_{zz} \end{bmatrix} . \quad \text{(Eq. 2.3.4)}$$

As the scalar b-value, $\boldsymbol{B}$ can be computed off-line (Mattiello et al., 1994). For example, for a PGSE gradient $\boldsymbol{G}(t) \rightarrow \left[G_x(t), G_y(t), G_z(t)\right]$ with duration $\delta$ and a time interval between diffusion pulses Δ, the elements $b_{ij}$ of the $\boldsymbol{B}$ matrix can expressed as:

$$b_{ij} = \gamma^2 G_i G_j \delta^2 \left(\Delta - \frac{1}{3}\delta\right). \quad \text{(Eq. 2.3.5)}$$

Alternatively, the matrix $\boldsymbol{B}$ can also be computed from the scalar b-value and then combined with an 3x3 matrix containing the gradient direction information, as shown in the following equation:

$$\boldsymbol{B} = b.\begin{bmatrix} n_x n_x & n_x n_y & n_x n_z \\ n_x n_y & n_y n_y & n_y n_z \\ n_x n_z & n_y n_z & n_z n_z \end{bmatrix} \qquad \text{(Eq. 2.3.6)}$$

where $\boldsymbol{n} = (n_x, n_y, n_z)$ is the unit vector that describes the direction of the diffusion gradient $\boldsymbol{G}$.

Since $\boldsymbol{D}$ and $\boldsymbol{B}$ are symmetric matrixes, Eq. 2.3.3 can be simplified and expanded in function of individual $D_{ij}$ and $b_{ij}$ elements as:

$$S = S_0 \exp\{ - b_{xx}D_{xx} - b_{yy}D_{yy} - b_{zz}D_{zz} - 2b_{xy}D_{xy} - 2b_{xz}D_{xz} - 2b_{yz}D_{yz}\}, \qquad \text{(Eq. 2.3.7)}$$

or in order of the scalar b-value and the gradient direction as:

$$S(\boldsymbol{n}, b) = S_0 \exp\left\{-b \sum_{i=1}^{3}\sum_{j=1}^{3} n_i n_j D_{ij}\right\}$$
$$= S_0 \exp\{ - bn_1n_1D_{11} - bn_2n_2D_{22} - bn_3n_3D_{33} - 2bn_1n_2D_{12} - 2bn_1n_3D_{13} - 2bn_2n_3D_{23}\}, \qquad \text{(Eq. 2.3.8)}$$

where index 1, 2, and 3 correspond to the spatial dimensions *x*, *y* and *z*.

The six independent elements of the diffusion tensor can be extracted from at least $N$=7 DWIs. This 7 images have to be recorded for at least six non-collinear diffusion gradients directions and at least 2 different b-values which can include b=0 s.mm$^{-2}$.

The estimation of diffusion parameters from DWIs can be achieved by different methods (Koay et al., 2006) which varies from linear least squares approaches to non-linear ones. Moreover depending on the robustness of each method, one can add constrains which can force the diffusion metrics to physically and biologically plausible ranges or weights to correct non-uniformity of diffusion-weighted signal variance. Examples of some of these approaches are described below.

**Linear least squares**

Eq. 2.3.3 can be converted to a linear framework by taking the logarithm on each side of the equation (Basser et al., 1994b). For example, from the expansion described on Eq. 2.3.8 we can write:

$$\ln S(\boldsymbol{n}, b) = \ln S_0 - bn_1n_1D_{11} - bn_2n_2D_{22} - bn_3n_3D_{33} \\ -2bn_1n_2D_{12} - 2bn_1n_3D_{13} - 2bn_2n_3D_{23}\ . \qquad \text{(Eq. 2.3.9)}$$

By defining a column vector $\boldsymbol{L}$ with all the logarithms of the $N$ measures of $S(\boldsymbol{n}, b)$, writing the $N \times 7$ matrix $\boldsymbol{A}$ with the information of the b-values and gradients directions used in each $\boldsymbol{N}$ measurements as:

$$\boldsymbol{A} = \\ b.\begin{bmatrix} b^{(1)}n_1^{(1)}n_1^{(1)} & b^{(1)}n_2^{(1)}n_2^{(1)} & b^{(1)}n_3^{(1)}n_3^{(1)} & 2b^{(1)}n_1^{(1)}n_2^{(1)} & 2b^{(1)}n_1^{(1)}n_3^{(1)} & 2b^{(1)}n_2^{(1)}n_3^{(1)} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b^{(N)}n_1^{(N)}n_1^{(N)} & b^{(N)}n_2^{(N)}n_2^{(N)} & b^{(N)}n_3^{(N)}n_3^{(N)} & 2b^{(N)}n_1^{(N)}n_2^{(N)} & 2b^{(N)}n_1^{(N)}n_3^{(N)} & 2b^{(N)}n_2^{(N)}n_3^{(N)} & 1 \end{bmatrix} \qquad \text{(Eq. 2.3.10)}$$

and defining the vector $\boldsymbol{X}$ which elements are the diffusion tensor elements and the logarithm of $S_0$:

$$\boldsymbol{X} = [D_{11} \quad D_{22} \quad D_{33} \quad D_{12} \quad D_{13} \quad D_{23} \quad \ln S_0\ ]^T\ , \qquad \text{(Eq. 2.3.11)}$$

Eq. 2.3.9 can be rewritten as:

$$\boldsymbol{L} = \boldsymbol{AX}. \qquad \text{(Eq. 2.3.12)}$$

From Eq. 2.3.12 the unknown parameters on $\boldsymbol{X}$ can be estimated from the standard linear least squares solution:

$$\widehat{\boldsymbol{X}} = [\boldsymbol{A}^T\boldsymbol{A}]^{-1}\boldsymbol{A}^T\boldsymbol{L} = \boldsymbol{A}^+\boldsymbol{L} \qquad \text{(Eq. 2.3.13)}$$

where $\widehat{\boldsymbol{X}}$ is the estimation of vector $\boldsymbol{X}$ and $\boldsymbol{A}^{+}$ the pseudoinverse of matrix $\boldsymbol{A}$. This approach was followed on several studies due to its implementation simplicity and for providing algorithms with high computation speed efficiency (Jones & Cercignani, 2010).

**Weighted Linear least squares**

The linear least squares solution described by Eq. 2.3.13 assumes that the elements of vector $\boldsymbol{L}$ have errors with independent and identically distributed variance. In fact, when signal to noise ratio is sufficiently high, measures of the diffusion-weighted signal can be assumed to have a uniform variance. However, it is important to note that using a linear framework we are not using the diffusion-weighted signals but their log-transformed versions. For standard error propagation techniques (Jones & Cercignani, 2010) one can deduce that assuming diffusion-weighted signals with uniform variance of $\sigma_{DW}^2$, the variance of the log-transformed version $\sigma_{\ln DW}^2$ can be described by:

$$\sigma_{\ln DW}^2 = \frac{\sigma_{DW}^2}{DW^2}, \qquad \text{(Eq. 2.3.14)}$$

where $DW^2$ is the noise free version of $S$. Approximating the exact value of $DW$ by the measured value $S(\boldsymbol{n}, b)$ and measuring $\sigma_{DW}^2$ from noise power estimation algorithms, (e.g. Aja-Fernandez, Alberola-Lopez, & Westin, 2008), one can compute an covariance matrix $\boldsymbol{\Sigma}$ where diagonal elements are the values of $\sigma_{\ln DW}^2$. The information of matrix $\boldsymbol{\Sigma}$ can be incorporated on the estimation of $\boldsymbol{X}$ using the more general linear least squares solution (Jones & Cercignani, 2010) given by the following equation:

$$\widehat{\boldsymbol{X}} = \left[\boldsymbol{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\right]^{-1}\boldsymbol{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{L}\,. \qquad \text{(Eq. 2.3.15)}$$

This approach should provide a more accurate estimation once each log-transformed diffusion-weighted signal is weighted by its variance.

**Non-Linear least squares**

For the non-linear least squares solution, the tensor is estimated directly from Eq. 2.3.3 using well established methods as the Levenberg-Marquardt non-linear regression (Jones & Basser, 2004). Such non-linear approaches involve iterative algorithms where the

elements of diffusion tensor are adjusted till the errors between the predicted signals and the measured ones are minimized. The objective function to be minimized can be described as:

$$f(\boldsymbol{X}) = \frac{1}{2}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}\left(S(\boldsymbol{n}^i, b) - S_0 \exp\sum_{j=1}^{6} A_{ij}X_j\right)^{\boldsymbol{2}} \qquad \text{(Eq. 2.3.16)}$$

Since this non-linear fit deals directly with the diffusion-weighed signals, the variance of the measured insets should be uniform (if signal to noise ratio is sufficiently high) even without applying any kind of approximation as assumed by the weighted linear solution described in Eq. 2.3.15. Therefore, it is expected that non-linear least squares approaches will have a better accuracy than the ones mentioned above. Nevertheless, iterative algorithm can be disadvantageous for being very computationally demanding. A strategy to reduce such disadvantage is to initialize the iterative process with results obtained from Eq. 2.3.13 or 2.3.15. This also have the advantages of avoiding convergence to a global minimum (Jones & Cercignani, 2010).

**Constrained linear and non-linear least squares**

Constrains can be applied in the minimization approach to avoid that data influenced by noise motion and artefacts result on physical or biological implausible values of diffusion (Koay et al., 2006). This can be applied on the non-linear approach described above (Eq. 2.3.16) or on the linear framework by minimization the following function:

$$f(\boldsymbol{X}) = \boldsymbol{\frac{1}{2}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}\left(\ln\frac{S(\boldsymbol{n}^i, b)}{S_0} - \sum_{j=1}^{6} A_{ij}X_j\right)^{\boldsymbol{2}} \qquad \text{(Eq. 2.3.17)}$$

However, when applying constrains both linear and non-linear techniques will be computationally demanding once they both require iterative approaches or non-linear formulations. For example, constrains can be incorporated on each step of an iterative algorithm so that $\boldsymbol{X}$ is forced to converges to a solution according to imposed restrictions. Alternatively the diffusion tensor can be directly constrained on the DTI formulation matrix using for example the Cholesky parameterization (Koay et al., 2006). Assuming

that all and only positive matrices can be decomposed into the product of a lower triangular matrix, the positive diffusion tensor can be decomposed as:

$$\begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{bmatrix} = \begin{bmatrix} R_0 & 0 & 0 \\ R_3 & R_1 & 0 \\ R_5 & R_4 & R_2 \end{bmatrix} . \begin{bmatrix} R_0 & R_3 & R_5 \\ 0 & R_1 & R_4 \\ 0 & 0 & R_2 \end{bmatrix} =$$

$$\begin{bmatrix} {R_0}^2 & R_0R_3 & R_0R_5 \\ R_0R_3 & {R_1}^2 + {R_3}^2 & R_1R_4 + R_3R_5 \\ R_0R_5 & R_1R_4 + R_3R_5 & {R_2}^2 + {R_4}^2 + {R_5}^2 \end{bmatrix} . \quad \text{(Eq. 2.3.18)}$$

Consequently the optimization problem on Eq. 2.3.16 or Eq. 2.3.17 can be rewritten in a way that elements $R_i$ are iteratively estimated, instead of the elements $D_{ij}$. Diffusion tensor will then be estimated using Eq. 2.3.18 and since computed by the product of a lower triangular matrix it will mandatory be defined with a positive matrix, reflecting physically plausible values.

### 2.3.2. DTI invariant Measures

From the estimated diffusion tensor, several rotationally invariant scalar measures can be calculated. These parameters can be as important as identifying fibre direction (Pierpaoli & Basser, 1996). Examples of invariant parameters are the mean diffusivity (*MD*), the fractional anisotropy (*FA*), axial diffusivity ($\lambda_\parallel$), and radial diffusivity ($\lambda_\perp$).

Mean diffusivity can be calculated as Eq. 2.3.19 and gives an overall of the mean-squared displacement of molecules i.e. average of diffusion ellipsoid size (Denis Le Bihan et al., 2001):

$$MD = \frac{\mathrm{Tr}(\boldsymbol{D})}{3} = \frac{D_{11}+D_{22}+D_{33}}{3} = \langle\lambda\rangle = \frac{\lambda_1+\lambda_2+\lambda_3}{3} \quad \text{(Eq. 2.3.19)}$$

where 'Tr' is the matrix trace.

The axial diffusivity gives the magnitude of the diffusion along the principal ellipsoid component, so it is equal to $\lambda_1$:

$$\lambda_\parallel = \lambda_1 \quad \text{(Eq. 2.3.20)}$$

For example in white matter, since the first principal axis $\boldsymbol{\varepsilon_1}$of the diffusion ellipsoid is align with the fibre direction, $\lambda_\parallel$ is of interest because corresponds to the diffusion value

along the direction of fibres, and thus it is believed to be an index of axonal integrity (Helpern et al., 2011).

In the other hand radial diffusivity is the mean diffusion along the orthogonal components relative to the principal ellipsoid axis, which can be estimated as the average between $\lambda_2$ and $\lambda_3$:

$$\lambda_{\perp} = \frac{\lambda_2 + \lambda_3}{2} \qquad \text{(Eq. 2.3.21)}$$

In opposite to the axial diffusivity, in white matter $\lambda_{\perp}$ is related to the average of the diffusion values in the directions across the fibre barriers which consists on structures as the axonal myelin, and thus it is believed to be an index of myelin integrity (Helpern et al., 2011).

Fractional anisotropy (Pierpaoli & Basser, 1996) measures the anisotropy in a range between 0 (case of isotropic diffusion, $\lambda_1 = \lambda_2 = \lambda_3$) to 1(fully anisotropic $\lambda_1 >> \lambda_2 = \lambda_3 = 0$):

$$FA = \frac{\sqrt{3[(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2]}}{\sqrt{2(\lambda_1{}^2 + \lambda_2{}^2 + \lambda_3{}^2)}} \quad , \qquad \text{(Eq. 2.3.22)}$$

therefore structures that give diffusion high anisotropy proprieties, as single oriented populations of white matter fibres, are related to FA values close to one, while brain regions which are not oriented on specific direction, as grey matter regions, are related to low or even null values of FA.

### 2.3.3. Effects of noise on DTI

MRI as any other imaging technique is influenced by noise (Sijbers, den Dekker, Scheunders, & Van Dyck, 1998). The complex valued data acquired by MRI coils are corrupted by noise that is typically well described by a Gaussian distribution on both real and imaginary space. When the complex data in K-space is converted to an image in spatial domain using Fourier reconstruction, the Gaussian noise in real space $g_r$ and imaginary space $g_i$ is converted to Rician noise which corrupted intensity $R$ of each voxel image is given by:

$$R = \sqrt{\left(\frac{S}{\sqrt{2}} + g_r\right)^2 + \left(\frac{S}{\sqrt{2}} + g_i\right)^2} \qquad \text{(Eq. 2.3.23)}$$

being $S$ the noise free intensity.

The DWIs as MRI images are also contaminated by Rician noise, and therefore, for a correct interpretation of the data extracted from DTI, it is important to know the influences and impacts of noise on the estimates of the diffusion tensor and on the invariant metrics extracted. In particularly, noise can be responsible to biased or even physically implausible estimates. Physically implausible values of diffusion occur when corrupted diffusion-weighted intensity $S$ for a given b-value is larger than the diffusion-weighted intensity acquired using a smaller b-value. In such case, the mono-exponential decay of $S$ with the increase of the b-value is not satisfied; instead the noisy outlier suggests an increase which corresponds to a negative directional diffusion coefficient. To exemplify this artefact, four simulations were preformed which are shown in Figure 2.8. Panel A and B corresponds to the expected directional diffusion coefficient value of $0.1\times10^{-3}$ $mm^2.s^{-1}$, while panel C and D corresponds to the expected directional diffusion coefficient value of $1.0\times10^{-3}$ $mm^2.s^{-1}$. Diffusion-weighted signals were simulated based on b-values of 0 and 1000 $s.mm^{-2}$ for panels A and C, and 0 and 3000 $s.mm^{-2}$ for panels B and D. After corrupted with Rician noise using Eq. 2.3.23, the diffusion-weighted signals for each pair of b-values were marked with blue, if a positive value re-estimate of diffusion is obtained, or marked with red if the diffusion estimate had an implausible negative value (diffusion coefficient re-estimates are computes using Eq. 2.2.11). The histograms of the estimates of diffusion coefficient are plotted on the bottom of each panel.

From the results, one can note that implausible estimates are present on panels A and B, i.e. for lower diffusions, and are more frequent for lower b-values, panel A. Interesting, despite not showing negative implausible values, diffusion on panel C and D seems to have large proportions of underestimated values. This underestimations were previously reported in 2004 by Jones and Basser (Jones & Basser, 2004) on the study of the rectified noise effects on DTI. Rectified noise corresponds to the noise on voxel where noise free intensity should be zero. On DWIs, this case corresponds to background voxels or voxels where the diffusion-weighted signal had completely decay. As the MRI intensity is always a positive number, the Rician error in such voxels will raise to a minimum signal measurable. Consequently, while high diffusion-weighted signals (which are related to low b-values or low diffusions) are unaffected, diffusion-weighted

signals that are close to the noise floor will be overestimated, and a directional diffusion coefficient extracted for such data will be underestimated.

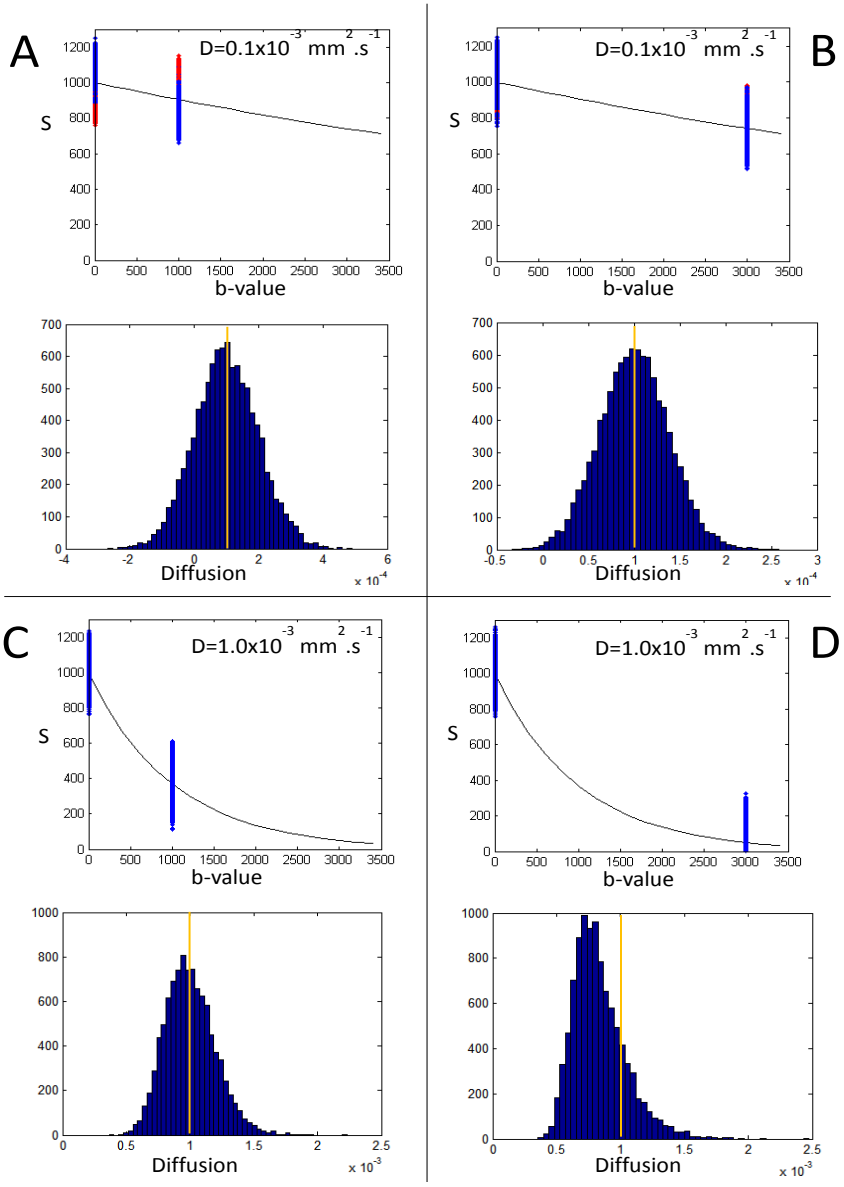



Figure 2.8 – Plots of diffusion-weigthed intensity as a function of b-values, for two different values of directional diffusivity: $0.1x10^{-3}$ $mm^2.s^{-1}$ (panels A and B) and $1.0x10^{-3}$ $mm^2.s^{-1}$(panels C and D). Blue points are Rician noise corrupted diffusion-weighted intensities (SNR 15). The histograms on each panel correspond to diffusion estimated from two pairs of b-values points. The pairs of intensities that resulted in negative values of diffusion are plotted in red. Panels A shown relative large amounts of implausible values, while panel D diffusion estimated that are bias by rectified noise. The values of this stimulation were based on the values used by (Jones & Basser, 2004).

Assuming a white matter voxel with a single oriented population of fibres, one can note that rectified noise will have a larger effect on the direction parallel to the fibres once this direction corresponds to larger diffusion. Thus, for b-values significantly high, the first principal eigenvalue $\lambda_1$ will be underestimated while $\lambda_2$ and $\lambda_3$ unaffected. In such

way, anisotropy will be apparently reduced and values of FA will be underestimated (Jones & Basser, 2004).

Another well document artefact is the eigenvalue repulsion (Pierpaoli & Basser, 1996). Since the decomposition of the diffusion tensor to eigenvalues and eigenvectors (Eq. 2.1.18) involves a sorting algorithm (i.e larger eigenvalues are classified as the value of $\lambda_1$, while the smaller value to $\lambda_3$), a systematic sampling bias is introduced which makes isotropic structures appear more anisotropic (Pierpaoli & Basser, 1996). Consequently, isotropic media as the grey matter will show an erratic anisotropy (Figure 2.9). In the anisotropic case of single oriented population of fibres on white matter, since $\lambda_1$ is much larger than $\lambda_2$ and $\lambda_3$, the likelihood of misclassification of $\lambda_1$ is lower. However there is still a bias between the classification of $\lambda_2$ and $\lambda_3$, and therefore anisotropy is still overestimated.

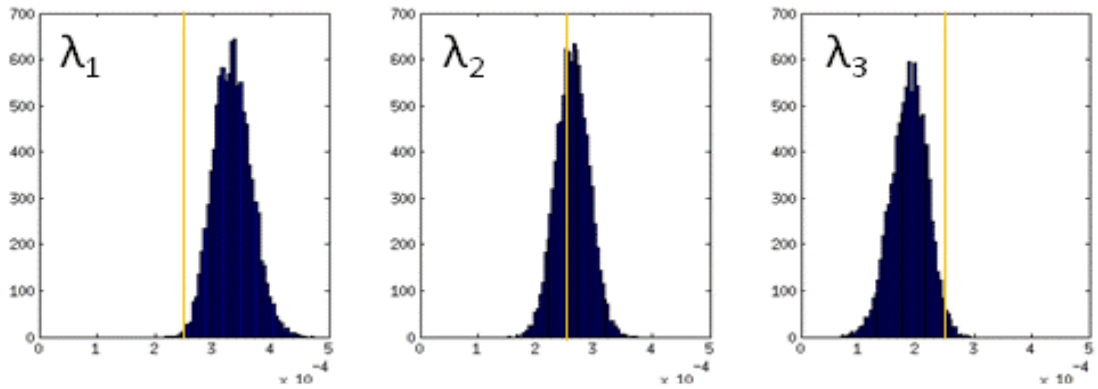



Figure 2.9 – Histograms of the estimates of $\lambda_1$, $\lambda_2$, and $\lambda_3$ for an isotropic diffusion voxel. Although all noise free eigenvalues were equal ($2.5 \times 10^{-4}$ $mm^2.s^{-1}$), estimates of $\lambda_1$ shown the overestimated values while $\lambda_3$ the underestimated ones. Estimates were extracted from data corrupted with Rician noise (SNR=15) using the ordinary linear less squares (Eq. 2.3.13) and based on pairs of b-values 0 and 1000 $s.mm^{-2}$ over 30 directions.

From the last two paragraphs, one can note that FA is underestimated by the effects of rectified noise and overestimated by the repulsion effects. Since these artefacts depend on b-value, directional diffusion and anisotropy, FA will be overestimation or underestimation depending on such factors. As example, Figure 2.10 shows the dependence of estimated of FA with the b-value for a fixed diffusion tensor and signal to noise ratio, where overestimations are predominant from lower b-values while underestimations are present for higher b-values.

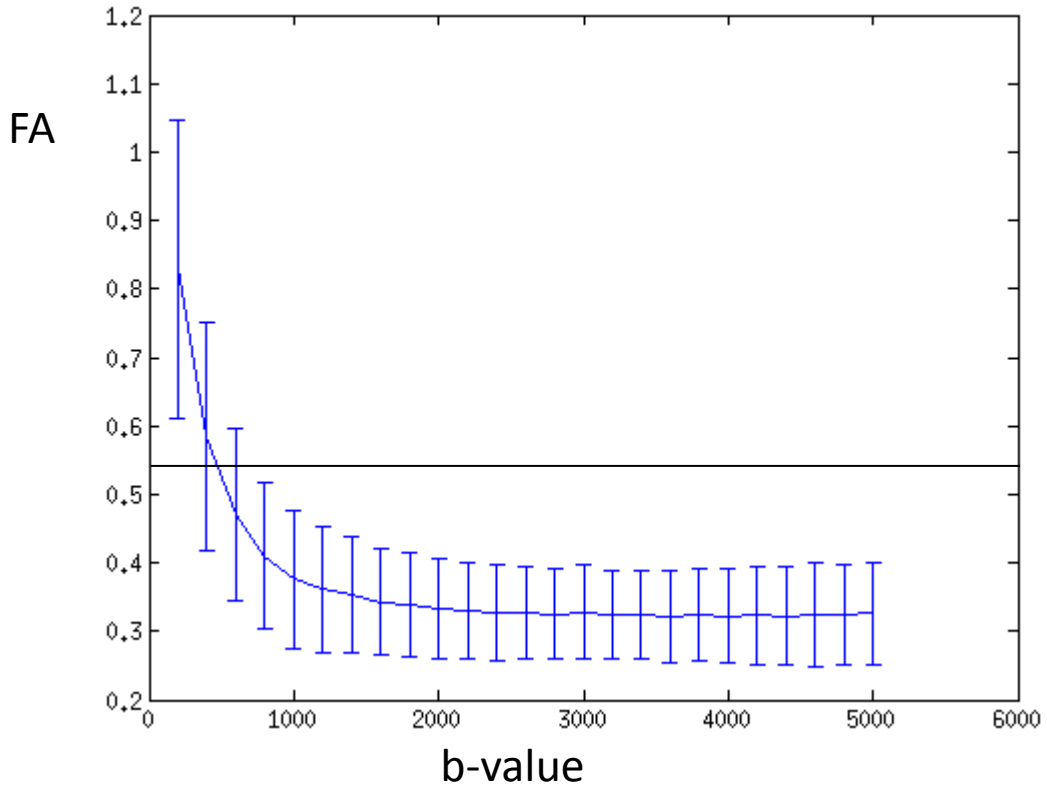



Figure 2.10 – Plots of FA estimates as function of b-values. The values of the diffusion tensor was adapted from (Qi, Wang, & Wu, 2008) which noise FA value is 0.54. Each re-estimate of FA was extracted for a diffusion tensor extracted from Rician corrupted diffusion-weighted signals (SNR=15) sampled using the b-value zero and the non-zero b-value marked in the plot over 30 directions. Diffusion tensor was extracted with an ordinary linear less squares Eq. 2.3.13. Similar plots can be seem in (Jones & Basser, 2004).

## 2.4. Diffusion Kurtosis Imaging

As mentioned previously, DTI is based on the assumption that displacement of water molecular is described by a 3D Gaussian distribution (Basser, 1995). However water in biological structures have shown to present non-Gaussian diffusion behavior (Thoralf Niendorf, Dijkhuizen, Norris, van Lookeren Campagne, & Nicolay, 1996). As example, for time intervals on the order of tens of milliseconds, diffusion displacement probability distribution seems to deviate substantially from the Gaussian distribution which is believed to be a consequence of the restrictions imposed on water translation by the microstructures barriers (Jensen et al., 2005).

Once higher b-values correspond to higher displacements timeframes, the non-Gaussian behaviour will be more evident using diffusion-weighted sequences with larger b-values. As a consequence of this, if one plots the log-transformation of the diffusion

weighted signal in order of the b-value, a gradual increased deviation from the linear relationship is observed when increasing the b-value. This profile of deviation was experimentally demonstrated in both white and grey matter (Lu et al., 2006), see Figure 2.11.

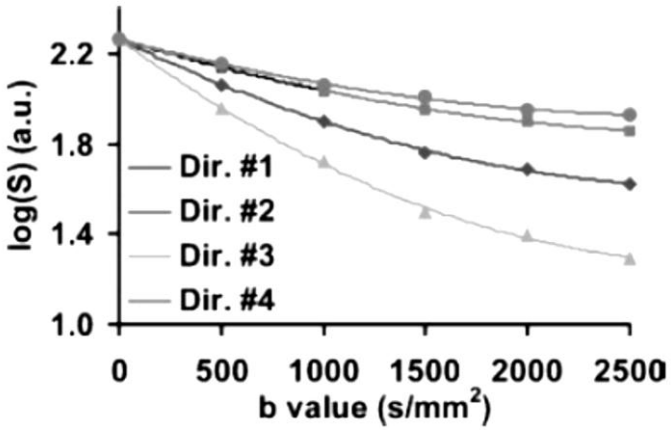



Figure 1.11 - Natural logarithm of DW signals vs. b-value for a selected ROI corresponding to voxels from corpus callosum. Nonlinear decays were observed for all four directions presented in the figure. This figure was adapted from (Lu et al., 2006).

Several approaches to characterize the non-Gaussian proprieties have been developed in the last years (Clark et al., 2002; Özarslan & Mareci, 2003; Tuch, 2004; Yablonskiy et al., 2003). One of the techniques that shown promising results is diffusion Kurtosis Imaging (DKI) (Jensen et al., 2005; Lu et al., 2006). DKI is an expansion of DTI, where the diffusion tensor is estimated together with a $4^{th}$-order 3D kurtosis tensor. Kurtosis $K$ is a measure that quantifies the deviation of the water diffusion profile from Gaussian distribution for which $K$=0. For positive Kurtosis the displacement distribution will have a more sharply peaked profile, such as a squared Gaussian, while for negative Kurtosis the distribution will be least sharply. Although mathematically kurtosis can be ranged to negative values, biological tissues shown only to present positive kurtosis (Wu & Cheung, 2010).

### 2.4.1. The formulation beyond DKI

Mathematically, distributions can be characterized by cumulants $\kappa_i$ (Jensen et al., 2005). The cumulant can be related by the central moments of the distribution, as mean $\mu$ and variance $\mu_2$. In particularly the first three cumulants are equal to the first three central moments which corresponds to the mean, variance and skewness (the latter momentum quantifies deviations from the central localization of the peak distribution). Kurtosis $K$

can be calculated from the second and fourth cumulant, $\kappa_2$ and $\kappa_4$, or from the second and fourth central momentum, $\mu_2$ and $\mu_4$ by

$$K = \frac{\kappa_4}{{\kappa_2}^2} = \frac{\mu_4}{{\mu_2}^2} - 3, \qquad \text{(Eq. 2.4.1)}$$

and is the dimensionless metric that gives the deviation of a Gaussian function. Assuming a PGSE sequence which $\delta$ is assumed to be much smaller than Δ, an general expression of the relationship between the diffusion signal decay and the diffusion coefficient can be derived as a summation of the cumulants $\kappa_p$:

$$\ln\left[\frac{S(g)}{S_0}\right] = \sum_{p=1}^{\infty} \kappa_p \frac{(i\gamma g\delta)^p}{p!} \qquad \text{(Eq. 2.4.2)}$$

Since the asymmetry of displacement of probability can be neglected (Wu & Cheung, 2010), all odd order cumulants can be approximated to zero and Eq 2.4.2 can be re-expressed as:

$$\ln\left[\frac{S(g)}{S_0}\right] = -\kappa_2 \frac{(\gamma g\delta)^2}{2} + \kappa_4 \frac{(\gamma g\delta)^2}{4!} + \kappa_6 \frac{(\gamma g\delta)^2}{6!} + \cdots \qquad \text{(Eq. 2.4.3)}$$

Given that diffusion D is defined by

$$D = \frac{\kappa_2}{2\Delta}, \qquad \text{(Eq. 2.4.4)}$$

Eq. 2.4.1 can be rewritten as

$$\kappa_4 = 4KD^2\Delta^2 . \qquad \text{(Eq. 2.4.5)}$$

Applying Eq. 2.4.4 and Eq. 2.4.5 into Eq. 2.4.3, and truncating all right handle terms above the second one, the relationship between the signal of a DWI and the *b*-value for a gradient direction can be approximated by:

$$S \approx S_0 exp\left\{-b.D_{app} + \frac{1}{6}b^2 {D_{app}}^2 K_{app}\right\} , \qquad \text{(Eq. 2.4.6)}$$

where $D_{app}$ and $K_{app}$ are called the apparent diffusion and kurtosis once they dependent on the direction of the applied diffusion gradient. For 3D anisotropic medium, kurtosis has to be defined by a 3x3x3x3 matrix to fully characterize non-Gaussian behaviour of the diffusion tensor $\boldsymbol{D}$. The elements $D_{ij}$ of the diffusion tensor and the elements $W_{ijkl}$ of the kurtosis tensor $\boldsymbol{K}$ can be related with $D_{app}$ and $K_{app}$ by the following equations:

$$D_{app} = \sum_{i=1}^{3}\sum_{j=1}^{3} n_i n_j D_{ij} \qquad \text{(Eq. 2.4.7)}$$

$$K_{app} = \frac{MD^2}{{D_{app}}^2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} n_i n_j n_k n_l W_{ijkl} \qquad \text{(Eq. 2.4.8)}$$

Combining Eq. 2.4.7 and Eq. 2.4.8 on Eq. 2.4.6, one can be write an equation relating the individual elements of diffusion and kurtosis tensor with the signal decay:

$$S(\boldsymbol{n}, b) = S_0 \exp\left[-b\sum_{i=1}^{3}\sum_{j=1}^{3} n_i n_j D_{ij} + \frac{1}{6}b^2 (MD)^2 \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} n_i n_j n_k n_l W_{ijkl}\right] \qquad \text{(Eq. 2.4.9)}$$

In analogy with the diffusion tensor, the kurtosis tensor is symmetric, therefore it can be fully described with only 15 independent elements and Eq. 2.4.8 can be simplified as Eq. 2.4.10 (Lu et al., 2006).

$$K_{app} = \frac{MD^2}{{D_{app}}^2}[n_1^4 W_{1111} + n_2^4 W_{2222} + n_3^4 W_{3333} + 4n_1^3 n_2 W_{1112} + 4n_1^3 n_3 W_{1113} + 4n_1 n_2^3 W_{1222} + 4n_2^3 n_1 W_{2223} + 4n_1 n_3^3 W_{1333} + 4n_2 n_3^3 W_{2333} + 6n_1^2 n_2^2 W_{1122} + 6n_1^2 n_3^2 W_{1133} + 6n_2^2 n_3^2 W_{2233} + 12n_1^2 n_2 n_3 W_{1123} + 12n_1 n_2^2 n_3 W_{1223} + 12n_1 n_2 n_3^2 W_{1233}] \qquad \text{(Eq. 2.4.10)}$$

Since DKI can be parameterized by 22 unknown elements (six independent elements for the diffusion tensor, 15 for the kurtosis tensor and one relative to the non-

diffusion weighted value $S_0$), it requires at least 22 DWIs on which 15 different gradient direction and three b-values have to be acquired to fully characterize the kurtosis tensor.

Similar to DTI, the parameters of DKI can be estimated using different approaches. Eq. 2.4.9 can also be converted to a linear framework where ordinary or weighted linear less squares can be applied. On the hand, one can fit the DKI unknown parameters directly to Eq. 2.4.9 using iterative non-linear methods. Constrains can be also introduced to the linear and non-linear frameworks to ensure that diffusion and kurtosis tensors are defined as plausible values. The formulations of such approaches are discussed below.

**Linear least squares**

After linearization by applying log transformations, Eq. 2.4.9 can be rewritten using matrixes:

$$\boldsymbol{L} = \boldsymbol{A}_k \boldsymbol{X}_k \qquad \text{(Eq. 2.4.11)}$$

where $\boldsymbol{L}$ (similar to Eq. 2.3.14) is a column vector containing all the log-transformed values of measured diffusion-weighed intensities for all diffusion gradient directions and b-values, while $\boldsymbol{A}_k$ consist on the parameters depending on the acquisition sequence parameters,

$$\boldsymbol{A_k} = \begin{bmatrix} -b^{(1)}\left(n_1^{(1)}\right)^2 & \cdots & -b^{(1)}n_1^{(1)}n_2^{(1)} & \cdots & \frac{1}{6}\left(b^{(1)}\right)^2\left(n_1^{(1)}\right)^4 & \cdots & \frac{12}{6}\left(b^{(1)}\right)^2 n_1^{(1)}n_2^{(1)}\left(n_3^{(1)}\right)^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -b^{(N)}\left(n_1^{(N)}\right)^2 & \cdots & -b^{(N)}n_1^{(N)}n_2^{(N)} & \cdots & \frac{1}{6}\left(b^{(N)}\right)^2\left(n_1^{(N)}\right)^4 & \cdots & \frac{12}{6}\left(b^{(1)}\right)^2 n_1^{(1)}n_2^{(1)}\left(n_3^{(1)}\right)^2 \end{bmatrix} \qquad \text{(Eq. 2.4.12)}$$

and $\boldsymbol{X}_k$ contains all DKI 22 unknown parameters

$$\boldsymbol{X_k} = [D_{11} \quad D_{22} \quad D_{33} \quad D_{12} \quad D_{13} \quad D_{23} \quad V_{1111} \quad \ldots V_{1112} \quad \ldots V_{1122} \ldots V_{1233} \quad \ln S_0\,]^T \qquad \text{(Eq. 2.4.13)}$$

where $V_{ijkl}$ is given by

$$V_{ijkl} = MD^2 \times W_{ijkl} \, . \qquad \text{(Eq. 2.4.14)}$$

Assuming that the elements of the vector $\boldsymbol{L}$ have errors with independent and identically distributed variance, the linear least square solution can be obtained from Eq. 2.4.11 as:

$$\widehat{\boldsymbol{X}}_{\boldsymbol{k}} = \left[\boldsymbol{A}_k{}^T \boldsymbol{A}_k\right]^{-1} \boldsymbol{A}_k{}^T \boldsymbol{L} = \boldsymbol{A}_k{}^+ \boldsymbol{L} \qquad \text{(Eq. 2.4.15)}$$

where $\widehat{\boldsymbol{X}}_k$ is the estimation of vector $\boldsymbol{X}_k$, and $\boldsymbol{A}_k{}^+$ the pseudoinverse of matrix $\boldsymbol{A}_k$.

**Weighted Linear least squares**

For a more general linear solution, Eq 2.4.11 can be solved as

$$\widehat{\boldsymbol{X}}_k = \left[\boldsymbol{A}_k{}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{A}_k\right]^{-1} \boldsymbol{A}_k{}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{L} \, . \qquad \text{(Eq. 2.4.16)}$$

Similarly to the DTI weighted linear least squares version mentioned in section 2.3.1, the covariance matrix $\boldsymbol{\Sigma}$ allows the introduction of weights to correct the non-uniform variance nature of the log-transformed versions of diffusion-weighted measures. For example assuming a uniform variance on the non-logarithm transformed signals, $\boldsymbol{\Sigma}$ can be computed as a diagonal matrix where diagonal elements are given by Eq. 2.3.16.

**Non-linear least squares**

In analogy to the formulation described on Eq. 2.3.16, Eq. 2.4.11 can be solved directly by fitting values on the unknown parameters using iterative algorithms as applying Levenberg-Marquardt nonlinear fitting, as proposed on first studies on diffusion kurtosis imaging (Jensen et al., 2005; Lu et al., 2006). As starting point one can use the linear least squares solutions mentioned above.

**Constrained linear least squares**

An example of constrained linear least squares was suggested by Tabesh et al. (2011) which not only ensures that all directional diffusivities are positive but also all directional

kurtosis is in an acceptable range. This constrains can be described with the following equations:

$$D(\boldsymbol{n}) \geq 0\ , \qquad \text{(Eq. 2.4.17)}$$

$$K(\boldsymbol{n}) \geq 0\ , \qquad \text{(Eq. 2.4.18)}$$

$$K(\boldsymbol{n}) \leq \frac{3}{b_{max} D(\boldsymbol{n})}. \qquad \text{(Eq. 2.4.19)}$$

As mentioned on previous sections, diffusion should to be positively defined for all spatial direction to be physically plausible (constrains described by Eq. 2.4.17). Although mathematically feasible minimum kurtosis of a probability distribution is -2, multi-compartment diffusion models and empirical evidence suggest that biological values of kurtosis are always positive, and therefore the minimum directional kurtosis is constrained to zero by Eq. 2.4.18 (Tabesh et al., 2011) . Moreover, to ensure that all directions the diffusion-weighted signal is always a strictly decrease function of b-value, the maximum kurtosis should satisfy Eq. 2.4.19.

Defining Eq. 2.4.17, 2.4.18, and 2.4.19, the constrained can be imposed when solving the linear least squares problem using convex quadratic algorithms, where the DKI unknown parameters are fitted until the right-hide of Eq. 2.4.11 closely matches the left-hand side and simultaneously satisfying Eq. 2.4.17, Eq. 2.4.18 and 2.4.19. This problem can be formulated as:

$$\begin{aligned} &Minimize\ \|\boldsymbol{A}_k\boldsymbol{X}_k - \boldsymbol{L}\|^2 \\ &such\ that\ \boldsymbol{C}\boldsymbol{X}_k \leq \boldsymbol{d}, \end{aligned} \qquad \text{(Eq. 2.4.20)}$$

where $\boldsymbol{d}$ is a $3N$ column vector of zeros and $(3N) \times 22$ matrix $\boldsymbol{C}$ represents the linear constrains which is given by:

$$\boldsymbol{C} = \begin{bmatrix} -\left(n_1^{(1)}\right)^2 & \cdots & -n_2^{(1)}n_3^{(1)} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\left(n_1^{(N)}\right)^2 & \cdots & -n_2^{(N)}n_3^{(N)} & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & -\left(n_1^{(1)}\right)^4 & \cdots & -n_1^{(1)}n_2^{(1)}\left(n_3^{(1)}\right)^2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & -\left(n_1^{(N)}\right)^4 & \cdots & -n_1^{(N)}n_2^{(N)}\left(n_3^{(N)}\right)^2 & 0 \\ -3\left(n_1^{(1)}\right)^2/b_{max} & \cdots & -3n_2^{(1)}n_3^{(1)} & \left(n_1^{(1)}\right)^4 & \cdots & n_1^{(1)}n_2^{(1)}\left(n_3^{(1)}\right)^2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -3\left(n_1^{(N)}\right)^2/b_{max} & \cdots & -3n_2^{(N)}n_3^{(N)} & \left(n_1^{(N)}\right)^4 & \cdots & n_1^{(N)}n_2^{(N)}\left(n_3^{(N)}\right)^2 & 0 \end{bmatrix}$$

(Eq. 2.4.21)

To constrain the lower boundary of plausible values of kurtosis as an value $\boldsymbol{k_{min}}$ rather than zero, the last *N* elements of vector *d* have to be defined as

$$\left[\boldsymbol{K_{min}D}\left(\boldsymbol{n}^{(1)}\right)^2 \quad \cdots \quad \boldsymbol{K_{min}D}\left(\boldsymbol{n}^{(N)}\right)^2\right] \qquad \text{(Eq. 2.4.22)}$$

where *D* is the diffusion tensor which have to be approximated a priori. For this priori approximation the linear least squares solution can be used (Tabesh et al., 2011).

### 2.4.2. DKI invarant Mesuares

For DKI invariant parameters can be defined from both tensors diffusion and kurtosis tensors. The same invariant metrics described in section 2.3.2 can be extracted for the diffusion tensor. In this section four more invariant parameters relative to the kurtosis tensor is described below: 1) the mean kurtosis (*MK*); 2) the axial kurtosis ($k_{//}$), 3) the radial kurtosis ($k_{\perp}$), and 4) the kurtosis fractional anisotropy ($FA_k$) (Wu & Cheung, 2010). It is notable that these latter invariant parameters are very important to extract information from the kurtosis tensor, since the interpretation of individual elements of the kurtosis are yet to be explored (Hui, Cheung, Qi, & Wu, 2008).

Analogous to MD, MK gives a measure of the overall kurtosis which can be estimated as the average of directional kurtosis:

$$MK = \frac{1}{N}\sum_{i=1}^{N}\left(K_{app}\right)_i \qquad \text{(Eq. 2.4.23)}$$

Thus, from the elements$W_{ijkl}$, MK can be estimated by first calculating all directional kurtosis $\left(K_{app}\right)_i$ using Eq. 2.4.7 and Eq. 2.4.8, and then using Eq. 2.4.23.

Similar to the directional diffusivity $\lambda_{//}$ and $\lambda_{\perp}$, one can compute the values of kurtosis along the directions perpendicular and parallel to the first principal axis of the diffusion ellipsoid, which is of interest for white matter regions since it gives additional information of the axonal and myelin integrity (Helpern et al., 2011). The first step on estimating the directional kurtosis $k_{//}$ and $k_{\perp}$ is finding the values of kurtosis along each direction of diffusion ellipsoid principal axis $\boldsymbol{\varepsilon_1}$, $\boldsymbol{\varepsilon_2}$, and $\boldsymbol{\varepsilon_3}$. This is achieved by rotating the kurtosis tensor $\boldsymbol{K}$ from the standard Cartesian coordinate system to the coordinate system formed by the three orthogonal eigenvectors of $\boldsymbol{D}$ (Hui et al., 2008):

$$\widehat{W}_{ijkl} = \sum_{i'=1}^{3}\sum_{j'=1}^{3}\sum_{k'=1}^{3}\sum_{l'=1}^{3} e_{i'i}e_{j'i}e_{k'i}e_{l'i}W_{i'j'k'l'} \qquad \text{(Eq. 2.4.24)}$$

where $e_{ij}$ are the elements of the 3D rotation matrix defined by the diffusion eigenvectors $\boldsymbol{\varepsilon_1}$, $\boldsymbol{\varepsilon_2}$, and $\boldsymbol{\varepsilon_3}$:

$$\boldsymbol{e} = \begin{bmatrix} \varepsilon_{1x} & \varepsilon_{2x} & \varepsilon_{3x} \\ \varepsilon_{1y} & \varepsilon_{2y} & \varepsilon_{3y} \\ \varepsilon_{1z} & \varepsilon_{2z} & \varepsilon_{3z} \end{bmatrix} \qquad \text{(Eq. 2.4.25)}$$

and $\widehat{W}_{ijkl}$, the elements of the rotated kurtosis tensor. Having the values of $\widehat{W}_{ijkl}$, the values of kurtosis over the diffusion ellipsoid three principal axes ($i$=1, 2 and 3) can be computed with the following equation:

$$\kappa_i = \frac{MD^2}{{\lambda_i}^2}.\widehat{W}_{iiii}, \qquad \text{(Eq. 2.4.26)}$$

and finally, axial kurtosis ($k_{//}$) can be estimated by the values of kurtosis along the first principal diffusion ellipsoid axis:

$$\kappa_{\parallel} = \kappa_1 \qquad \text{(Eq. 2.4.27)}$$

while radial kurtosis ($k_\perp$) can be computed by the average between the values of the two axis related to the smaller diffusion values:

$$\kappa_\perp = \frac{\kappa_2+\kappa_3}{2} \qquad \text{(Eq. 2.4.28)}$$

In analogy with diffusion FA, the anisotropy of directional kurtosis can be conveniently defined as:

$$FA_k = \frac{\sqrt{3[(\kappa_1-\langle\kappa\rangle)^2+(\kappa_2-\langle\kappa\rangle)^2+(\kappa_3-\langle\kappa\rangle)^2]}}{\sqrt{2({\kappa_1}^2+{\kappa_2}^2+{\kappa_3}^2)}} \qquad \text{(Eq. 2.4.29)}$$

where

$$\langle\kappa\rangle = \frac{1}{3}\sum_{i=1}^{3}\kappa_i \ . \qquad \text{(Eq. 2.4.30)}$$

$FA_k$ is close to one if one value of $\kappa_i$ is much larger than the other two, and close to zero if all values of $\kappa_i$ are similar to each other.

### 2.4.3. Effects of noise on DKI

Since DKI follows a different fitting model relative to DTI, the Rician noise on the DWIs will have different consequences than the ones described on 2.3.3. As DKI metrics of non-Gaussianity are not based on a biological model, its interpretation must be explored to interpret the biological meaning of the results (Fieremans et al., 2011). In the same way, subtle effects of noise, motion and artefacts are yet be explored to distinguish the differences induced by biological changes from the differences induced by artefacts. Though, it is already well known that noise can result on large amount of negative values of kurtosis as well as implausible large values (Tabesh et al., 2011). In addition, despite being more sensitive to errors due to the larger number of unknown parameters to be estimated (Kay, 1993), previous studies shown that DKI provide more robust and unbiased data when compared to DTI (Veraart et al., 2011a). For instance, in opposite to DTI that shown to be dependent on the selected b-value due to Rectified noise (see section 2.3.3 for more details), the more general model proposed on DKI shown to provide diffusion measures less sensitive to different b-values (Veraart et al., 2011). Such

is demonstrated on Figure 2.12, where diffusion-weighed intensities are simulated from values of diffusion ($0.74x10^{-3}mm^2.s^{-1}$) and kurtosis coefficient (0.86) for isotropic media (grey matter). After corrupting the diffusion-weighted intensities by adding Rician noise (SNR=50), the values of diffusion coefficient were re-estimated using ordinary linear least solutions for DTI and DKI (panel A and B) based on three b-values (0, b1/2 and b1). With the increase of b-value, diffusion from DTI seems to be erratic overestimated, while DKI show unbiased values i.e. the mean of the estimates matches the expected value ($0.74x10^{-3}mm^2.s^{-1}$) marked with the back line. This pattern is similar to empirical observations in previous studies (Veraart et al., 2011a).

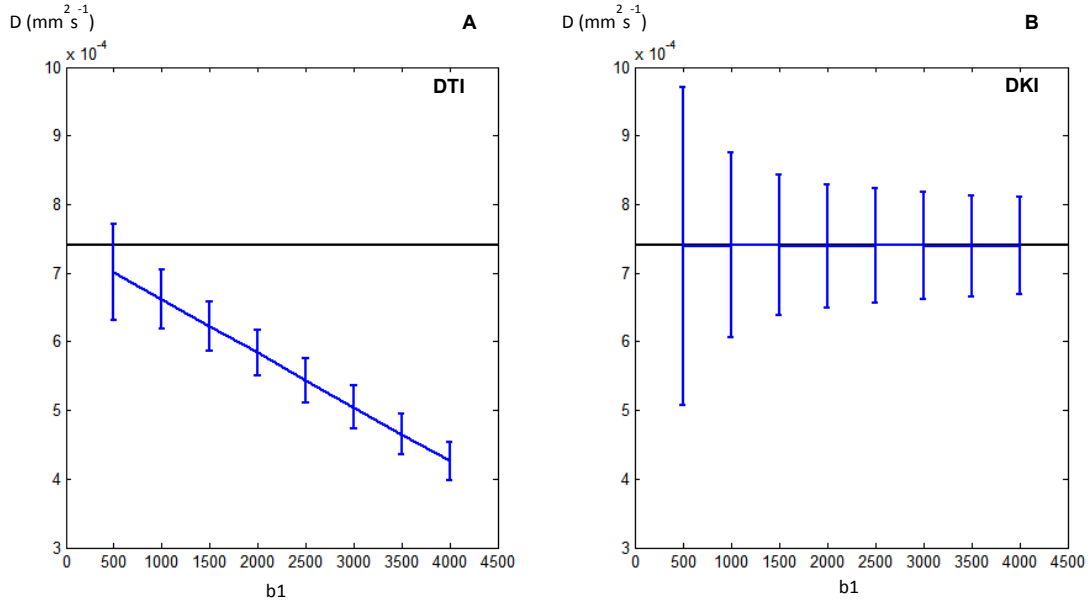



Figure 2.12 – Comparison of noise affects between DTI (panel A) and DKI (panel B). Each plots correspond to the mean of 10000 diffusion values which were extracted from linear least square versions of DTI and DKI on simulated diffusion-weighted signal from three b-values (0, b1, and b1/2) and corrupted with Rician noise (SNR=50). Simulates were based on known values of diffusion ($0.74x10^{-3}mm^2.s^{-1}$) and kurtosis (0.86) from an isotropic media which corresponds to gray matter values adapted from (Jensen et al., 2005). The expected value of diffusion is marked in both plot by the vback line. Despite DKI being more sensitive to noise (larger error bars), it doesn't produce biased data or values dependent on the b-value as the data extracted by DTI.

In section 2.3.3 deviation of the linear relationship between log-transformed diffusion-weighted signal and b-values was mentioned being an effect of rectified noise. Although DKI assumed that microstructural barriers are responsible for non-Gaussian behaviour which reflects nonlinearities, it is not denied that rectified noise can influence in increases on the degrees of non-Gaussianity. Therefore, as mentioned on previous studies, it is possible that kurtosis can be positive bias (Jones & Basser, 2004). However

such overestimation of kurtosis with the influence of Rician noise never was experimentally demonstrated. This will be novelty addressed in chapter 3 (Effects of rectified noise in kurtosis). In addition, in chapter 5 the consequence of the eigenvalues repulsion on DKI will experimental shown for the analysis of the CamCAN data.

# Chapter 3

# Extraction of the Tensor and Tensor-derived Measures in DKI

## 3.1. Introduction

As mentioned in the previous chapter, after acquisition of diffusion-weighted images, DTI and DKI can be applied by using different fitting approaches. The accuracy of the diffusion and kurtosis metrics will depend on the robustness of the fitting approach selected. Some are more sensitive to the noise effect than others and by increasing its complexity one can even developed procedures to avoid implausible estimates (Koay, Carew, Alexander, Basser, & Meyerand, 2006; Tabesh, Jensen, Ardekani, & Helpern, 2011).

For instance, in the context of DTI, non-linear approaches seems to provide diffusion estimates less biased from the rectified noise effects than the ordinary and weighted linear solutions (Jones & Basser, 2004). Latter in 2006, constrains were added to DTI linear and non-linear frameworks using Cholesky decomposition (Koay et al., 2006), where the constrained non-linear approach showed a better performance in estimating the diffusion tensor without generating negative implausible values of diffusion or implausible values of FA (values larger than one). The Cholesky decomposition of the diffusion tensor can be applied to DKI techniques to ensure that estimates of diffusion are positively defined (Verrart, Van Hecke, & Sijbers, 2011b); however this does not guarantee that kurtosis estimates will fall in a plausible range.

For the first studies on DKI, non-linear methods were used to extract the diffusion and kurtosis tensors (Jensen, Helpern, Ramani, Lu, & Kaczynski, 2005; Lu, Jensen,

Ramani, & Helpern, 2006). Although such approaches can be more robust in dealing with noise, since they can account for the noise Rician nature directly (Veraart et al., 2011b), they do not ensure that diffusion and kurtosis values are in an acceptable range. In particular, following these approaches can result in negative diffusion estimates, implausible negative of kurtosis, or kurtosis values that are implausible high (Tabesh et al., 2011; Veraart et al. 2011b). To deal with this problem, a constrained algorithm on a linear framework was proposed by Tabesh et al. (2011) which results in data with lower errors and similar to data extracted from unconstrained linear methods using larger amounts of DWIs, i.e. the larger accuracy of this method is potential to allow the use of less data to produce similar results to previously proposed unconstrained methods, and consequently allows the reduction of scanning time.

In this chapter, different fitting procedures to estimate the quantitative parameters for DKI will be studied using simulated data corrupted by Rician noise. These procedures will include the ordinary (OLS) and weighted linear least squares (WLS) solutions, a non-linear least squares algorithm (NLS) using the Levenberg-Marquardt fitting, a DKI adaption of a non-linear algorithm were diffusion tensor is rewritten by its Cholesky decomposition (chNLS) to ensure positive diffusion estimates, and the constrained linear method (CLS) proposed by Tabesh et al. (2011). An additional linear least squares method is studied for direct extraction of MD and MK, which was named as the direct linear least squares fit (DLS). Despite not allowing estimates of other invariant parameters, this simple method can potentially overcome some noise artefacts on MD and MK values, since they are extracted from DWIs on a single step.

In this chapter, novel results on the development of DKI estimation framework to increase the robustness in the presence of noise are presented. First of all, previous studies never compared such diversity of fitting methods for DKI. Moreover, some noise influences for these fitting methods will be clarified. At the end of the chapter it will be decided which DKI processing algorithm should be used on the DKI procedures for the Cam-CAN project.

## 3.2. Methods

### 3.2.1. Single voxel simulates

In this chapter the accuracy of the different fitting methods are tested firstly on simulated data based on values for a single voxel of white matter. The typical values of the diffusion and kurtosis tensors for this white matter voxel are adapted from (Qi, Wang, & Wu, 2008) which are adjusted to ensure that all the given directional values of diffusion and kurtosis are positive defined with the absence of noise. Simulations of diffusion-weighted signal are therefore obtained using the DKI model (Eq. 2.4.9) by defining the following elements: $S_0 = 500$, $D_{11} = 0.1783 \times 10^{-3}$, $D_{22} = 0.1459 \times 10^{-3}$, $D_{33} = 0.4028 \times 10^{-3}$, $D_{12} = 0.0011 \times 10^{-3}$, $D_{13} = 0.0125 \times 10^{-3}$, $D_{23} = 0.0034 \times 10^{-3}$ , $W_{1111} = 0.5698$, $W_{2222} = 0.3208$, $W_{3333} = 2.6049$, $W_{1112} = -0.1142$, $W_{1113}$=-1.1521, $W_{1222} = 0.3944$, $W_{2223} = -0.5409$, $W_{1333} = 0.7220$, $W_{2333}$ =0.777, $W_{1122}$ =0.2865, $W_{1133}$ =0.4700, $W_{2233} = 0.2065$, $W_{1123} = -0.0005$, $W_{1223} = 0.1655$, and $W_{1233} = -0.4157$. The gradient directions and b-values used on these simulations are based on the Cam-CAN's acquisition scheme, i.e. 30 gradient directions for b-values 1000 and 2000 s.mm$^{-2}$ and 3 images acquired with a b-value 0 s.mm$^{-2}$. At last, each simulated diffusion-weighted signal is corrupted with different levels of Rician noise using Eq. 2.3.21.

### 3.2.2. All Brain simulation

As mentioned in sections 2.3.3 and 2.4.3, the noise artefacts can be dependent on the noise free values of diffusion and kurtosis tensors. Therefore, for an analysis that is more comparable to real brain data, it is necessary to test the methods on a range of values similar to the range found in real data. To achieve this goal, plausible ranges of diffusion and kurtosis tensors of an entire human brain are simulated, using the procedure represented by the diagram on Figure 3.1. Data for a single Cam-CAN subject was selected[1] (panel A of Figure 3.1) and to largely decrease the noise effects, motion and distortion artefacts, DWIs are smoothed using the SPM8 function *spm_smooth* (panel C of Figure 3.1). The full-width-at-half-maximum (FWHM) of the Gaussian kernel used on the *spm_smooth* function was set to 5 mm, which corresponds to a value two times higher

[1] Cam-CAN data is acquired on a 3T Siemens Trio with a twice-refocused-spin-echo, 30 diffusion gradient directions for each b-values 1000 and 2000s.mm$^{-2}$ and three images acquired using b-value 0, TR=9100ms,

than the used in the most recent DKI studies (e.g. Falangola et al., 2008; Jensen et al., 2011). This excess of smoothing is applied to ensure that implausible values are significantly reduced and, despite corrupting the resolution of data, this should provide a range of brain values of diffusion and kurtosis that should be similar to an ideal noise free data set.

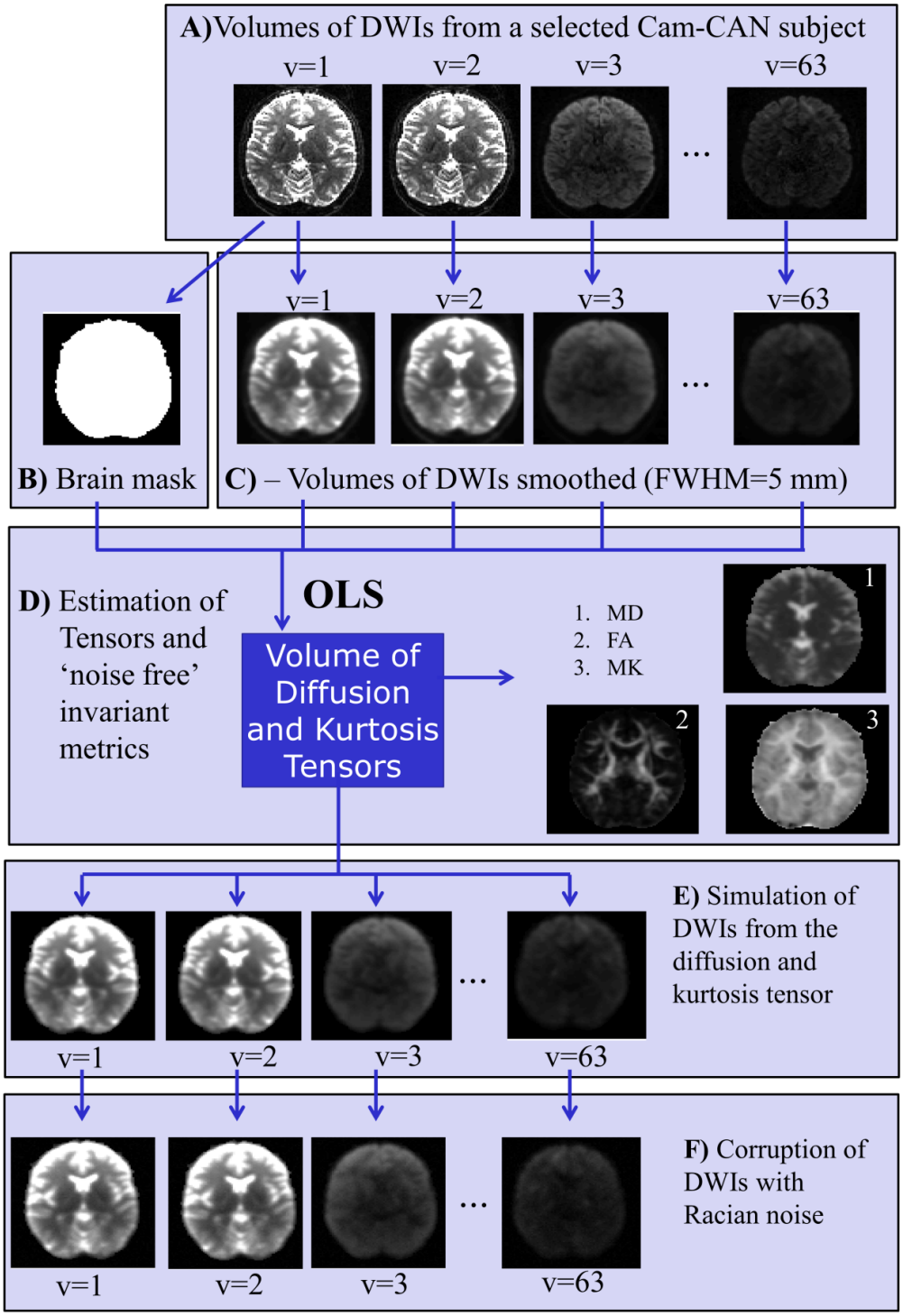



Figure 3.1 – Procedure used to simulate entire brain volumes of DWIs.

After smoothing the DWIs, the range of diffusion and kurtosis tensor values are extracted only from the brain voxels using the OLS method, and the values of MK, MD, and FA are extracted using Eq. 2.3.19, Eq. 2.3.22, and Eq. 2.4.23 (panel D of Figure 3.1). The brain voxels are selected based on a brain mask computed from the first volume of DWI (which was acquired with b-value=0 s.mm$^{-2}$) using FSL brain extraction tool (Smith, 2002), panel B of Figure 3.1. Despite being a simplistic approach, OLS should be enough to provide accurate results due to the high reduction of artefact and noise effects. For instance, Figure 3.2 shows the histograms of MK, MD, and FA extracted from the diffusion and kurtosis tensors estimated from the 5 mm smoothing diffusion-weighed data using OLS. One can observe that with the exception of ten implausible values of MK (which is negligible since it corresponds to 5x10$^{-3}$% of the number of simulated voxels) all remaining values ranged within plausible intervals.

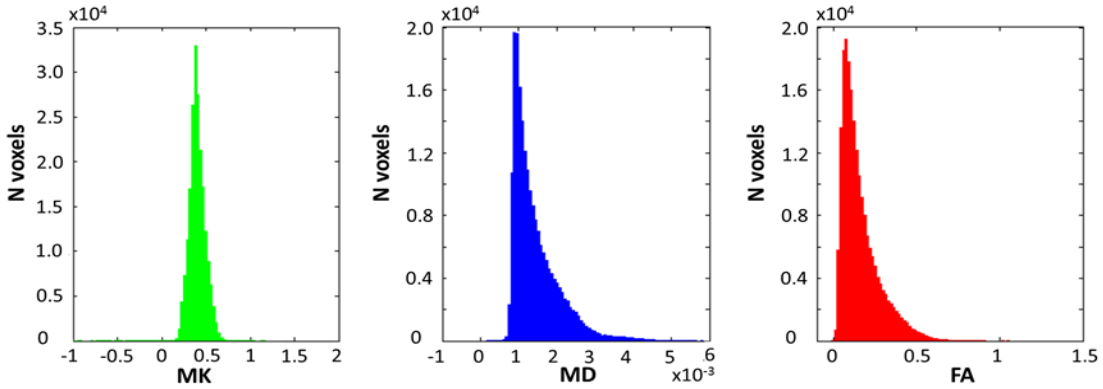



Figure 3.2 – Histograms of noise free values of MK, MD and FA obtained from the estimates of diffusion and kurtosis tensors distributed in a similar range of an entire brain data set.

Rician corrupted DWIs are then simulated from all the brain simulated diffusion and kurtosis tensors using Eq. 2.4.9 (based on the Cam-CAN's b-values and gradient directions, panel E of Figure 3.1) and Eq. 2.3.23 (based on Rician noise with uniform standard deviation of 5, which corresponds approximately to a signal to noise ratio of 100 for typical values of white matter on a DWI acquired using a b-value of 0 s.mm$^{-2}$, panel F of Figure 3.1).

### 3.2.3. Tensor fitting methods

For the simulated DWIs for both single white matter voxels and all brain ranged voxels, the values of diffusion and kurtosis tensors were re-estimated using the OLS, WLS, NLS, chNLS, and CLS. With the exception of chNLS, the details of the formulation of each

method are described in chapter 2 section 2.4.1. For chNLS the elements of the diffusion tensor were decomposed as discussed in section 2.3.1 and incorporated in a DKI non-linear formulation. Instead of the value of zero suggested by Tabesh et al. (2011), the lower boundary of directional kurtosis for CLS was set to -3/7 which corresponds to the theoretical kurtosis limit for regions that consist of water confined to spherical pores (Barmpoutis & Zhuo, 2011). Although recent papers argue that only positive values of kurtosis are plausible (e.g. Tabesh et al., 2011), this lower directional kurtosis boundary was selected to give more flexibility to the CLS fit. After having the estimates of the diffusion and kurtosis tensors for the five tested methods, the values of MD, FA, and MK were computed using Eq. 2.3.17, Eq. 2.3.20, and Eq. 2.4.23.

### 3.2.4. MK and MD direct fitting method

For the proposed method DLS, the MD and MK are directly estimated by fitting the diffusion weighted signals for all directions and b-values of the selected acquisition scheme in the following equation:

$$\log S = \log S_0 - b.MD + \frac{1}{6}b^2 MD^2 MK \,. \qquad \text{(Eq. 3.2.1)}$$

As Eq. 3.2.1 is written in a linear framework, the unknown parameters MK, MD, and $\log S_0$ can be solved by linear least squares which solution can be written as:

$$\widehat{\boldsymbol{X}}_d = \boldsymbol{A}_d{}^{+}\boldsymbol{L} \qquad \text{(Eq. 3.2.2)}$$

where $\boldsymbol{L}$ is a column vector containing all the log-transformed values of measured diffusion-weighed intensities, $\boldsymbol{A}_d{}^{+}$ is the pseudo inverse matrix of $\boldsymbol{A}_d$ which consist on the know acquisition sequence parameters:

$$\boldsymbol{A}_d = \begin{bmatrix} -b^{(1)} & \frac{1}{6}\left(b^{(1)}\right)^2 & 1 \\ \vdots & \vdots & \vdots \\ -b^{(N)} & \frac{1}{6}\left(b^{(N)}\right)^2 & 1 \end{bmatrix}, \qquad \text{(Eq. 3.2.3)}$$

and $\widehat{\boldsymbol{X}}_d$ an column vector with the estimates of the unknown parameters:

$$\widehat{\boldsymbol{X}}_d = [MD \quad MK \times MD^2 \quad \log S_0] \,. \qquad \text{(Eq. 3.2.4)}$$

### 3.2.5. Evaluation of results and selection of the optimal fitting method

In this chapter, the accuracy of the fitting methods is determined based on the analysis of three invariant parameters: MK, MD, and FA. For quantifying the differences between the noise free parameters ($\theta$) and the re-estimated ones ($\hat{\theta}$) the root mean square error (*RMSE*) is computed using the following equation (Tabesh et al., 2011):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{\theta}_i - \theta_i)^2}{n}} \qquad \text{(Eq. 3.2.5)}$$

where $n$ is the number of estimated values (which corresponds to 202315 voxels for the real brain-based simulation and to 10000 voxel for each level of Rician noise tested on the single voxel analysis).

Notably, *RMSE* provides a metric of accuracy only based on the global differences between expected values and estimated ones, and therefore can depends on the error mean and/or standard deviation. For example, *RMSE* can have smaller values even for biased data if the estimates' standard deviation is low; or larger values for unbiased data if standard deviation is high. Therefore, for a better description of the accuracy of each method, the mean error (*M*) and error standard deviation (*SD*) are also computed. When computing *M*, *SD*, and *RMSE*, the values of MK are concatenated with a lower threshold of -2 to avoid excessively large *RMSE* values (Tabesh et al., 2011).

Since Cam-CAN will involve the processing of data for 700 subjects, the computational processing time is a relevant issue on the selection of the optimal fitting method. Therefore, the required time to process DWIs volumes for each method is also analysed in addition to the accuracy metrics.

## 3.3. Results

### 3.3.1. Single voxel simulation

For the single simulated voxel analysis, OLS, WLS, NLS, and chNL show similar performances. For example, the histograms presented in Figure 3.3, which correspond to 10000 samples corrupted with Rician noise standard deviation of 5 (SNR=100), show similar profiles across these four methods for all invariant metrics.

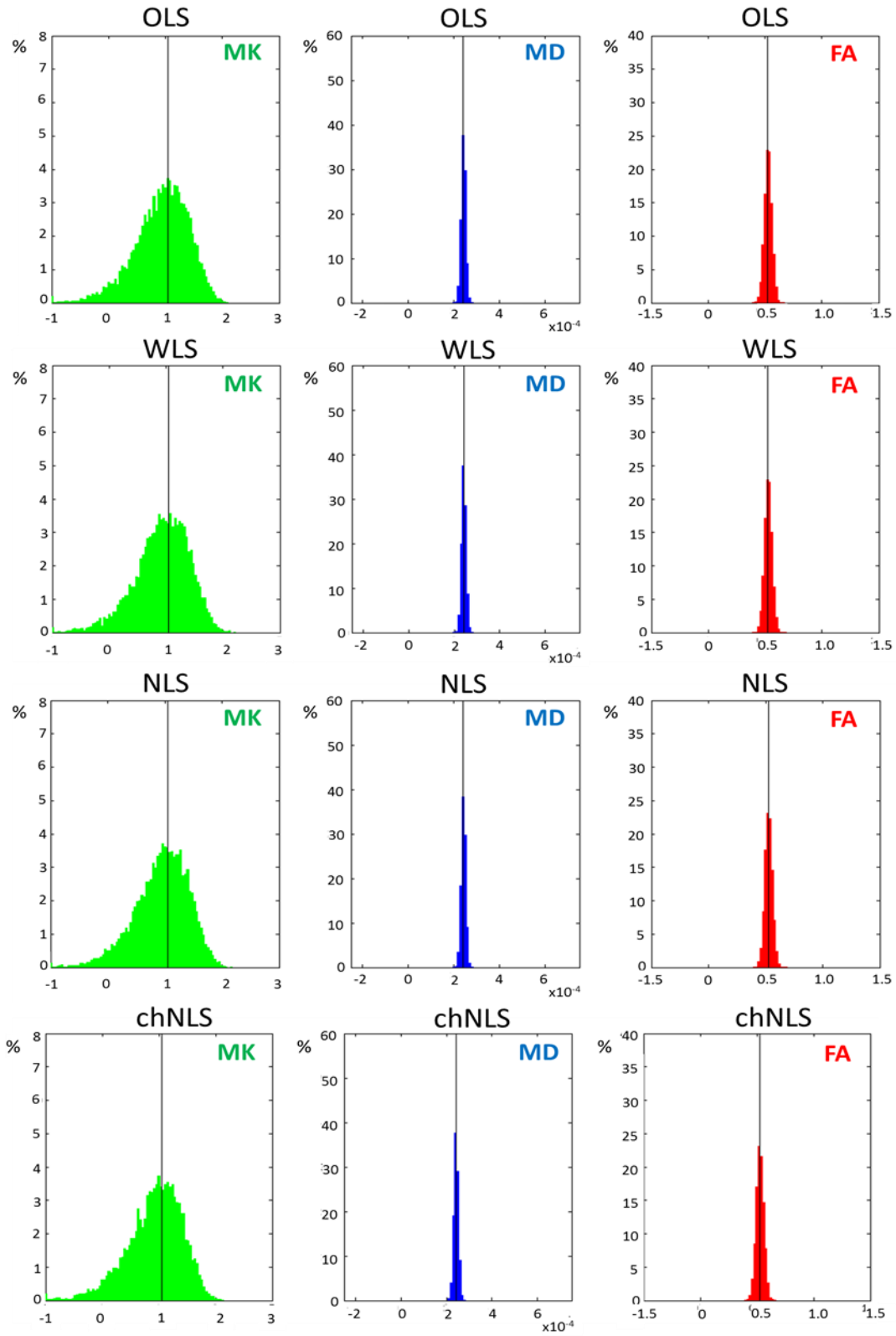



Figure 3.3 - Histogram of 10000 samples of MK, MD, and FA values estimated from OLS, WLS, NLS and chNLS. The values were extracted from simulated data corrupted by Rician noise with standard deviation of 5 (SNR=100). Histogram are plotted in the interval [-a, 3a], where a is the expected values of MK, MD, or FA which are marked with the back lines on each panel.

The histograms on Figure 3.3 are all plotted in the interval [-a, 3a], where a is the expected values of MK, MD, or FA marked with the back lines on each panel. Thus, from Figure 3.3 it can be observed that values of kurtosis are more sensitive to noise than values of diffusion (in particularly MK is the invariant metric that present significant amounts of implausible values).

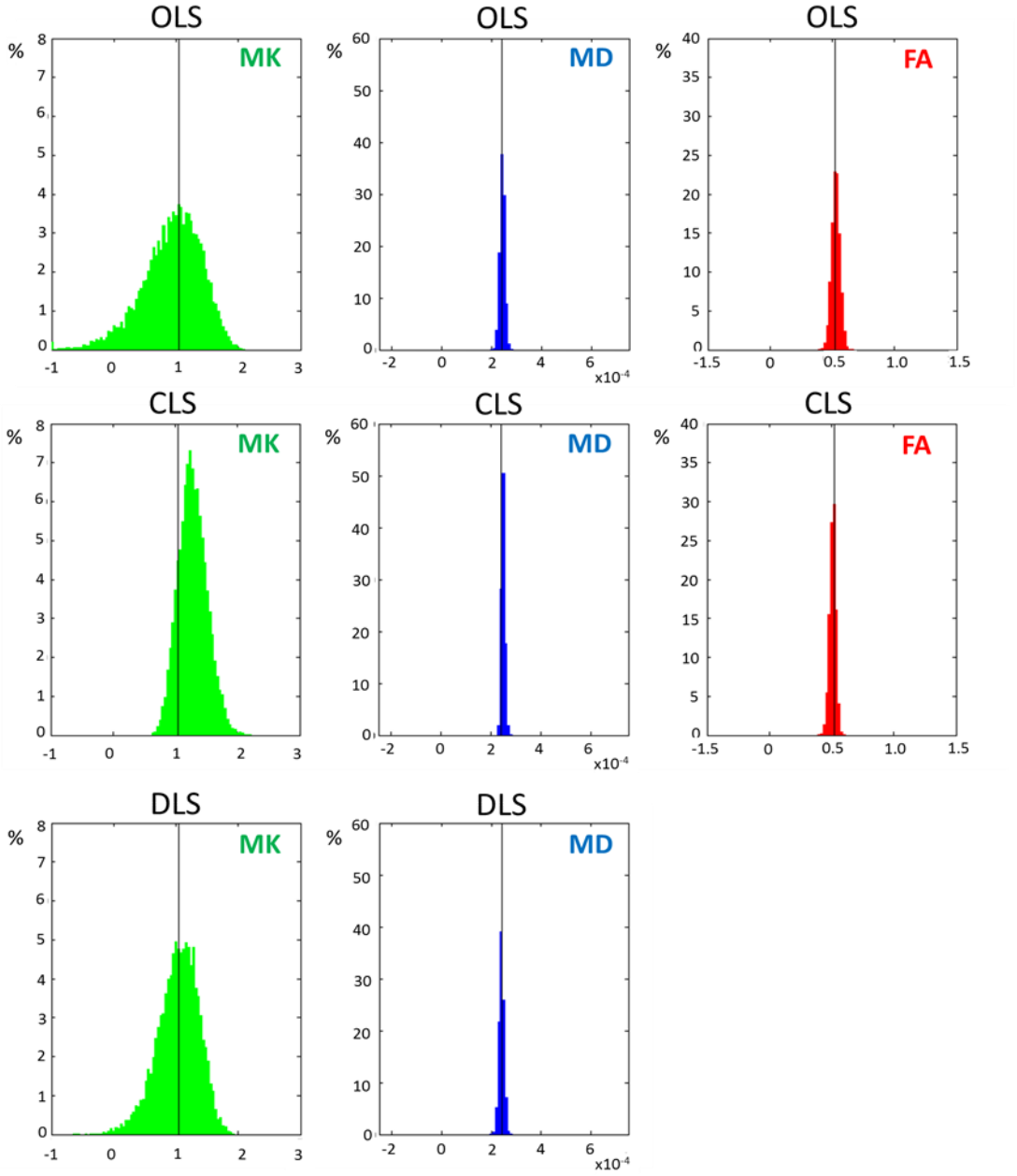



Figure 3.4 - Histogram of 10000 samples of MK, MD, and FA values estimated from OLS, CLS, and DLS. The values were extracted from simulated data corrupted by Rician noise with standard deviation of 5 (SNR=100). Histogram are plotted in the interval [-a, 3a], where a is the expected values of MK, MD, or FA which are marked with the back lines on each panel.

The histograms of CLS and DLS, which shown different profiles relative to OLS, are represented in Figure 3.4. Both CLS and DLS methods resulted on smaller dispersion of MK errors relative to the methods on Figure 3.4. While DLS still outcomes with negative values of MK, CLS seem to be successful on avoiding them. However, CLS show to produce biased results of MK. Notably, FA is not presented for DLS once this method does not provide estimates of metrics rather than MD and MK.

Figure 3.5 shows the values of *RMSE*, *M* and *SD* for the MK values with increasing Rician noise levels. The curves of WLS, NLS and chNL are not presented since they overlap with the curve of the OLS. The values of *M* show that OLS and DLS seem to underestimate the values of MK whereas CLS seems to overestimate them. For all the methods, *SD* increases with the increase of Rician noise level. CLS is the method that shows the smallest *SD*, followed by the DLS method and OLS. For the values of *RMSE*, OLS shows larger values than DLS and CLS, while the values for CLS are the smallest.

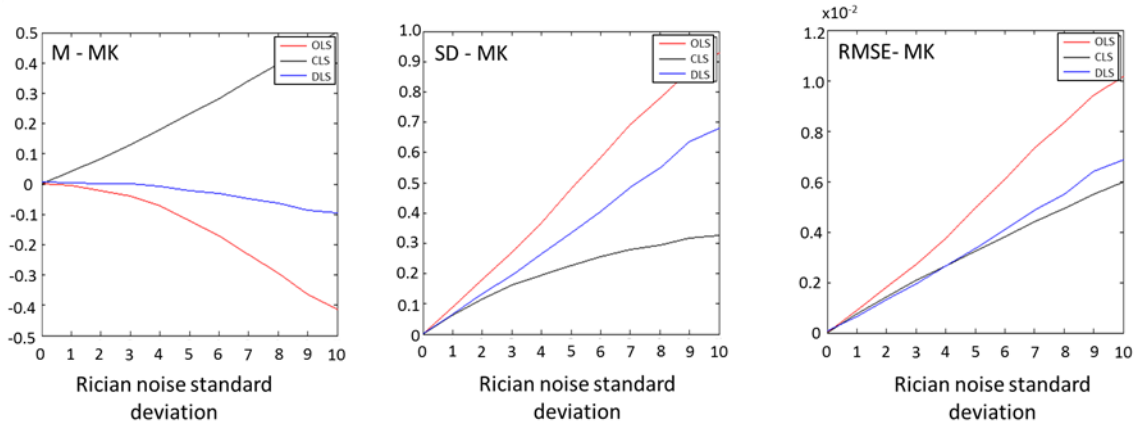



Figure 3.5 - MK values of error mean *M*, error standard deviation *SD*, and root mean square errors *RMSE* computed from 10000 samples for each level of Rician noise.

Figure 3.6 shows the behaviour of MD and FA with the increase of Rician noise level. The scales of the plots in Figure 3.6 can be compared with Figure 3.5 (for the all mean error plots it is used an scale based on the differences of 50% relative to the noise free MK, MD, and FA, while for the error of standard deviation *SD* and *RMSE* the scales are based on differences of 100% and 1.2% respectively). Therefore, once more it is show that MK values are more sensitive to the influence of Rician noise. For the values of MD, DLS seems to produce a value that overlaps the values of OLS (upper panels of

Figure 3.6). Comparative to the other methods, CLS seem to produce values of MD that are more biased (i.e non-null values of *M*); however with inferior values of *SD* and comparable values of *RMSE*. The values of *RMSE* and error *SD* for FA are slightly smaller for the values extracted by CLS (lower panels of Figure 3.6); however CLS is related to underestimated values of FA (negative values of FA error mean *M*).

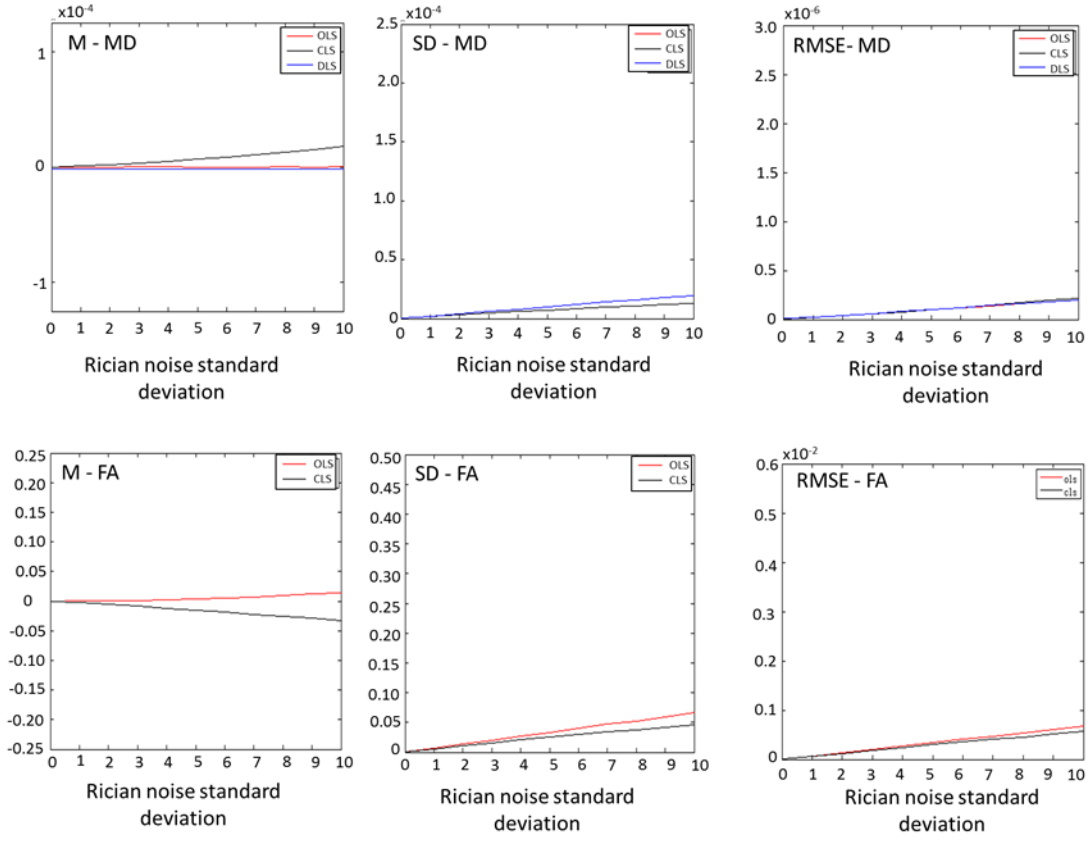



Figure 3.6 – MD and FA values of error mean *M*, error standard deviation *SD*, and root mean square errors *RMSE* computed from 10000 samples for each level of Rician noise.

### 3.3.2. All brain simulation

Similarly to the single voxel analysis, OLS, WLS, and NLS show similar behaviours on all brain simulated data. Therefore, figures for WLS and NLS estimates are not shown below[2]. Figure 3.7 shows the scatter plots between the noise free and Rician corrupted values of MK, MD, and FA extracted using OLS and chNLS. For both methods, MK values seem to have larger error dispersion than MD values which scatter plots are closer to the identity line. Notably, FA values show also large dispersion on scatter plots, in particularly noise free values with low anisotropy seem to be largely influenced by

[2] However scatter plots between noise free and Rician corrupted values for WLS and NLS can be seen in Appendix A

overestimations. Although MK metrics obtained from chNLS are also identical when compared to the extracted values from OLS, MD and FA from chNLS shows slight differences. For the OLS method, lower anisotropy overestimation can even reach values larger than the plausible limit of one. Although these overestimations of anisotropy are still present on the values extracted from chNLS, values of FA for this method do not range out of the plausible value interval. Noise corrupted MD estimates for chNLS never seem to reach values inferior to 0.4 $mm^2.s^{-1}$

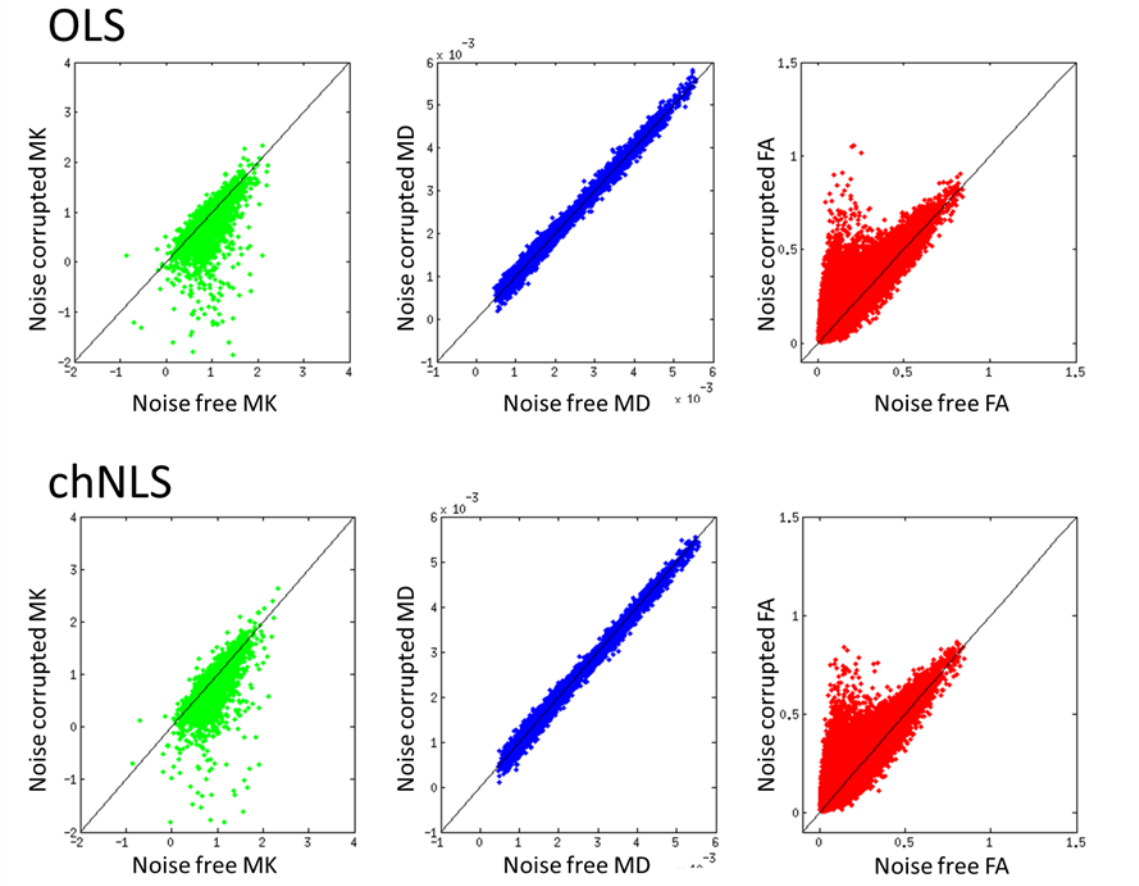



Figure 3.7– All brain MK, MD, and FA scatter plots between noise free values and estimates extracted from DWIs corrupted by Rician noise with standard deviation of 5 (SNR=100) using both OLS and chNLS fitting approaches.

Figure 3.8 shows that both the CLS and DLS produces MK scatters plots with lower dispersions than the methods shown in Figure 3.7. While values of MK from DLS are more uniformly distributed along the identity line, values of MK from the CLS seem to present regions of overestimation, in particular for high noise free values of MK. DLS produces similar MD profiles to the one presented on Figure 3.7. The scatter plots show that CLS produced the MD values that are not only negatively biased but also have the

largest error dispersion relative to the all six analysed methods. Despite a worst behaviour when estimating MD, CLS is the method that shows lower overestimation of FA values.

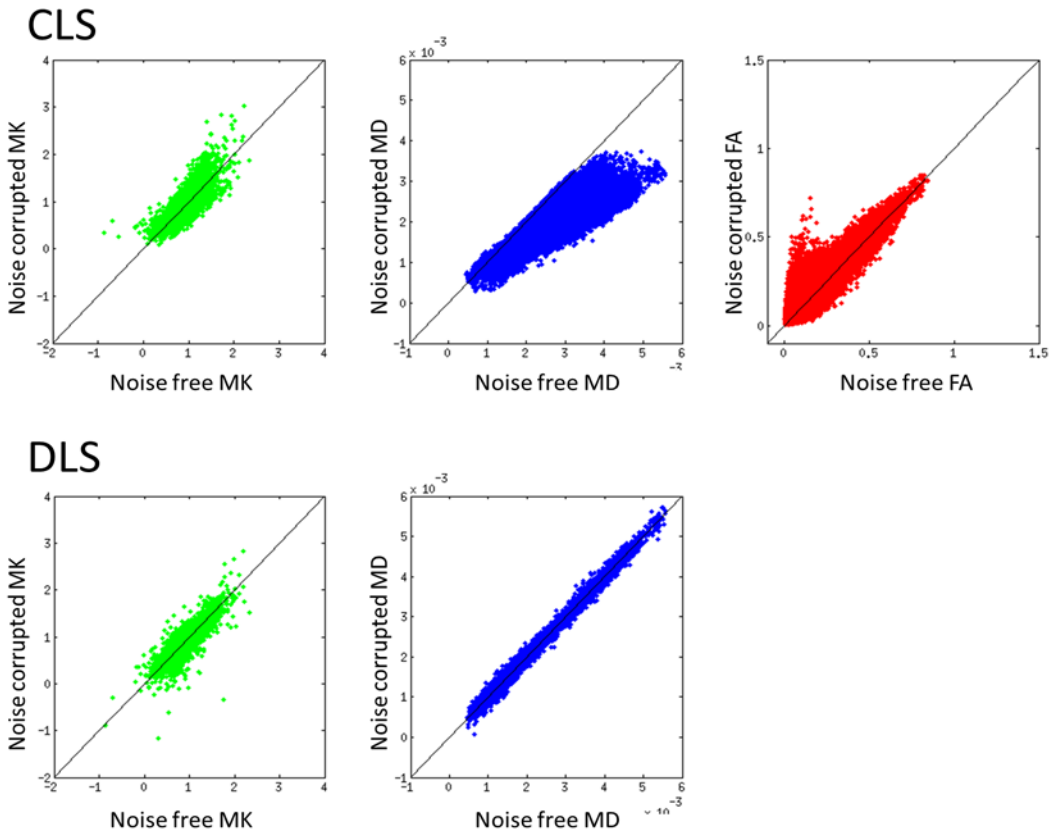



Figure 3.8 – All brain MK, MD, and FA scatter plots between noise free estimates and values extracted from DWIs corrupted by Rician noise with standard deviation of 5 (SNR=100) for both CLS and DLS fitting approach. Notably that scatter plot of FA is not presented for the DLS fitting method once this does not produce estimates of metrics rather than MK and MD.

Table 3.1 shows the mean error *M*, error standard deviation *SD* and root mean square error *RMSE* of the values of MK, FA, and MD extracted using the six fitting method tested on this chapter. In addition, the time that each method required to process all 202315 brain voxels are in the last column. All the values of *M*, *SD*, and *RMSE* across the four methods OLS, WLS, NLS, and chNLS are identical. Notably, for the CLS method MK mean error showed a negative value (which is contrary to the overestimations observed previously on the simulations based on the single voxel). In addition, CLS' absolute values of *SD* and *RMSE* for MK are lower than the ones produced by all other tensor fit methods (i.e. OLS, WLS, NLS, and chNLS). On the other hand, values of MD are higher for CLS. Comparing the MD and MK values with the

ones extracted from the tensor fitting methods, DLS produced the absolute lower values of *M*, *SD*, and *RMSE*.

Table 3.1 Values of mean error *M*, error standard deviation *SD* and *RMSE* of the invariant maps MK, MD, and FA extracted from the six fitting methods (OLS, WLS, NLS, chNLS, CLS, and DLS). This values correspond to the all brain simulated data corrupted by Rician noise with sd=5 (SNR=100). The computational time required for processing the 202315 all brain voxels are presented on the last column of the table. Each value has an estimation error of 5 units in the last decimal algorism presented on each cell of the table.

| | **MK** | | | **MD** ($x10^{-4}mm^2.s^{-1}$) | | | **FA** | | | **Time** |
|---|---|---|---|---|---|---|---|---|---|---|
| **methods** | **M** | **SD** | **RMSE** | **M** | **SD** | **RMSE** | **M** | **SD** | **RMSE** | **(s)** |
| OLS | - 0.014 | 0.071 | 0.072 | 0.002 | 0.388 | 0.388 | 0.029 | 0.048 | 0.056 | 170 |
| WLS | - 0.001 | 0.069 | 0.070 | - 0.078 | 0.438 | 0.438 | 0.028 | 0.046 | 0.053 | 170 |
| NLS | - 0.010 | 0.069 | 0.070 | - 0.020 | 0.386 | 0.386 | 0.028 | 0.046 | 0.053 | 4970 |
| chNLS | - 0.010 | 0.070 | 0.070 | - 0.020 | 0.384 | 0.384 | 0.028 | 0.046 | 0.054 | 5010 |
| CLS | - 0.004 | 0.058 | 0.059 | - 0.624 | 2.144 | 2.234 | 0.024 | 0.040 | 0.054 | 1470 |
| DLS | - 0.005 | 0.048 | 0.049 | 0.001 | 0.389 | 0.389 | - | - | - | 35 |

At last, Figure 3.9 shows the scatter plots between values of free noise MD and values of MK re-estimated from the DWIs corrupted with the Rician noise level of SD=5 (green points on each panel). Each plot corresponds to a different fitting method: OLS (panel A), DLS (panel B), and CLS (panel C). Over each panel, scatter plots of the noise free values of MK vs MD are plotted on black to provide a reference to the noise effects. DLS is the method that shown a smaller deviation from the noise free profile (panel B). While OLS shows to produce underestimates of MK for lower values of MD (panel A), CLS shown to generate MK overestimates for larger values of MD (panel C).

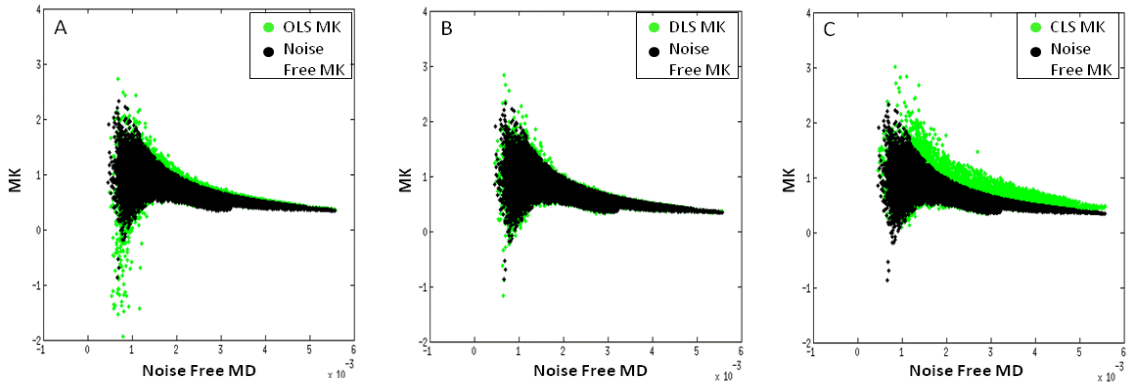



Figure 3.9 – Scatter plots between values of free noise MD and values of MK (back points correspond to noise free values of MK which are plotted over the green points corresponding to the MK results from data corrupted with Rician noise with SD=5). The three different panels are related to three different fitting methods: OLS (panel A), DLS (panel B), and CLS (panel C).

Relative to the processing times, the newly proposed DLS is the method less computationally demanded, followed by the tensor linear least squares solutions OLS and WLS. Adding constrains on the linear approach, CLS shown to be nine times more computational demand than the unconstrained linear frameworks. The most computational demanding methods are the ones that followed the non-linear framework i.e. the NLS and chNLS methods which required respectively 4970 s and 5010 s to process all brain voxels.

## 3.4. Discussion

**Similar performances between OLS, WLS, and NLS**

The ordinary linear least squares (OLS) solution erroneously assumes that log-transformed samples have a uniform variance $\sigma^2_{\ln DW}$. However, as previously mentioned on section 2.3.1, a more accurate approximation is to assume uniform variance $\sigma^2_{DW}$ of diffusion-weighted signals on the non-transformed framework. Following this, the log-transformed variance $\sigma^2_{\ln DW}$ is described by Eq. 2.3.14, which is re-written below:

$$\sigma^2_{\ln DW} = \frac{\sigma^2_{DW}}{DW^2}\,. \qquad \text{(Eq. 2.3.14)}$$

Analysis Eq. 2.3.14, one can note that samples with smaller diffusion-weighted signals in the log-transformed framework should have larger variance $\sigma^2_{\ln DW}$. Ignoring this by using the OLS solution, smaller log diffusion-weighted samples are incorrectly associated to smaller variance values that it should be theoretically assumed. Consequently, samples with larger signal attenuation are strongly weighted when the OLS approach is used. As demonstrated experimentally on section 2.3.3, larger signal attenuations are the samples more influenced by the rectified noise. Therefore, OLS is logically more influenced by artefacts than other fitting methods as the WLS and NLS. WLS uses Eq. 2.3.14 to weight the log-transformed diffusion signals. However, it is expected that WLS is still more sensitive to Rician noise than non-linear approaches as the NLS since it uses the measured diffusion-weighted signal as the value of $DW$ on Eq. 2.3.14, which is corrupted by Rician noise. This limitation is well avoided with NLS by handling the diffusion-

weighted samples directly on the non-transformed space and thus without requiring any approximation from weighing them.

Despite previous DTI studies had experimentally demonstrated that NLS provides better estimates than the ones provided by WLS and OLS (Jones & Basser, 2004; Jones & Cercignani, 2010), in this chapter the results for the single voxels simulations and for the all brain simulated data clearly shows that OLS, WLS, and NLS on DKI have similar behaviours, e.g. Figure 3.3 and Table 3.1. These can be explained by the different formulation followed by DTI and DKI. Once assumed that water displacements are well described by Gaussian probability distributions, DTI is not formulated to handle any degree of non-linearity, and non-linearity induced by rectified noise will be a significant source of DTI artefacts. Therefore using methods than handle differently the Rician noise variance will naturally show different performances. On the other hand, assuming the existence of deviation of water Gaussian distribution, all degree of non-linearity introduced by the existence of microstructures barrier as well as the introduced by Rician noise will be modelled in DKI by values of kurtosis. Consequently, the bias induced by rectified noise is less significant on diffusion tensors extracted from DKI and the previous differences shown on OLS, WLS, and NLS in DTI studies are neglected.

**Effects of Cholesky tensor decomposition on DKI**

On the single voxel simulations, incorporating the diffusion tensor Cholesky decomposition on DKI formulation did not shown evident differences on the estimates of MK, MD, or FA metrics relative to the ones computed from NLS, WLS, or OLS methods (Figure 3.3). However slight changes are observed on the all brain simulation on MD and FA scatter plots (Figure 3.4). In particular, noise corrupted MD estimates never seem to reach values inferior to 0.4 $mm^2.s^{-1}$ which is an effect of the corrections of implausible negative values of diffusion by chNLS. More interesting is the reduction of overestimations of FA. For the FA scatter plots of OLS, WLS, and NLS, values of anisotropy are largely overestimated for lower free noise values of FA, and therefore, lower values of anisotropy on real data can be mistakenly classified as a voxel from high anisotropy tissue. This artefact can be associated by the repulsion effects reviewed on section 2.3.3 as well as negative estimates of diffusion which on Eq. 2.3.20 it will not only reflect on large values of FA but also values larger than the plausible maximum limit of one (Koay et al., 2006). Consequently, avoiding implausible negative values of diffusion by describing the diffusion tensor on its Cholesky framework, chNLS is able to

reduce overestimates of FA. Notably, the simulations based on the values of diffusion and kurtosis for a single voxel is not significant to detect such differences since they are based on relative high values of FA.

Despite the advantages provided by chNLS, values of FA are still corrupted by overestimations due to repulsion effects, and constrains imposed on the diffusion tensor do not seem to affect differently the estimates of MK (Figure 3.7). Moreover, chNLS is the tested fitting method that shown to require more computational processing time (see Table 3.1). Therefore, chNLS is not considered a perfect alternative to the OLS, WLS, and NLS on DKI's framework.

**Noise effects on CLS**

In 2011, Tabesh et al. showed that the CLS formulation (named as the CLLS-QP algorithm on their paper) ensures that the diffusion and kurtosis estimates remain within a physically plausible range and MK estimates provides inferior values of *RMSE* relative to the estimates from the OLS solution. Despite these observations having been supported by the results presented in this chapter (Figure 3.5 and Table 3.1), some problems while using CLS were also observed. In particular, this method produces MK values that are positively biased for larger values of noise free kurtosis and negatively biased MD for larger values of noise free diffusions (Figure 3.8). Moreover, the underestimates of MD seem also to be associated with larger error variance, while MK overestimates are also related to large noise free values of MD (panel C of Figure 3.9). Interestingly, in opposite to the values present on Figure 3.5, CLS mean error *M* for the all brain simulation presented a negative value (Table 3.1). This suggests that unlike large values of noise free MD and MK, MK estimates for lower noise free values of MD and MK are in fact underestimated. Such will have a larger impact on quality metrics computed for all brain voxels once they corresponds to the type of voxels with larger population (see Figure 3.2).

The diversity of artefacts on CLS can be related to the imposed constrains on the CLS formulation. Despite MK values for CLS do not show any negative implausible value (Figure 3.4 and Figure 3.8), the underestimation suggested on Table 3.1 can be related to imperfection on the correction provided by the lower kurtosis constrain (Eq. 2.4.18) on two cases of corrupted voxel' intensity: 1) intensities of values which were inferior to its noise free version but did not range out of the plausible interval values of kurtosis; and 2) values that on the initial state of CLS algorithm show negative estimates

and after convergence were wrongly corrected to an inferior value relative to its noise free version. Although less likely, the same effects on the second type of voxel's intensity can produce some of the MK overestimations observed for smaller values of noise free MD presented on Figure 3.9. Regarding the limitations of the upper kurtosis constrain (Eq. 2.4.19); these can explain the artefacts for the higher free noise values of diffusion and kurtosis. For data extracted for OLS (panels A of Figure 3.9), it can be observed that larger values of kurtosis overestimates are associated to higher values of free noise MD ($>1.5x10^{-3}mm^2.s^{-1}$). Once OLS correspond to the initial state before CLS' convergence algorithm, the higher values of free noise MD values are the ones that will be constrained by the upper kurtosis constrain. Observing panel C of Figure 3.9, one can see that overestimations are not reduced after applying CLS, instead they are even more overestimated. This suggests that the imposed upper kurtosis limit is unable to avoid implausible values of kurtosis; alternatively it decreases diffusion values in order to satisfy Eq. 2.4.19, and causing MD to be largely underestimated (Figure 3.8).

Due to the excess of artefacts, CLS should not be used on Cam-CAN analysis without further improvements on its formulation. For instance, CLS could show better performances by testing different values on the variables of constrains. In future a trade-off between CLS accuracy and the reduction of the artefact introduced by constrains could be studied.

**Effects of the Rician nature of noise on DKI**

As shown in section 2.4.3, DKI is more sensitive to noise than DTI however it produces more accurate diffusion parameters. In this chapter, the results also suggests than kurtosis tensor is more sensitive to noise than the diffusion tensor. For example, despite diffusion and kurtosis tensors being estimated on the same DKI model, result from Figure 3.5 shows that standard deviation of MK reach values which correspond to differences near to 50% of the MK expected value, while the standard deviation for MD reach only amplitudes around 5% (Figure 3.6).

The larger noise sensitivity for MK parameters can be intuitively explained by its geometrically meaning. Similar to diffusion which is related to the slope of log diffusion-weighted curve decay (see section 2.3.3), the kurtosis is related to its concavity. Consequently, kurtosis is more sensitive to Rician noise than diffusion as a concavity is more easily corrupted by noise than the slops on general quadratic models. To illustrate this, Figure 3.10 shows 10000 curves simulated for values of scalar kurtosis and diffusion

from diffusion-weighted samples corrupted with Rician noise. This simulations are performed based on two values of free noise diffusion, $0.5x10^{-3}mm^2.s^{-1}$ (panel A and C) and $1.5x10^{-3}mm^2.s^{-1}$ (panel B and D), while kurtosis was always set to a value of 1. All curves associated to negatives re-estimations of kurtosis are plotted on red which shows to correspond to negative concavities. Despite none of the re-estimated curves seem to be related to implausible values of diffusion (i.e. positive slops), Figure 3.10 presents estimates of implausible kurtosis on the panels associated with the lower directional diffusion value. Interestingly, results from Figure 3.10 can also explain why MK is likely underestimated for the fitting methods OLS, WLS, NLS, chNLS, and CLS. As the average of the negative bias directional kurtosis, MK computed by Eq. 2.4.23 will be likely biased.

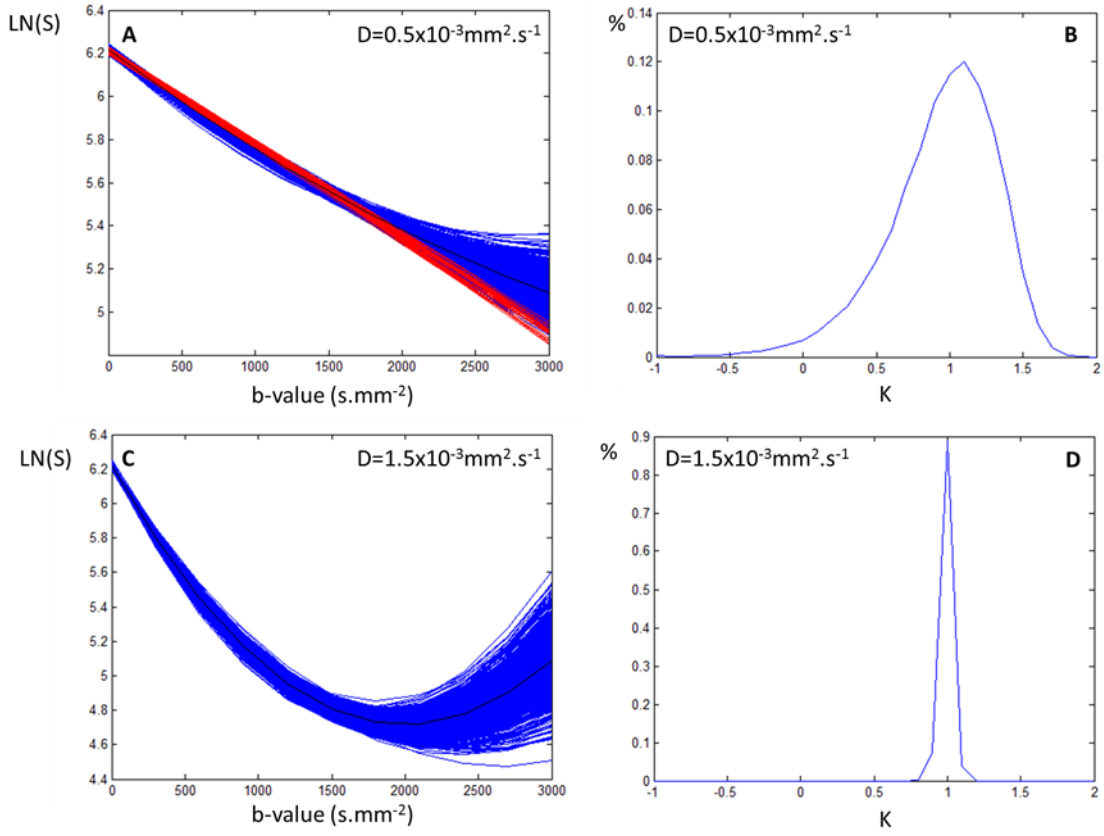



Figure 3.10 – Re-estimation of diffusion-weighted signal decay in order to b-value. Panel A and C shows curves for directional diffusions of $0.5x10^{-3}mm^2.s^{-1}$ and $1.5x10^{-3}mm^2.s^{-1}$, while on panels B and D shows respectively the histograms of the re-estimated values of kurtosis which noise free values is 1. The curves are based on estimates of three DWIs sampled corrupted with Rician noise (SNR=100) for three b-values samples: b-value 0, 1000, and 2000 $s.mm^{-2}$

Figure 3.10 also reveals that smaller values of diffusion are related to larger concavity errors, which can explain the larger value dispersions shown in Figure 3.9 for

lower values of free noise diffusion, as well as the large standard deviations shown on Figure 3.3 and Figure 3.4 relative to the single voxels simulations, noting that the free noise MD value of $0.1x10^{-3}mm^2.s^{-1}$ is related to the lower MD values represent of the all brain simulated voxels (Figure 3.2). Notably this show the importance of performing more general DKI evaluation analysis as the one based on the simulation on an entire brain. Despite obvious pitfalls on the all brain simulation performed in this chapter, as not ensuring that its noise free conditions is completely free from implausible values of diffusion and kurtosis over all voxels, the all brain simulations provide a way to evaluate simultaneously the effects of noise on a significant range of diffusion and kurtosis values.

Regarding the effects of Rician noise, the only question which was not clearly answered is if the rectified noise induces overestimation of non-linearity as assumed by Jones and Basser (2004) , and consequently resulting on overestimation of kurtosis (for more details see section 2.4.3). If this was a significant effect on brain data, overestimations of MK were visible on large free noise values of diffusion. However such is not largely observed on Figure 3.9 for panel A and B, instead such MK range of values seems to associated to values that data is less corrupted by bias (the overestimations for the CLS on panel C were already explained above on '*Noise effects on CLS*'). Therefore, one can conclude that such overestimations are not significant for the b-values used by Cam-CAN.

**Better performance of DLS**

Extracting MD and MK values directly from DWI images using a single step provides estimates with larger accuracy. DLS is the approach that shows smaller impacts of bias introduced by noise (Figures 3.4, 3.8, and 3.9), it shows smaller error variances when compared to OLS, WLS, NLS and chNLS (Figure 3.4 and Table 3.1), and it was less sensitive to the increase of Rician noise levels (first panel of Figure 3.5). Moreover this algorithm is computationally very efficient (Table 3.1).

The improved accuracy of DLS can be related to several reasons. First DLS fit is less sensitive to noise since it involves an inferior number of unknown parameters to be estimated (Kay, 1993). Secondly, for not being associates to tensor estimates, fitting directly MK and MD will not have any influence of artefacts as the eigenvalues repulsion effects mentioned by Pierpaoli and Basser (1996). More importantly, as they not involve any average of directional metrics that can be highly corrupted by Rician noise bias, the MK and MD estimates will be more robust against under or overestimation effects.

Although related to better performances, DLS cannot extract other invariant metrics rather than MD and MK. Once for Cam-CAN metrics as FA, directional diffusivity $\lambda_{//}$ and $\lambda_{\perp}$, and directional kurtosis $\lambda_{//}$ and $\lambda_{\perp}$ can provide important features to detect ageing changes on anisotropy or axonal and myelin integrity, DLS cannot be seen as a perfect alternative to tensor fit methods for not provide the same amount of information. Therefore, on later chapter OLS will be used to extract all promising invariant maps (as the alternatives of tensor fitting method did not show satisfactory advantages), while DLS will be used as a more accurate method to validate the finding given by MK and MD metrics. It is also important to note that OLS will not be used to study ageing changes without the evaluation of the proper pre-processing steps to reduce the impacts of noise (that will be performed on the next chapter).

## 3.5. Chapter Summary

In this chapter, novel findings on the behaviour of DKI by using different fitting methods are reported. First of all, it was demonstrated that WLS and NLS produce similar corrections than the OLS solution on DKI, which is opposite to results previously reported on DTI studies, (e.g. Jones & Basser, 2004; Jones & Cercignani, 2010). Secondly, it was shown that constrains on the diffusion tensor using Cholesky decomposition did not have any impact on the estimates of MK. Despite being one of the approaches more followed in recent papers (e.g. Fieremans, Jensen, & Helpern, 2011; Gao et al., 2012), the CLS method based on the approach suggested by Tabesh et al. (2011) introduces artefacts on the estimates of the diffusion and kurtosis tensors, for example large values of kurtosis seems to be overestimated while correspondent diffusion values are largely underestimated. The simple proposed method DLS to extract directly MK and MD in a single step shown to provide estimates less biased and less sensitive to the increase of noise levels. Relative to the other fitting methods analysed, OLS shows be most adequate to extract the diffusion and kurtosis tensor for the Cam-CAN project. DLS will be used as a more accurate method to support the finding given by MK and MD metrics. However due to the high kurtosis sensitivity to Rician noise on DWIs, OLS will not be used to study ageing changes without the proper pre-processing steps to reduce the impacts of noise.

In addition, new DKI artefacts were reported on this sections which include the larger kurtosis sensitivity to noise relative to diffusion metrics and the high degree of underestimation of kurtosis for low values of diffusion. It was also shown that Rician noise does not introduce significant overestimation on kurtosis for the b-values used by Cam-CAN i.e. b-values inferior or equal to 2000 $s.mm^{-2}$.

# Chapter 4

# Optimization of DKI Pre-processing

## 4.1. Introduction

As shown in the last chapter, the accuracy of DKI is highly influenced by the Rician MRI noise, making it inappropriate to use methods for extracting the diffusion and kurtosis tensors without first applying algorithms to increase the DWI's signal to noise ratio (SNR) and algorithms to remove artefacts as the ones produced by subject motion or eddy currents.

SNR can be improved by averaging neighbouring voxels using image regularization. For instance, smoothing DWIs with a 3D Gaussian kernel is considered an essential pre-processing step (Jones & Cercignani, 2010). In addition to increasing SNR, a Gaussian kernel filter is also shown to reduce the impacts of motion and distortion artefacts (Jones & Cercignani, 2010). Another simple way to improve SNR is to decrease the image resolution. The Cam-CAN datasets have a voxel size of $2x2x2mm^3$, but resolution can be decreased using downsampling algorithms. Previous studies have suggested more complex algorithms for noise removal, including approaches that directly account for the diffusion Rician noise bias, such as techniques based on a Bayesian framework (Basu, Fletcher, & Whitaker, 2006) or on non-local means algorithms (Wiest-Daesslé, Prima, Coupé, Morrissey, & Barillot, 2008). Although the latter approaches had shown potential in providing more accurate correction of data corrupted by Rician noise than simpler regularization techniques, the Rician nature of the noise distribution can be compromised by the influence of motion or eddy currents artefacts.

While eddy current distortion can be significantly cancelled by using specific diffusion MRI sequences, such as the Twice Refocused Spin Echo (TRSE) mentioned in

section 2.2.3, head motion artefacts are experimentally inevitable. The most common algorithms for head motion and eddy current correction are based on image registration methods (Rohde, Barnett, Basser, Marenco, & Pierpaoli, 2004) where a reference image is used for correcting all DWIs. For example, a frequently adopted approach is to register all DWIs relative to the intensity of a DWI acquired with b-value of 0 $s.mm^{-2}$ using affine transformations with twelve degrees of freedom (to correct simultaneously motion and residual eddy correct artefacts) or rigid body transformation with six degrees of freedom (to correct only motion artefacts). Despite these algorithms still being common in diffusion MRI studies, (e.g. Gao et al., 2012), studies indicate that such approaches are inaccurate (e.g. Bastin, 1999). The contrast of each DWI image is b-value and direction dependent and therefore correction algorithms based on image intensity are not able to distinguish the inherent image differences from artefact distortions.

The aim of this chapter is to investigate adequate pre-processing steps to reduce the impact of noise, motion and image artefacts on the Cam-CAN diffusion-weighted data. Firstly the frequently used registration based method for artefact removal is tested. Since the Cam-CAN's DWIs were acquired using a TRSE, the effects of eddy currents are reduced, and thus the registration method is based on a six degree of freedom rigid transformation for only correcting motion artefacts. Secondly, the effects of applying a 3D Gaussian kernel filter are studied. The extent of noise removal as well as the expected ageing effects will depend on the filter size which is characterized by the full-width-at-half-maximun (FWHM) of the Gaussian kernel (Jones & Cercignani, 2010). Therefore, the optimal value of FWHM is defined as the one which is more sensitive to the age differences and provides a better trade-off between image resolution and quality. The Gaussian kernel filter for noise removal is selected on this study instead of more robust approaches as the ones proposed by Basu et al. (2006) or Wiest-Daesslé et al. (2008) since the assumption of Rician noise cannot be ensured due to the corruption effects of motion artefacts. In order to study possible advantages on DWIs with larger voxel sizes, the effects of the Gaussian kernel filter with different values of FWHM are also tested on data acquired and downsampled to a voxel resolution of $2.7x2.7x2.7mm^3$. Finally the most promising pre-processing steps will be combined with the direct linear least squares fitting method described in the previous. Based on the results of this chapter, the most adequate DKI procedure to detect age differences will be selected and applied to the Cam-CAN datasets acquired so far. The preliminary ageing results are reported in Chapter 5.

## 4.2. Methods

### 4.2.1. Intensity registration-based algorithms to correct head motion

For head motion correction, DWIs for each Cam-CAN subject are registered to the first brain volume of DWIs acquired, which corresponds to a b-value of 0 $s.mm^{-2}$. This was done using the SPM8 (Statistical Parametric Mapping, University College London, UK) function *spm_realign*. In order to analyse the influences of the contrast differences that are only dependent on the b-value, the registration algorithm is applied to the brain DWIs simulated on chapter 3 without being corrupted with Rician noise. Then the algorithm is tested on real data from 28 Cam-CAN subjects that visually did not show large amount of motion during image acquisition. As mentioned in the previous chapters, Cam-CAN data for each subject was based on 63 volumes of DWIs which corresponded to 30 gradient directions for b-values 1000 and 2000 $s.mm^{-2}$ and 3 images acquired with a b-value 0 $s.mm^{-2}$. Each acquired image has a voxel size of 2x2x2 $mm^3$.

### 4.2.2. Comparison of different FWHMs for smoothing Cam-CAN data

After analysing if the registration algorithm for head motion correction is adequate on diffusion weighted data, DWIs for 28 Cam-CAN subjects are segmented into brain and non-brain voxels using FSL brain extraction tool (Smith, 2002). Then a Gaussian filter is applied on the brain extracted data using the SPM function *spm_smooth*. To find the optimal value of kernel size, data is filtered using eight different values of FWHM (values equally spaced between 0.5 mm and 4 mm). From the non-smoothed and smoothed versions of the data, the kurtosis tensors for each voxel image is estimated using ordinary linear least squares (see section 2.4.1) and the MK values are computed using Eq. 2.4.23. In this chapter we will focus only on the MK values since they have been shown to be the more sensitive to noise effects than diffusion parameters (chapter 3). The quality of the extracted MK images for all different FWHM values are quantified by the percentage of negative values of kurtosis, signal to standard deviation ratio, and contrast to standard deviation ratio (Kingsley & Monahan, 2005).

Contrast to standard deviation ratio (CSR) is similar to the well know contrast-to-noise ratio (CNR) which quantifies the data's ability to distinguish two types of tissue,

and thus larger values of CNR will correspond to a better quality. CNR is used to distinguish populations that have uniform characteristics. However, when analysing tissues on real data using regions of interest, the ability to distinguish two types of tissue will depend on biological heterogeneity. For this reason, the concept of CSR ratio noise is formulated to deal with experimental data (Kingsley & Monahan, 2005), and for values of MK it can be computed using the following expression:

$$CSR(MK) = \frac{\langle MK\rangle_1 - \langle MK\rangle_2}{\sqrt{var(MK_1) + var(MK_2)}} \qquad \text{(Eq. 4.1)}$$

where $\langle MK\rangle_1$, $\langle MK\rangle_2$, $var(MK_1)$, and $var(MK_2)$ are the means and variances of MK of the two regions of interest (ROI) corresponding to two different tissues. Similarly the concept of signal to standard deviation ration (SSR) replaces the (SNR) and is computed for a tissue type *i* as:

$$SSR(MK) = \frac{\langle MK\rangle_i}{\sqrt{var(MK_i)}} \qquad \text{(Eq. 4.2)}$$

In this study, the two types of tissues analysed are white and grey matter. In order to define these two tissue types brain masks are obtained by segmenting each subject's structural images with voxel resolution of 1 mm using the SPM functions *spm_preproc* and *spm_pre2n* (Cam-CAN structural images were acquired using magnetization prepared rapid acquisition with gradient echo - MPRAGE), panel A of Figure 4.1. The images are then registered to DWI space using the functions *spm_coreg* and *spm_reslice* (the reference image for the DWI space is selected as the mean of the volumes acquired with b=0s.mm$^{-2}$), panel B of Figure 4.1. Notably, the *spm_coreg* function is based on morphological similarity rather than maximization of the correlation between image intensities, therefore the intrinsic contrasts differences between b=0 s.mm$^{-2}$ DWIs and structural MPRAGE have a reduced effect on corrupting the performance of this registration technique.

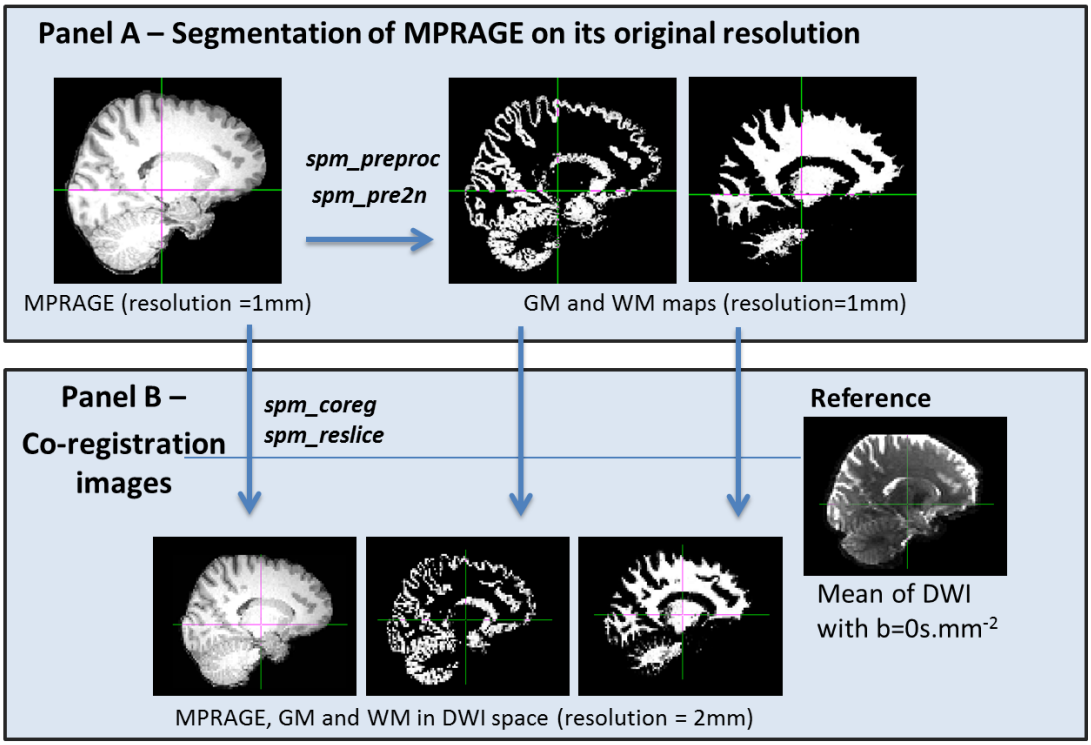



Figure 4.1 – Example of the two steps to compute grey and white matter maps for one Cam-CAN subject: first the grey and white matter maps are obtained by segmentation of MPRAGE images with resolution of 1mm (panel A); secondly images are moved to DWI space by co-registration using as reference volume the mean of DWIs acquired with b-value of 0 $s.mm^{-2}$ (panel B).

To decrease the biological heterogeneity within tissue types, white and grey matter voxels will only be analysed on the neocortex by defining an ROI in stereotactic coordinate space (Evans et al., 1992) using the template from the Monteal Neurological Institute (MNI), Figure 4.2 panel A. This ROI is registered to each subject's b=0 $s.mm^{-2}$ DWI space using a transformation matrix computed by FSL non-linear registration *fnirt*, and then, combined with the white and grey matter maps. Since the FSL registration technique is based on image intensity, the transformation matrix to convert the neocortex ROI to the DWI space should be computed between images with the same type of contracts for a better performance. Therefore instead of using the MNI template and a volume of DWI, the transformation matrix is computed based on registration between the MNI and the MPRAGE previously converted to the DWI's space, panel C of Figure 4.2.

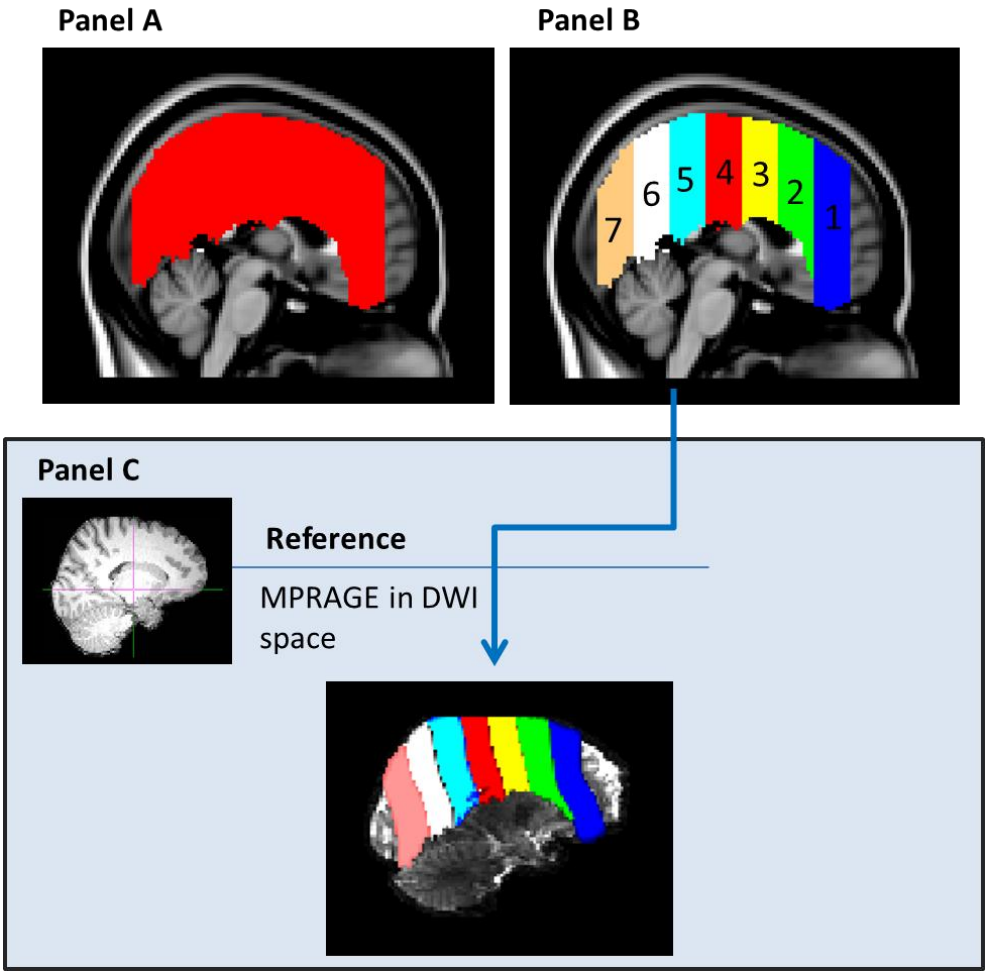



Figure 4.2 – Sagittal view of the neocortex ROI defined in a MNI stereotactic coordinate space (panel A). To analyse age differences, neocortex ROI is subdivided on 7 sub-regions (panel B). The neocotex ROI are converted to each individual subject DWI space using the transformation matrix resulted on the registration of the MNI template to the MRAGE subject images previously converted to the DWI's space (panel C).

As shown in previous studies, the sensitivity of diffusion MRI images depends on the filter size (Jones & Cercignani, 2010), thus, in addition to the metrics of image quality, the optimal Gaussian Kernel FWHM is selected according to its sensitivity to detecting age changes. For this purpose, the 28 CamCAN subjects are divided in two age groups, i.e. 14 young adults and 14 middle-aged adults, with age distributions as shown in Figure 4.3. The mean MK inter-group differences are quantified by p-values computed using a 2-sample t-test on both white and grey matter maps. As suggested by previous DTI studies (Sullivan & Pfefferbaum, 2006), brain ageing seems to occur spatially on different ranges, thus the neocortex ROI is divided in seven sub-regions which are separately analysed for ages differences. The sub-regions are enumerated in anterior-posterior order (Figure 4.2, panel B).

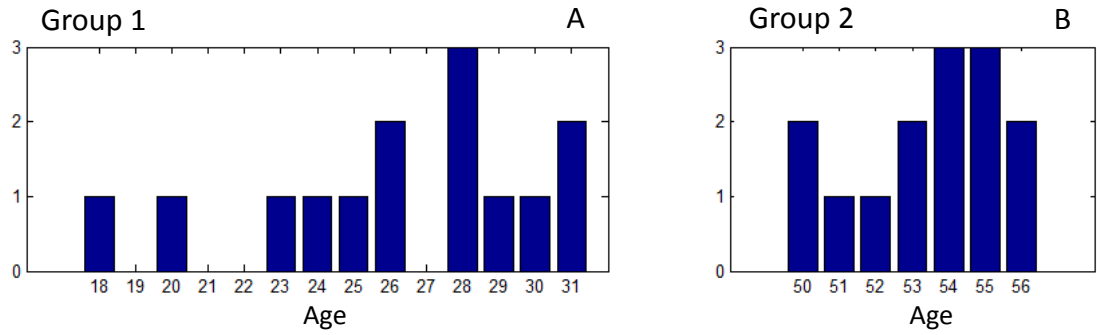


Figure 4.3 - Age distributions of the two age groups: young adults (mean 26.2, sd 3.9) on panel A, and middle-age adults (mean 53.4, sd 2.0) on panel B.

### 4.2.3. Reproducibility of DKI and effects of image resolution

Since MK is a new diffusion metric, its reproducibility should be verified. Therefore seven scan repetitions of the diffusion acquisition scheme selected by Cam-CAN are performed on a single male subject of 25 years old. The reproducibility is analysed by computing the within-subject coefficient, which is the standard deviation expressed as the percentage of the mean (Falangola et al., 2008). To account for the effects of the size of the Gaussian kernel used to reduce the impact of image artefacts and noise, the different values of FWHM tested on the last section are used to smooth this single subject's diffusion weighted dataset. Then the metrics of image quality (percentage of negative voxel of MK, CSR between white and grey matter, and white and grey matter SSR) are computed and analysed. Since for this subject no MPRAGE was recorded, white and grey matter maps are extracted directly from the mean of the DWIs acquired with a b-value of 0 $s.mm^{-2}$.

To study the potential benefit of data acquired with a larger voxel resolution, four independent datasets with voxel size of 2.7 mm were also acquired on the same subject using the Cam-CANs b-values and b-vectors. The reproducibility and image quality of this dataset are analysed using the same procedures and metrics mentioned in the previous paragraph.

With the objective to reduce the acquisition time, all the independent datasets are acquired using a partial brain field of view (FOV). Consequently the neocortex ROI presented on Figure 4.1 could not be used. Therefore a prefrontal brain ROI is selected directly on the DWI's space, as shown in Figure 4.4. Notably, this region of interest is similar to the ones used in previous DKI ageing studies (e.g. Falangola et al., 2008; Helpern et al., 2011), and similar to the one used in Chapter 5.

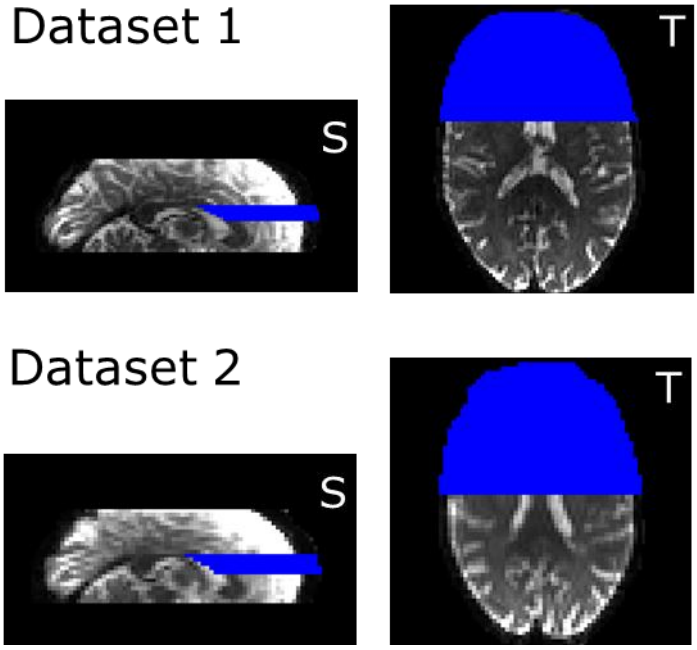



Figure 4.4 – Sagittal (S) and axial (A) views of the prefrontal ROI used to analyse eleven independent datasets acquired on the same male subject of 25 years old. Seven datasets are acquired with a resolution of 2mm (dataset 1), while the other four with a resolution of 2.7mm (dataset 2). Region of interest are defined directly on 5 b-value=0 $s.mm^{-2}$ axial slices by the regions extended from the most anterior point containing brain tissue to the dorsal border of the genu of the corpus callosum.

### 4.2.4. The effects of downsampling

The image quality as a function of the Gaussian kernel FWHM is also studied on downsampled data. This analysis will be performed on the 28 Cam-CAN subjects' data after it has been downsampled to a resolution of 2.7x2.7x2.7 $mm^3$ using a cubic spline in Matlab. A prefrontal ROI is used so that the analysis described in this section could be compared to the one described in the last section. This prefrontal ROI is defined in stereotactic coordinate space (Figure 4.5) and then registered to each subject's downsampled DWI space using the same procedures described in section 4.2.2. The image quality metrics and p-values related to the ageing differences are also computed for the non-downsampled data using the prefrontal ROI to provide a reference for the downsampled data.

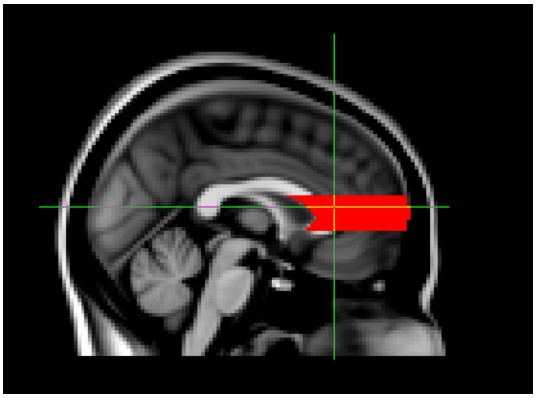


Figure 4.5 – Sagittal view of the prefrontal ROI defined in a MNI stereotactic coordinate space. ROI is defined on 10 selected axial slices by the regions extended from the most anterior point containing brain tissue to the dorsal border of the genu of the corpus callosum.

### 4.2.5. Benefits of using MK extracted from direct linear least squares

For all the pre-processing analysis described so far, MK values were extracted for the OLS fitting method. Therefore, to validate the results of different FWHMs for smoothing Cam-CAN data, the methodology described on section 4.2.2 is repeated for the MK values extracted using the direct linear least squares (DLS) described to the last chapter.

## 4.3. Results

### 4.3.1. Registration based algorithm to correct head motion

Figure 4.6 shows the six estimated parameters of the six degrees of freedom required to align the volumes of DWIs to the first volume acquired with b-value=0 $s.mm^{-2}$. The series of points on each panel are sorted in the same order that volumes were acquired (see on top of Figure 4.6). Despite the simulated volumes of DWI were not corrupted for motion artefacts, the six parameters required to align the volumes presented non-zero values, showing that additional translation and rotation are introduced to DWIs which does not correspond to subject motions. The parameters estimated for b-value 2000 $s.mm^{-2}$ show a larger variance than the ones estimated for b-value 1000 $s.mm^{-2}$. The same profile was observed on real data, of which 6 examples (3 for each age group) are plotted in Figure 4.7. One can observe that for both simulated and real data, translated parameters are in the order of 5 mm, while the rotation ones are in the order of 0.1º. The registration based methods for head motion correction are not applied in the analysis of

the following sections since they show to corrupt DWIs with additional motion and rotations rather than correcting for head motion artefacts.

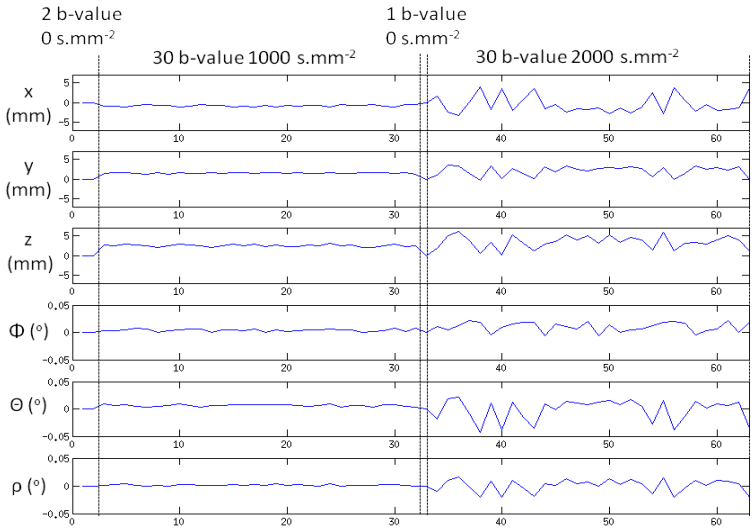



Figure 4.6 – Estimation of the six parameters of the rigid transformation to align 63 simulated volumes relative to the first volume which corresponds to DWIs with contrast of b-value 0 s.mm$^{-2}$. The b-values for the 63 volumes are marked on the top of the figure.

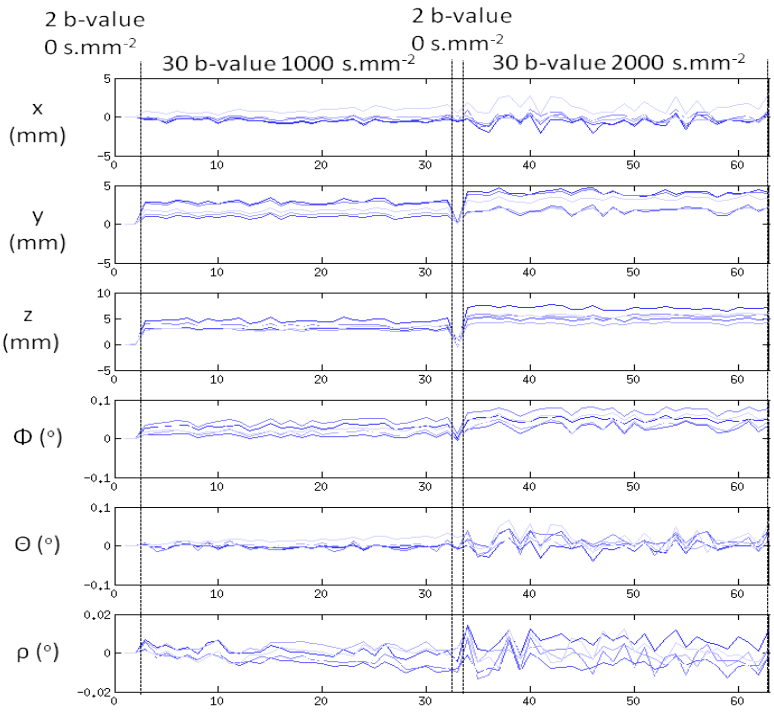



Figure 4.7 – Estimation of the six parameters of the transformation to align the 63 volumes relative to the first b-value 0 s.mm$^{-2}$ volume for six CamCAN's subjects. The b-values for the 63 volumes are marked on the top of the figure.

### 4.3.2. Optimal value of Gaussian kernel FWHM

For all 28 Cam-CAN subjects, Figure 4.8 represents the values of different quality metrics for MK images as a function of the Gaussian kernel size. On panel A and B, both white and grey matter showsSSR that increases with the increase of Gaussian FWHM. In particular, non-smoothed DWIs which correspond to the values of FWHM = 0 mm show low SSR values; however when the higher degree of smoothing is applied (i.e. FWHM = 4mm), the SSR values are on average 9 and 25 times larger than the non-filtered ones for white and grey matter respectively. CSR between white and grey matter also seems to be improved by smoothing the diffusion-weighted data (panel C). As seen for SSR, the profiles of CSR do not shown a linear dependence with FWHM. Larger improvements are more noticeable for values of FWHM between 0.5 and 2.5 mm. For higher FWHM, values of SSR seem to increase with a lower rate, while CSR seems to stabilize. MK values extracted directly from non-smoothed DWIs shown a high percentage of negative values which ranged between 2% and 4% of the total number of neocortex voxels (panel D). These implausible values are practically eliminated by using Gaussian kernels with FWHM larger than 2 mm.

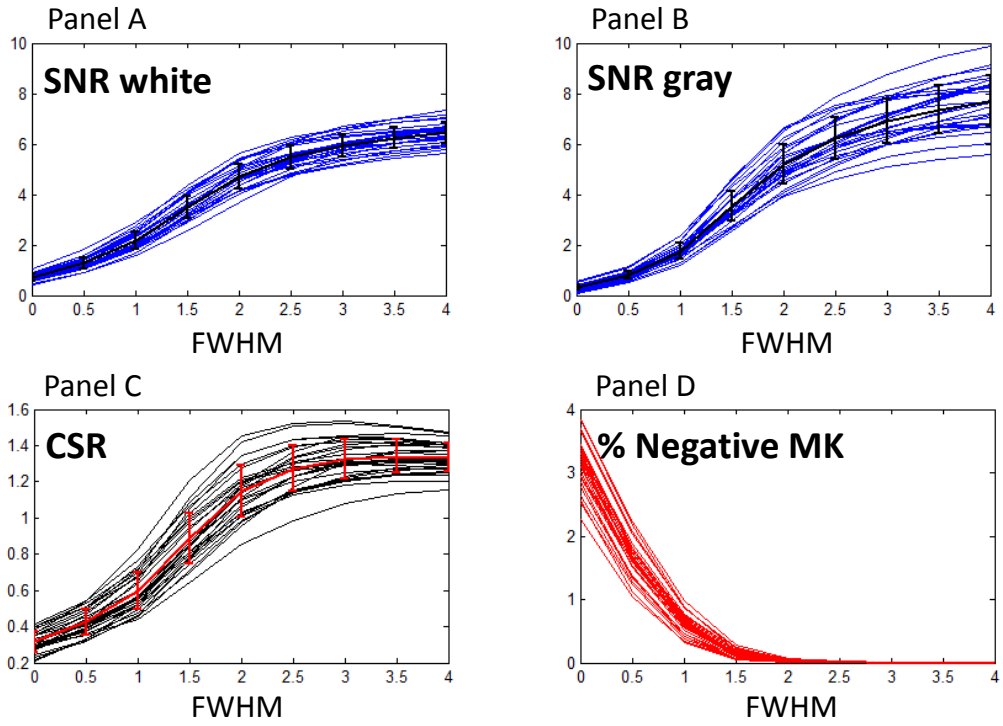



Figure 4.8 – MK image quality metrics in order to the values of FWHM. The black line and bars on panel A and B correspond to the SSR mean values and standard deviation across the 28 blue curves related to the 28 CamCAN subjects. The mean and standard deviation of CSR is presented in red on panel C, while each individual curve is plotted in black. Each individual percentage curve of negative MK is plotted in red on panel D.

Figure 4.9 shows the MK mean and standard deviation values for the two age groups – panel A and B show the results for white and grey matter for the subdivision 1 of the neocortex ROI, while C and D show the results for subdivision 2 (the results for the other subdivisions are shown in Appendix B). For all subdivisions of the neocortex and for both white and gray matter, mean MK seems to increase as a function of FWHM. Such increases are more pronounced for the lower FWHM values, and less significant for FWHM larger than 2.5mm. The p-values for the differences between the two age groups, which are plotted in grey on each panel, show significance age differences on the two most anterior and two most posterior neocortex subdivisions when Gaussian kernels with FWHM values between 2-3mm are applied.

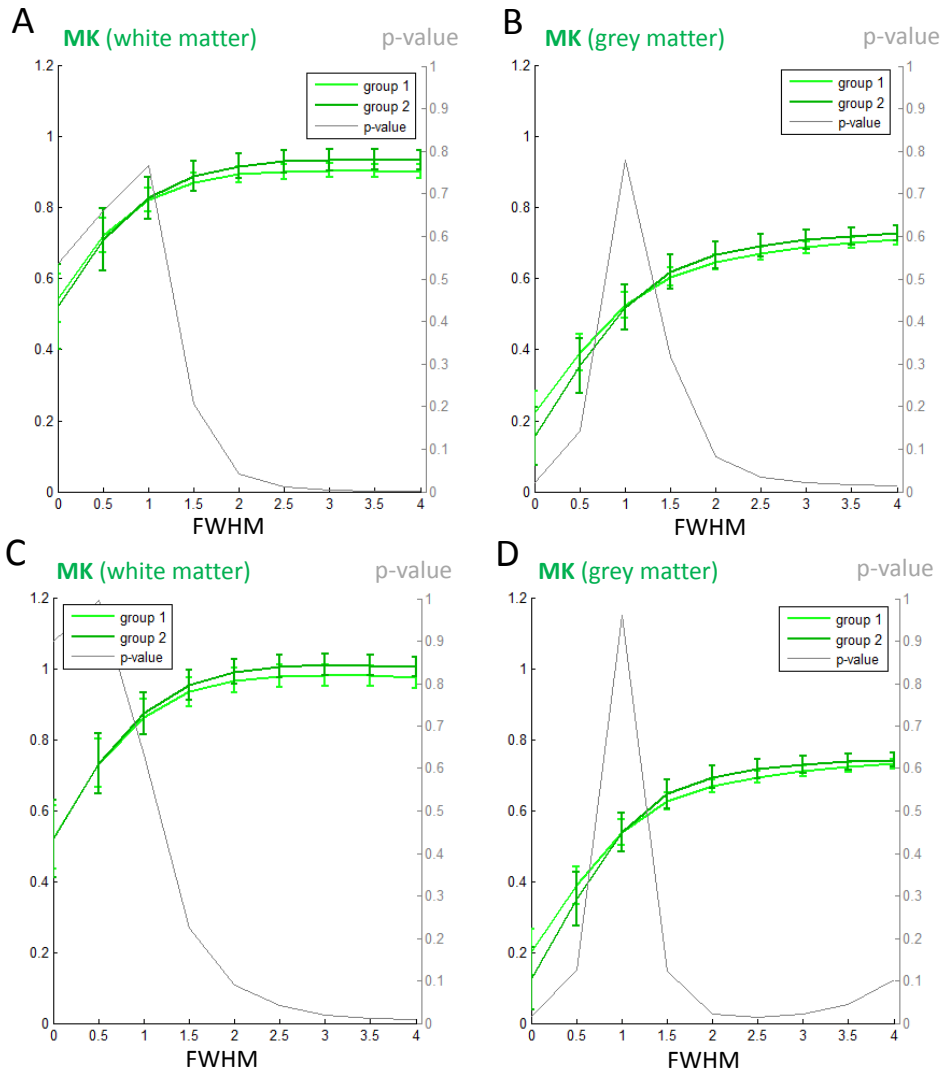



Figure 4.9 – Age differences of white and grey matter between the two ageing groups. Group 1 corresponds to 14 young adults (mean 26.2, sd 3.9) while group 2 corresponds to 14 middle age adults (mean 53.4, sd 2.0). The Panels corresponds to values of two prefrontal ROIs (subdivision 1 and 2 of the neocortex ROI shown on Figure 4.2 panel B).

### 4.3.3. Different Resolution and testing DKI reproducibility

Figure 4.10 shows the mean white and grey matter MK values for the 11 independent DWI datasets acquired on a single male subject of 25 year old (MK values for both white and grey matter are extracted from the ROIs shown in Figure 4.4). Panel A presents the results obtained for the DWIs acquired with resolution of 2 mm, while panel B shown the ones obtained for the DWIs acquired with resolution of 2.7 mm. Note that the SPM function *spm_smooth* did not to apply FWHM values that were inferior to 20% of the voxel resolution, and therefore on panel B values of MK are not shown for data acquired with resolution of 2.7mm for a kernel size of 0.5 mm (panel B).

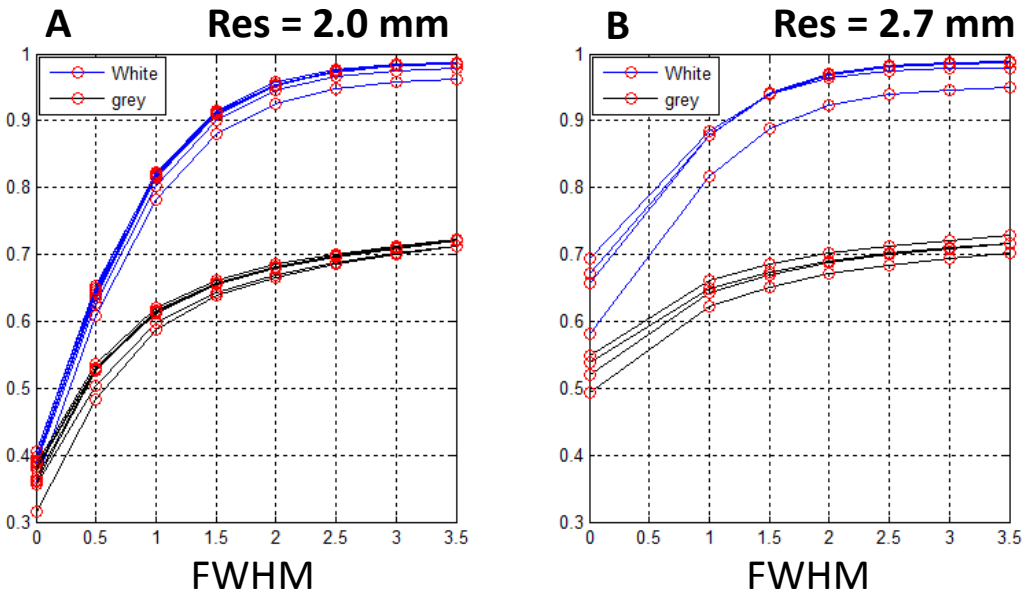



Figure 4.8 – Mean value of MK extracted for the 11 independent DWIs acquisitions for the male subject of 25 years old using a prefrontal ROI. Panel A shown the 7 repetitions of data acquired with resolution of 2.0mm, while panel B shown the 4 repetitions for the resolution of 2.7mm.

For the lower values of FWHM, Figure 4.10 shows that MK values are higher for data with resolution of 2.7 mm. However, for higher FWHM values the two resolutions seem to be similar. For example, Table 4.1 shows that values of MK mean and standard deviation are already similar for Gaussian kernels of 2.5mm. For this Gaussian kernel the within-subject coefficient of variation is presented in Table 4.2 which shows values inferior to 3% for both acquisition resolutions and for the two tissue types analysed.

Table 4.1 – MK mean ± standard deviation values computed for the white and grey prefrontal ROIs of DWIs acquired with resolution of 2.0 and 2.7 mm and filtered with a Gaussian kernel with FWHM of 2.5mm.

| Tissue type | Dataset 1 (2.0 mm) | Dataset 2 (2.7 mm) |
|---|---|---|
| **White matter** | 0.950 ± 0.010 | 0.969 ± 0.020 |
| **Gray matter** | 0.694 ± 0.006 | 0.700 ± 0.012 |

Table 4.2 – within-subject coefficient of variation values computed for the white and grey prefrontal ROIs of DWIs acquired with resolution of 2.0 and 2.7 mm and filtered with a Gaussian kernel with FWHM of 2.5mm.

| Tissue type | Dataset 1 (2.0 mm) | Dataset 2 (2.7 mm) |
|---|---|---|
| **White matter** | 1.1% | 2.1% |
| **Gray matter** | 0.9% | 1.7% |

The image quality metrics SSR, CSR and percentage of negative values of MK for the two acquired resolution are shown in Figure 4.11. With the exception of the MK negative percentage values for lower FWHM, the profiles for all image quality metrics with the increase of the Gaussian kernel seem to be similar between data acquired with resolution of 2.0 and 2.7 mm.

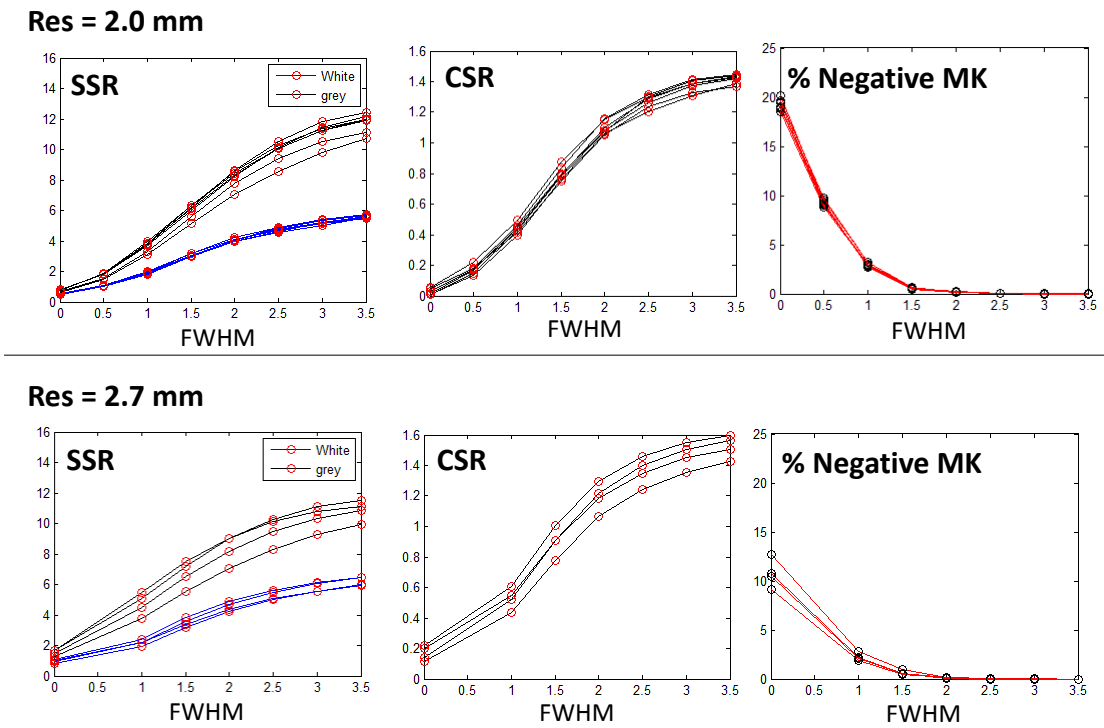



Figure 4.11 – SSR and CSR values of white and grey matter and percentage of negative MK for data acquired with two different resolutions. Superior panels correspond to the 7 repetitions acquired with a resolution of 2.0mm, while the inferior ones corresponds to the 4 repetitions acquired to a resolution of 2.7mm. Values are computed using the prefrontal ROI shown on Figure 4.4.

### 4.3.4. Downsampling effects

Figure 4.12 presents the image quality parameters for the 28 subjects' data as a function of the Gaussian kernel size for non-downsampled and downsample data. With the exception of slightly increase of CSR and SSR for all FWHM values and decreases of percentage of negative values of MK for the lower FWHM values, Figure 4.12 shows that image quality metrics of MK images extracted from downsampled data are similar to the ones extracted from data on its original resolution of 2 mm.

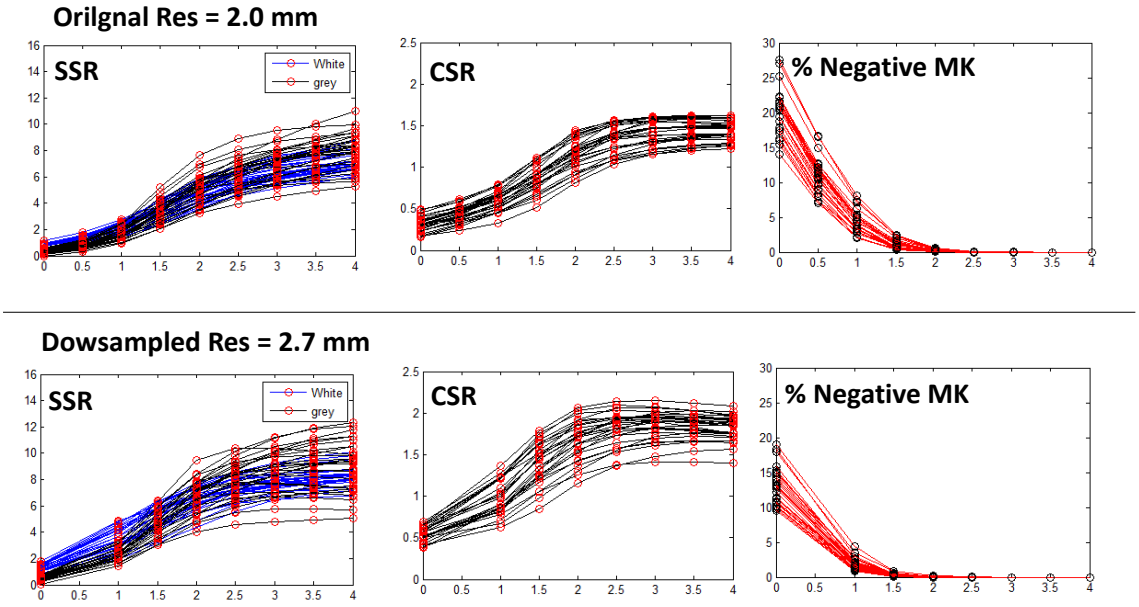



Figure 4.12 – SSR and CSR values of white and grey matter and percentage of negative MK extracted from two versions of the data of 28 CamCAN subjects: data with original resolution of 2.0mm and downsampled to a resolution of 2.7mm. Values are computed using the prefrontal ROI shown in Figure 4.5.

The prefrontal MK mean and standard deviation values for the two age groups are show in Figure 4.13. Panel A and C correspond to the white and grey matter values for data in its original resolution, while panel B and D show the white and grey matter values for the downsampled data. The p-values of the unpaired 2-sample t-test between the two age groups are plotted in grey in each panel. Both downsampled and non-downsampled data are more sensitive to age differences in white matter when smoothed with FWHMs larger than 2.0 mm.

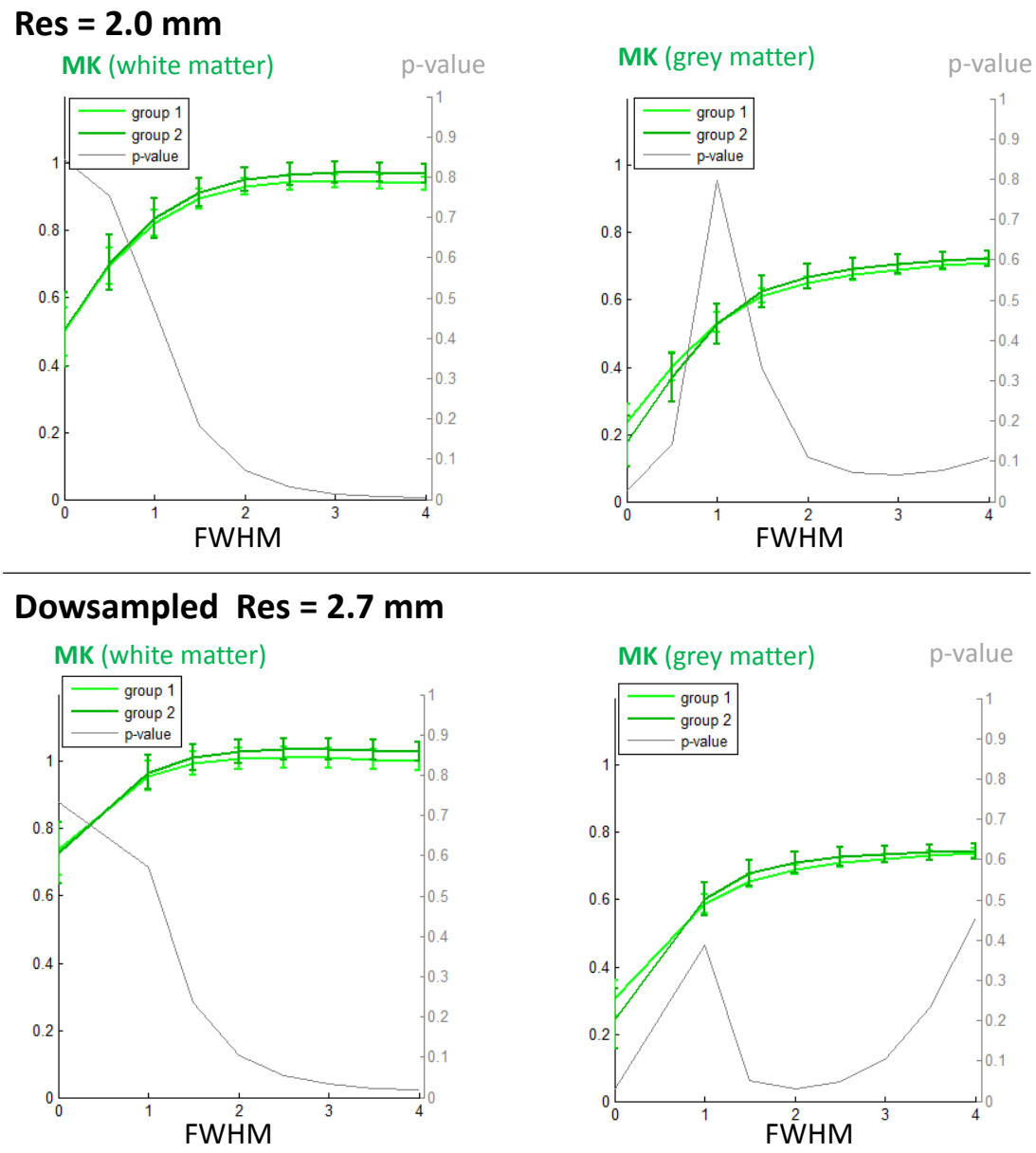



Figure 4.13 – Age differences of white and grey matter between the two ageing groups. Group 1 corresponds to 14 young adults (mean 26.2, sd 3.9) while group 2 corresponds to 14 middle age adults (mean 53.4, sd 2.0). Both downsampled and non-dowsample data are here analysed using the prefrontal ROI presented on Figure 4,5.

### 4.3.5. Applying direct linear fit on smoothed data

The age differences of mean MK between the two age groups of the neocortex ROI subdivisions 1 and 2 are shown in Figure 4.14, (the ageing differences analysis for the remaining ROI subdivisions are on Appendix C while the image quality metrics are on Appendix D). All image quality metrics show similar values to the ones presented in the previous sections; however, the direct linear fit shows to be more sensitive to age differences in particularly on white matter.

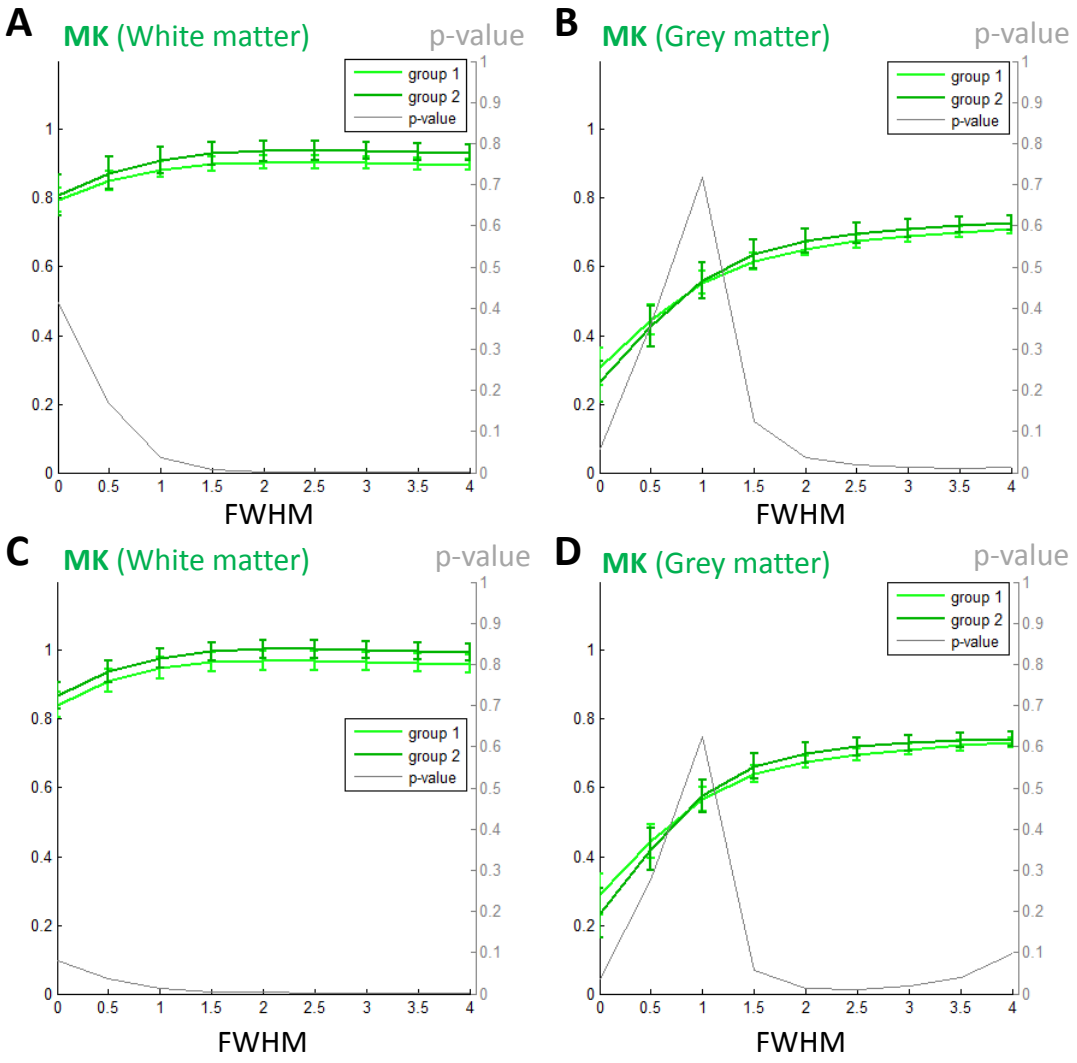



Figure 4.14 – Age differences of white and grey matter between the two ageing groups. Group 1 corresponds to 14 young adults (mean 26.2, sd 3.9) while group 2 corresponds to 14 middle age adults (mean 53.4, sd 2.0). Panel A and B corresponds to two prefrontal ROIs (subdivision 1 and 2 of the neocortex ROI plotted on Figure 4.2 panel B). MK data is extracted using DLS.

The improvements of applying different procedures for DKI can be assessed visually in Figure 4.15. For the image in panel A, kurtosis measures were extracted directly from the diffusion weighted images using the ordinary linear less squares (OLS) solution. Using this simplistic procedure, images will present a large number of voxels with implausible values, i.e. the dark regions on the image. By applying the direct linear least squares (DLS), the quality of the data is improved in particular on white matter regions as shown in panel B, and by introducing a Gaussian kernel in our pre-processing steps, the non-biological values along the cortex are removed (panel C). Finally, combining the Gaussian filter with an optimised parameter estimation procedure, results in images no longer show implausible values of kurtosis (panel D).

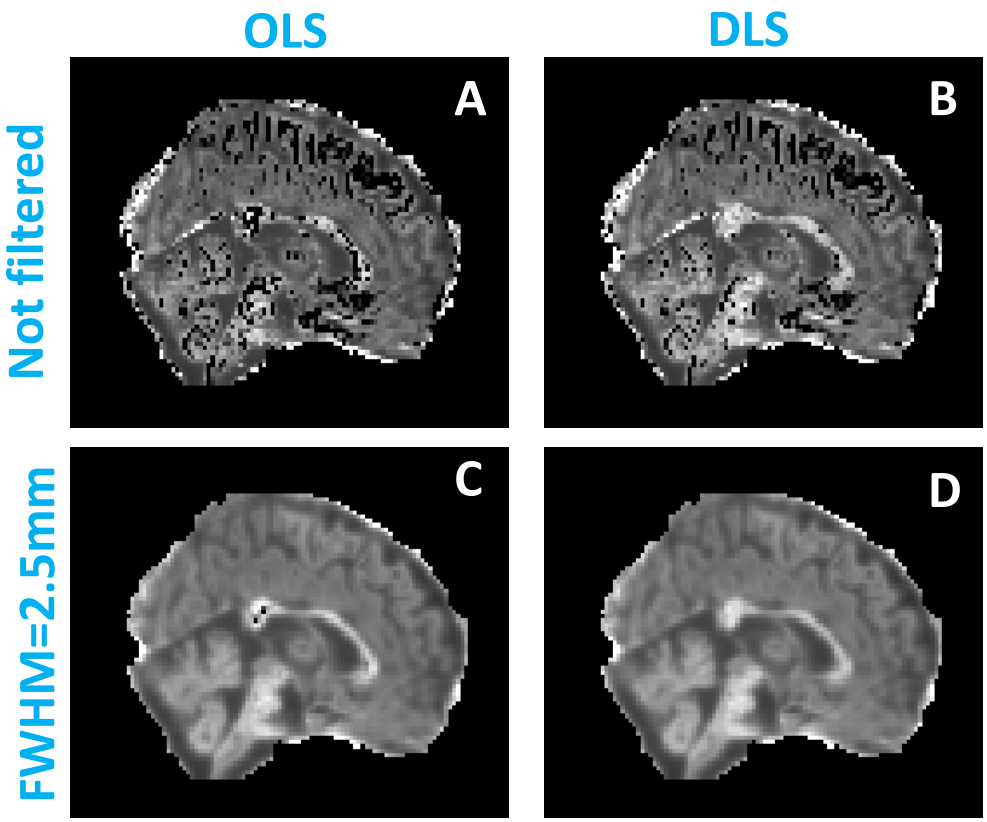



Figure 4.15 - MK maps of the same sagittal slice using different pre-processing and parameter extraction procedures.

## 4.4. Discussion

**Exclusion of motion correction algorithm on DKI**

The results presented in this chapter support the claim that a method for head motion correction based on registration algorithms is not adequate to correct DWI data since such algorithms cannot distinguish the inherent DWI contrast differences from head motion. For instance, despite brain volumes being simulated without any kind of artificial motion artefacts or even noise, the results in Figure 4.6 show that the registration-based method erroneously determines alignment parameters that reach values of the order of 5mm for the translation parameters. Interestingly, one can observe that the larger amplitudes of parameter estimates are related to the higher b-values of 2000 $s.mm^{-2}$ which is a consequence of larger diffusion-weighted contrast differences relative to the volume acquired with a b-value of 0 $s.mm^{-2}$. In addition, the variability seen across the 30 volumes acquired with different directions and the same b-value reflects the significant contrast differences observed for different gradient directions. Moreover, similar profiles

of parameter estimates were obtained on real data (Figure 4.7). This suggests that parameter estimates obtained by the rigid body transformation are essentially due to intrinsic contrast differences. Consequently, if these transformations were applied to the DWI volumes, one will be corrupting the images with additional translations and rotations instead of reducing the head motion effects.

**Optimal value of Gaussian kernel FWHM**

The results of this chapter show that improvements of image quality as function of the Gaussian kernel size do not have linear profiles (Figure 4.8). Larger improvement rate is observed on relatively small values of kernel FWHM (~1-2 mm). For example, by using a kernel with similar size than the image voxel resolution of 2 mm (which is assumed to have relative small consequences on image resolution), SSR on white and grey matter increases to values that are 6.82 and 17.03 times larger than the non-filtered data, however, when using a kernel of 3 mm which corresponds to a size 50% larger than the initial image resolution, SSR values for white and grey matter increases only to values 1.08 and 1.11 times larger comparatively to the data filtered with the Gaussian filter of FWHM = 2.5 mm. Consequently, the kernel size of 3mm can be considered to not provide a perfect trade-off between the increase of resolution and improvements in image quality. Based on the results of this chapter, a FWHM of 2.5 mm is suggested to be an optimal value to smoothing DWIs for DKI model fitting. In addition, this FWHM value shows the most significant age differences between the two age groups for all subdivisions of the neocortex ROI (Figure 4.9).

For the results in Figure 4.9, one can also observe that the mean values of both groups seem initially to increase which is a clear consequence of a gradual correction of implausible negative or underestimated values of MK. Interestingly, significant age changes ($p<0.05$) are visible when the MK values of the middle aged group are larger than the younger groups which was not the case of the non-smoothed data. For non-smoothed data, the middle aged group show values of MK that are inferior to the ones from the younger group, which can be a consequence of a larger tendency for older subjects to move during the acquisition of DWIs, which would result in a larger percentage of underestimated values of MK.

**Data acquired with resolution of 2.7 mm and DKI reproducibility**

For data not smoothed with a Gaussian filter, MK values acquired with a resolution of 2.7 mm seem to be larger than the ones obtained with a resolution of 2.0 mm (Figure 4.10) which reflects the inferior percentage of MK negative values when data is acquired with larger resolution (Figure 4.11). However, values of SSR for white and grey matter and CSR between these two types of tissue are not considerably larger for data recorded with resolution of 2.7 mm (Figure 4.11), suggesting that a large amount of non-negative underestimates of MK are still present. Moreover, after applying a Gaussian kernel with FWHM larger than 2 mm, MK characteristics seem to be similar from both resolutions (Table 4.1). This suggests that DWIs acquired with larger resolution do not present significant advantages for DKI apart from a faster acquisition time.

In contrast to what was assumed in previous studies (e.g. Tabesh, Jensen, Ardekani, & Helpern, 2011), these results suggest that data acquired with larger resolution do not necessary require the use of kernel size larger than 2.5 mm. In particular, the results shown in this chapter suggest that FWHM of 2.5 mm can be also considered an optimal value for smoothing data acquired with resolution of 2.7mm.

For the optimal value of FWHM = 2.5 mm, within-subject coefficient of variation shows that mean MK have errors inferior than 3% for both acquisition resolutions and both tissue types analysed which can be considered as being a sufficient low error (Table 4.2).

**Downsampling Effects**

By analysing SSR, CSR and percentage of negative MK values from downsampled data (Figure 4.12), similar profiles on the dependency of the Gaussian smoothing kernel size can be observed relative to the non–downsampled data mentioned above, and thus the optimal value of 2.5mm can be similarly considered. Despite downsampled data providing MK maps with slight improvements on the image quality metrics, these did not result in improved ability to detect age differences (Figure 4.13). This can be a consequence of the reduction of image resolution and the increase of the partial volume effects on downsampled data. Moreover, downsampling data without applying any smoothing seems to be insufficient to reduce the impacts of noise and artefacts, thus concluding that downsampling does not add advantages to the pre-processing of CamCAN diffusion MRI data.

**The direct linear least squares fit on smoothed data**

MK data extracted for the DLS did not show significant differences on the image quality metrics when compared to OLS; however DLS resulted in data that is more sensitive to age differences (Figure 4.14). Despite not avoiding all implausible estimates without smoothing of DWIs (see panel B of Figure 4.15), this fitting method provides smaller amounts of underestimated MK values as suggested by the global increase of mean MK on non-smoothed data (Figure 4.14). In addition, improvements of punctual voxels are visible in particular in white matter voxels which are related to larger amounts of OLS underestimates and negative values of MK since these voxels correspond to lower values of diffusion (Figure 3.7 of chapter 3). While the use of DLS provides greater improvements in white matter, the Gaussian smoothing visually shows a larger improvement for grey matter (Figure 4.15, panel C), and therefore the use of DLS on data smoothed with a Gaussian kernel of 2.5 mm can be seen as an optimal combination to extract MK images with better accuracy (Figure 4.15, panel D).

**Suggested DKI procedure for the Cam-CAN study**

As the registration based method for correcting head motion artefacts, the acquisition of images with larger voxels, and the downsampling algorithms did not shown to add accuracy to DKI, the only pre-processing steps that will be added on the preliminary DKI procedure for Cam-CAN is the Gaussian kernel with optimized FWHM of 2.5 mm. Therefore, as results of the analysis performed in this chapter and chapter 3, the suggested procedure to extract the diffusion and kurtosis tensors will consist on three steps:

1. Diffusion-weighted brain voxels are extracted from DWIs volumes using FSL brain extraction tool (Smith, 2002);

2. Brain DWIs volumes are smoothing using a 3D Gaussian kernel with FWHM of 2.5m;

3. Diffusion and kurtosis tensors are extracted using an ordinary linear least squares approach. In addition direct linear squares will be used to validate MK and MD estimates.

## 4.5. Chapter Summary

In this chapter some important aspects on pre-processing DWIs for DKI were discussed. Firstly it was experimentally demonstrated that head motion correction methods based on image correlation are not adequate to correct DWIs. Instead, this method corrupts DWIs with translation and rotation parameters that are uncorrelated with subject motion. Secondly, the FWHM Gaussian kernel optimization study shows that a kernel size of 2.5 mm provides an acceptable trade-off between image quality, resolution and it was shown to result in more significant age differences in the mean MK values. Despite this kernel size already being used in previous studies of DKI (e.g. Falangola et al., 2008; Jensen et al., 2011), in this chapter this kernel size is experimentally justified for the first time. Moreover, some studies used larger kernel sizes for data acquired with larger voxel sizes (e.g. Tabesh, Jensen, Ardekani, & Helpern, 2011), while the results of this chapter have shown that this might not be the optimal procedure. Despite downsampled data provides higher values of SSR and CSR, this step did not prove itself valuable in detecting age difference when compared to non downsampled data, and thus, it is considered to not add advantages on DKI for the Cam-CAN study. Finally it was shown that combining more accurate parameter estimations with a proper pre-processing method (i.e. DLS with a Gaussian smoothing with kernel size of 2.5mm) can provide the most robust procedure for DKI analysis. From the results of this chapter a three step procedure for DKI was suggested.

# Chapter 5

# Imaging Ageing using DKI

## 5.1. Introduction

Across the human stages of development and ageing, brain shows structural and morphological changes (Brody, 1992). For example, early MRI studies using conventional structural contrast have shown that cortical grey matter volume increases on early childhood and decreases with age after puberty (Giedd et al., 1999). White matter density has also been shown to increase during brain maturation, however it was only shown to decrease from early forties (Bartzokis et al., 2001). Brain changes can also differ in different anatomical regions. For example, grey matter density of parietal cortices seems to increase during childhood earlier than in the frontal cortices (Sowell, Thompson, Holmes, Jernigan, & Toga, 1999). In addition, visual, auditory and limbic cortices showed different age changing profiles than the frontal and parietal neocortices (Sowell et al., 2003).

Such volumetric changes are believed to be a consequence of a diversity of cellular maturation and degenerative events. Although conventional MRI contrasts are only sensitive to changes of macrostructures within the millimetre-sized voxel of T1, T2, or proton density-weighted images (Moseley, 2002), post mortem studies have shown that brain changes can be related to microscopic events as myelination or demyelination of axonal fibres, synaptic pruning, axonal lost, or cell shrinkage (e.g. Benes, Turtle, Khan, & Farol, 1994). Preliminarily studies using DTI have identified changes in diffusion parameters that can be directly associated with changes of microstructure. For instance, FA shows rapid increases of white matter anisotropy in neonatal and young childhood during brain development, which can be associated to myelination, and decreases from

early twenties onwards, which can be associated directly to demyelination processes (Moseley, 2002; Sullivan & Pfefferbaum, 2006). Mean diffusion (MD) is also sensitive to age changes. In white matter, diffusion seems to decrease until puberty and then it increases with age (Moseley, 2002). Such changes can be related to the same maturation and degenerations suggested by FA measures, i.e., the increase of myelin on axonal barriers during brain maturation results in larger anisotropy and a more restricted diffusion environment, while demyelination result in decreases of anisotropy and the medium less restricted to water diffusion. MD also shows ageing changes across lifespan on grey matter; however these cannot be related to microstructural changes measured by FA. As grey matter does not show anisotropy, FA is not sensitive to any maturation or ageing process in this type of tissue. In addition, the tissue anisotropy index given by FA is corrupted by tissue inhomogeneity from crossing fibres, to partial volume effects (Sullivan & Pfefferbaum, 2006), and to noise effects that bias FA as shown in chapter 3 of this thesis.

It has long been appreciated that diffusion-weighted images contain more information than the extracted by DTI (Falangola et al., 2008). In particular, the kurtosis tensor measured by DKI has been shown to provide estimates that can be regarded as an index of barrier complexity (Helpern et al., 2011). This index can be considered to be a more general index than FA since it does not depend on the degree of anisotropy of different brain regions (Falangola et al., 2008). DKI has already been used in a preliminary ageing study, where MK showed significant age correlations for both grey and white matter in the developing and ageing brain (Falangola et al., 2008). However, this previous study had noticeable limitations: reduced number of participants (24 subjects from 13 to 85), it analysed only a prefrontal region of interest, and the relationship between the differences found in the brain and the subject's cognitive abilities was not considered.

The aim of this chapter is to report preliminary ageing results from DKI on 103 subjects recruited by Cam-CAN. Despite being a preliminarily study for Cam-CAN, this chapter will already present novel results on ageing since it provides results based on a larger sample of subjects than the previous ageing study using DKI, which will allow for a more detailed description of the different stages of healthy ageing. This chapter also describes results from directional kurtosis, which were never shown in previous ageing studies. In addition to the analysis of mean values and histogram peak of a prefrontal region of interest, this chapter will also show a preliminary analysis of the FA, MK, and

MD ageing changes across different brain regions using the novel tract based spatial statistics (TBSS) technique suggested by (Smith et al., 2006).

## 5.2. Methods

### Subjects

The Cam-CAN subjects analysed in this preliminary study were selected from the 176 subjects recruited before March 2012. The age distribution for these subjects is shown in Figure 5.1 panel A. Since there is a large proportion of subjects with age above 48, some of the older subjects are randomly excluded on this preliminary study to ensure a more uniform age distribution. This random exclusion is performed on subjects aged over 48 so that each age is only represented by a maximum of two subjects. Additionally, eight subjects (ages 18, 22, 22, 23, 26, 42, 71, and 85) were excluded for visually presenting DWIs largely corrupted by motion artefacts. The age distribution of the 103 selected subjects is shown in Figure 5.1 panel B.

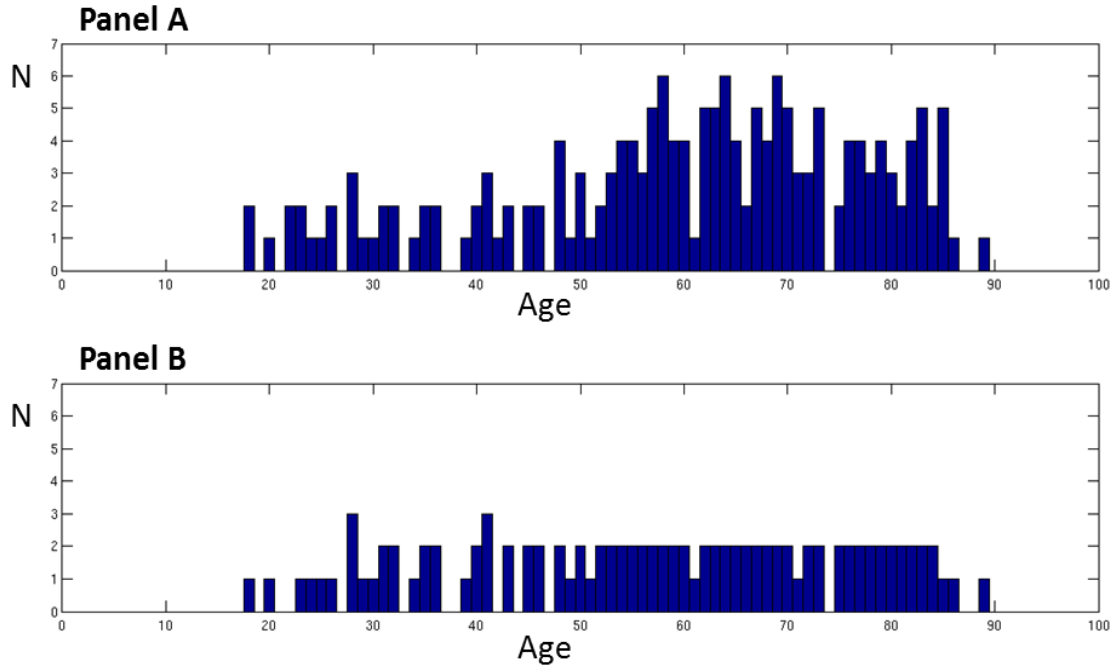



Figure 5.1 – Age distributions for two groups: panel A corresponds to the ages of all Cam-CAN's subjects acquired before March 2012; while panel B shows the ages of the 103 subjects selected for the preliminary Cam-CAN study described in this chapter.

**Diffusion Kurtosis Imaging**

Diffusion kurtosis analysis was performed on the 63 volumes of DWIs[3] for each subject of the selected group of 103 Cam-CAN subjects using the DKI procedures selected accordingly to the results of chapter 3 and 4, i.e.: (1) the brain diffusion-weighted voxels are selected using FSL brain extraction tool; (2) the diffusion-weighted brain intensities are smoothed using a 3D Gaussian kernel with FWHM of 2.5mm (SPM function *spm_smooth*); (3) diffusion and kurtosis tensors are extracted using an ordinary linear least squares approach (as described on section 2.4.1); (4) from the estimated diffusion and kurtosis tensors, the mean diffusion (MD), the fractional anisotropy (FA), the axial diffusivity ($\lambda_{\parallel}$), the radial diffusivity ($\lambda_{\perp}$), the mean kurtosis (MK), the axial kurtosis ($k_{//}$), the radial kurtosis ($k_{\perp}$) and kurtosis fractional anisotropy ($FA_k$) are extracted using the procedures described on section 2.3.2 and 2.4.2; (5) finally the direct linear least squares is used to validate the results from MK and MD once it shown to provide better performances by extracting this metrics.

**Diffusion Tensor Imaging**

To provide benchmark information relative to the diffusion kurtosis analysis, ageing changes over the 103 subjects are also analysed using a diffusion tensor model. In particular, since it is less sensitive to noise than DKI, DTI can be used to extract values of MD, FA, $\lambda_{\parallel}$, and $\lambda_{\perp}$ without requiring the Gaussian smoothing of DWIs to decrease the impacts of Rician noise. Therefore, the invariant maps extracted from DTI can be useful to validate the changes detected by MD, FA, $\lambda_{\parallel}$, and $\lambda_{\perp}$ after the blurring effects induced by the Gaussian filter on the DKI analysis. Since Cam-CAN is based on larger b-values than normally used for DTI, the diffusion tensor is estimated on the brain extracted DWIs using a non-linear fitting approach[4] (see section 2.3.1 or section 3.4) to reduce the bias introduced by Rectified noise as suggested by Jones & Basser (2004).

[3] DWIs were acquired on a 3T Siemens Trio with a twice-refocused-spin-echo, 30 diffusion gradient directions for each b-values 1000 and 2000s.mm-2 and three images acquired using b-value 0, TR=9100ms, TE=104ms, voxel size=2x2x2mm3, FOV=192x192mm2, 66 axial slices, number of averages=1.

4 The results for the DTI linear fit will be presented on Appendix E to highlight the importance of using non-linear approaches on DTI to reduce the artefact impacts when ignoring the non-Gaussianity of the displacement of water molecules on neural tissue.

**Region of interest analysis**

As ageing changes occur spatially with different profiles (e.g. Sowell et al., 2003), it is important to study separately the ageing changes on different brain regions. In the preliminary study reported on this chapter, a prefrontal brain region of interest (ROI) is selected since this region had shown significant age changes in previous DTI studies, (e.g. Sullivan & Pfefferbaum, 2006). In addition, this pre-frontal brain region was also used in the DKI ageing study published by Falangola et al. (2008), and therefore by studying the prefrontal brain, the results obtained by DKI on this chapter can be qualitatively compared to this previous study.

The prefrontal ROI is first defined in the Montreal Neurological Institute (MNI) template by selecting the region extended from the most anterior point containing brain tissue to the dorsal border of the genu of the corpus callosum on 10 selected axial slices, adding up to a total thickness of 1 cm (this ROI was already used for Chapter 4, and therefore it is represented in Figure 4.5). Then the prefrontal ROI is transformed to each subject's DWI space using the procedures described on section 4.2.2. To analyse separately white and grey matter, the maps for these two types of tissue are estimated by segmenting each subject's structural MPRAGE images using the SPM functions *spm_preproc* and *spm_pre2n*, and then registered to DWI space using the functions *spm_coreg* and *spm_reslice* (for more detail see section 4.2.2).

To study the volume age changes of white matter (WM), grey mater (GM), and cerebrospinal fluid (CSF), the percentage of these types of tissue on the prefrontal brain will be longitudinally analysed. The classification between different tissue types is based on the maps obtained by the SPM segmentation function (selecting a classification probability threshold of 70%). For each Cam-CAN subject, the WM and GM maps were then used to select the white and grey matter voxels to compute the frontal brain mean values of the metrics extracted by DKI and DTI.

**Analysis based on MK histograms**

Similarly to Falangola et al. (2008), the MK histogram profiles of the prefrontal ROI are analysed. The values of MK in the histograms ranged from -1 to 3 with bin sizes of 0.031. The values of peak position of the given histograms are also analysed. In this way, based on the assumption than MK histograms have peaks corresponding to grey matter (in the range 0.5-1) and white matter (range 1-2), the peak positions are estimated by first detecting the local histogram maximum on each tissue's typical interval and then, by

computing the weighted average position of the voxels contained within the upper 40% of the detected maximum value on the near neighbourhood (i.e. three bins for both left and right side of each peak). In addition to the analysis of the histograms from the entire prefrontal ROI, the histograms for the prefrontal segmented WM, GM, and CSF maps were also analysed. The peaks from the individual WM and GM histograms were then extracted by the weighted average position of the voxels contained in the upper 40% of the global maximum value found in each individual histogram over all histograms bins.

**Models to classify longitudinal age changes**

To classify the age longitudinal age changes of metrics based on mean values of ROIs and histogram peaks, step-down model selection is preformed to test cubic, quadratic, and linear age effects , i.e. if the cubic age effect is not significant relative to the quadratic model, the cubic term is removed and the quadratic term is tested against the linear one, and so on (Shaw et al., 2008). To avoid type 1 errors (rejection of a true null hypothesis) the rejection of a term on the multiple regression model is only considered if the statistical p-value is inferior to 0.001(Sowell, Thompson, & Toga, 2004).

**Analysis based on Tract Based Spatial Statistics**

In alternative to analysis based on the mean values and histograms of the prefrontal ROI, metrics extracted from DKI and DTI can be analysed using voxel-based analysis (Jones & Cercignani, 2010). This technique allows the simultaneous analysis of all brain voxels with the objective to spatially localise diffusion-related changes (Smith et al., 2006). For example, following similar approaches to voxel-based morphometry (VBM) techniques, maps of any invariant metric extracted from DTI or DKI can be registered to a standard space and then statistics are performed voxelwise to find the regions that correlate with a covariate of interest such as age. However, VBM based techniques have some pitfalls when applied to diffusion parameters. For instance, residual misalignments can decrease the analysis sensitivity or even suggest erratic results (Simon et al., 2005). Aiming to overcome the VBM limitations, tract based spatial statistics (TBSS) was proposed by Smith et al. (2006). According to Smith et al. (2006), the procedures for TBSS consist of four steps which are described on the following paragraphs.

The first step of TBSS is to align all subjects' FA maps to a common registration target using non-linear registration. To remove likely FA outliers, the FA maps are slightly eroded and the end slices are set to zero using the FSL script *tbss_1_preproc*

(http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS). The common target for the non-linear registration is selected as the subject identified as the "most representative" FA volume, i.e. the single individual target that minimises the amount of warping required for all other subjects to be aligned to it (Smith et al., 2006). This target is identified by registering all individual FA maps to all other FA volumes, and then the individual target is chosen as the individual FA with the minimum mean transformation required to realign all other subjects' FA maps. The process to identifying the registration target is performed using the FSL script *tbss_2_reg*. For the analysis performed on this chapter, it is used the FA maps extracted from non-linear DTI once this approach provides the FA metrics without being corrupted by the blurring effects of a 3D smoothing.

TBSS second steps consist on creating a map containing all white matter tracts that were common to all subjects which is named the mean FA skeleton. For convenience of the interpretation and display of the results, all data is converted to the MNI space. For this, the target selected on the previously steps is coregisted to the MNI space and the resulted transformation is combined to all transformation matrix related to each registed individual FA volume. This is performed using the FSL script *tbss_3_postreg* which gives as outputs all subjects' FA volumes transformed to the MNI space, a mean FA volume, and the mean FA skeleton. Regarding the mean FA skeleton, such is computed from the mean FA volume, by first estimate the local surface perpendicular to the tract directions and then by computing the centre of the tract using non-maximum-suppression (for more details on this see Smith et al., 2006). Having an estimate of the FA skeleton, this is thresholded in order to restrict the further analysis to voxels corresponding only to white matter voxels that were successfully aligned across subjects. The threshold can be set as an input on the FSL script *tbss_4_prestats*. For the mean FA skeleton generated from the 103 Cam-CAN's subjects, the threshold is adjusted to a value of 0.35 once it shown to be an adequate value to select typical white matter tracks as shown in Figure 5.2.

Automatically, the script *tbss_4_prestats* also performs the third step of TBSS' procedure which consists in creating an FA skeleton for each individual subject. These are based on the voxel positions of the thresholded mean FA skeleton, however now the intensity of each voxel of the skeleton is set as the subject's maximum FA perpendicular to the tract. This procedure aims to account for the misalignments between subjects after the initial nonlinear registration. TBSS allows also the projection of other invariant metrics to the skeleton. For this, one can use the script *tbss_non_FA* given as input a

folder containing all subject volumes for a specific metric. *tbss_non_FA* will apply the nonlinear registration (determined on TBSS' first and second steps) to the non-FA data, which will be followed by the projection of the maximum on the non-FA values on each individual skeleton (for more detail on this, see TBSS user guide at http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS, September 2012). On this chapter *tbss_non_FA* is applied to the preliminary study of age changes on MD and MK. The directly linear least squares fit is selected to extract this invariant metrics once this method shown to provide the more accurate DKI estimates of MD and MK on Chapter 3 and 4).

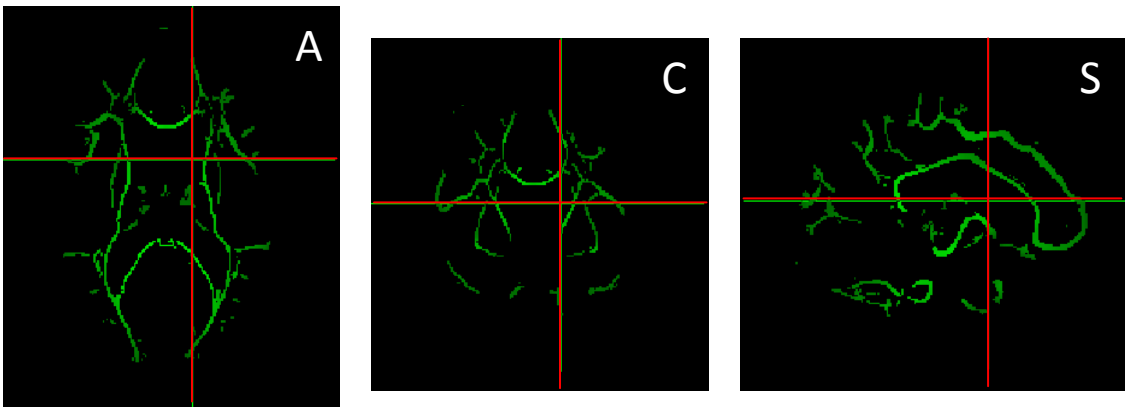



Figure 5.2 – FA 'skeleton' estimated from the 103 Cam-CAN subjects' and thresholded by a value of 0.35. The FA 'skeleton' is presented on an axial (A), coronal (C), and sagittal (S) views.

Having the individual skeletons, TBSS can finally be followed with the voxelwise statistics across subjects. For this, it is used the FSL's *randomise* tool to study the significance of partial regression coefficients of linear models using a permutation test (Anderson & Robinson, 2001). In this chapter, using the *randomise* software with 10000 permutations, it will be tested the brain regions that shown significance linear increases and decreases with age on FA extract from non-linear DTI, and on MD and MK maps extracts from DKI direct linear fit (significances were computed based on Threshold-Free Cluster Enhancement p-values (Smith & Nichols, 2009)).

## 5.3. Results

### 5.3.1. Analysis based on the prefrontal region of interest

**Percentage of WM, GM, and CSF voxels**

Figure 5.3 shows the percentage of WM, GM, and CSF voxels of the analysed prefrontal ROI. While the percentage of WM voxels seems to increase until the middle of the fifth decade, GM seems to decrease from the early twenties. The GM decreasing rate seems to accentuate for the last lifespan decades (50-90). Although CSF initially seems to remain stable, its volume significantly increases for decades from the forties.

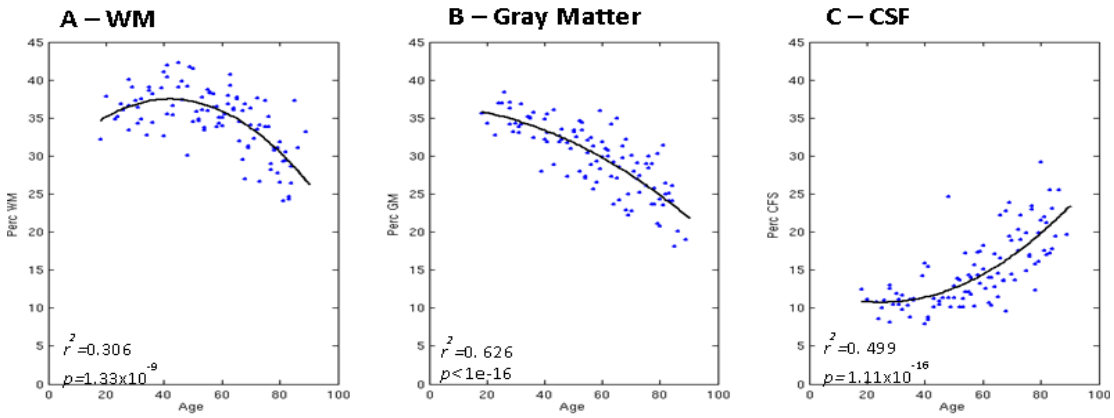



Figure 5.3 – Longitudinal changes on the percentages of voxels for WM, GM, and CSF on the prefrontal region of interest.

**DKI Ordinary linear least squares**

The longitudinal changes on MK, MD, and FA are shown in Figure 5.4 for white matter (panel A) and grey matter (panel B)[5]. FA shows a decrease of prefrontal brain white matter anisotropy since the early twenties. White matter MD remains stable during the first analysed decades, however it significantly increases after the fifties. MK shows a different profile when compared to white matter FA and MD. It increases until early forties and then decreases until the oldest ages. For grey matter, FA shows a constant low anisotropy while MD linearly increases. As for WM, MK shows a different profile from FA and MD on prefrontal brain GM. It shows a quadratic profile, increasing to the middle of the sixth decade and then it starts to decrease.

[5] $r^2$ statistic and p-values shown on the panels corresponds to the significance of the given step-down regression models against the null hypotheses that age does not have any influence the specific DKI or DTI analyses (p-values and $r^2$ statistic values are not shown for constant models once they already correspond to an age independent model).

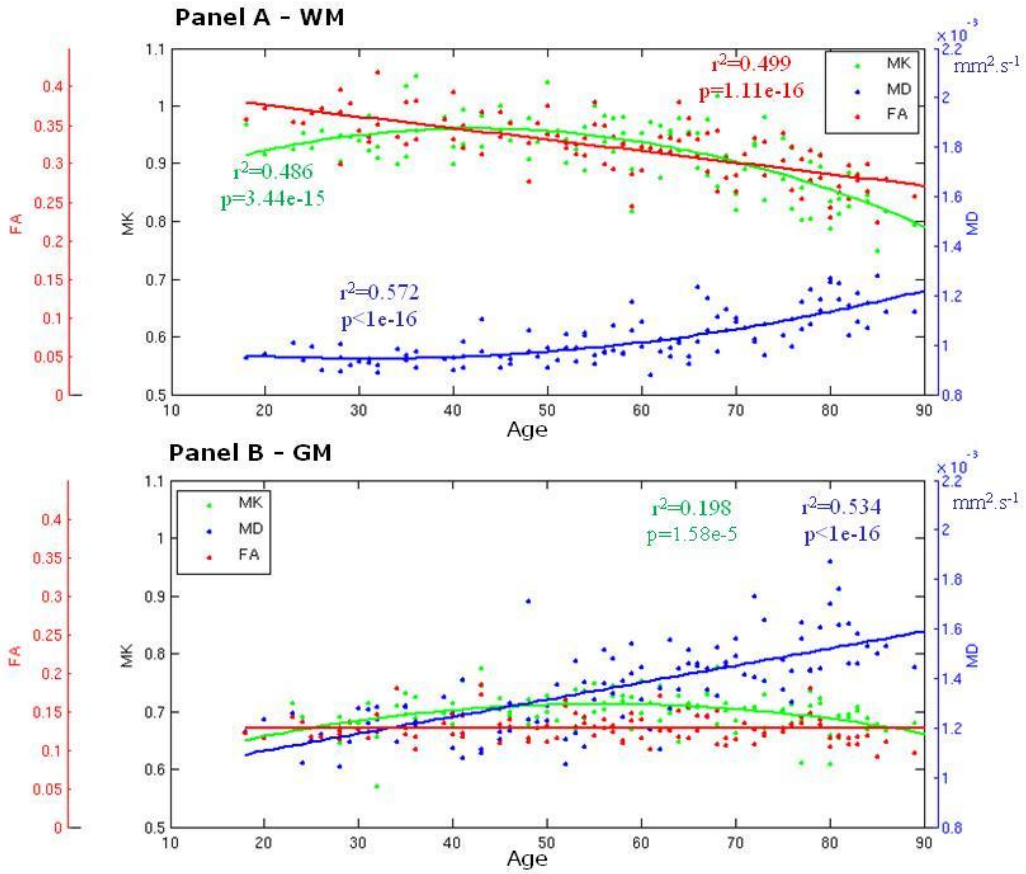



Figure 5.4 – Longitudinal changes on MK (green), MD (blue) and FA (red) for white matter (panel A) and grey matter (panel B).

Figure 5.5 shows the FAk changes for white (panel A) and grey (panel B) matter, which are similar in profile to the FA extracted from the diffusion tensor; however on panel A, FAk shows smaller values of $r^2$ statistic and larger p-value.

The diffusion and kurtosis axial and radial metrics are presented in Figure 5.6. Panel A and B represents the axial and radial metrics for white matter, where $\lambda_{\parallel}$ shows larger values than values $\lambda_{\perp}$, and $k_{||}$ shows lower values than $k_{\perp}$. Increases of diffusion and decreases of kurtosis seem to occur primarily on the axial directions (panel A) and not for radial directions (panel B). For grey matter (panel C and D), profiles are similar between $\lambda_{\parallel}$ and $\lambda_{\perp}$ and similar between $k_{\parallel}$ and $k_{\perp}$. Despite these profile similarities, $\lambda_{\parallel}$ seems to have slightly larger values than $\lambda_{\perp}$, while $k_{\parallel}$ values are slightly smaller than $k_{\perp}$. In addition, comparing the profiles on panels C and D with panel B on Figure 5.4, one can see that axial and radial profiles for both diffusion and kurtosis are similar to the grey matter profiles of MD and MK.

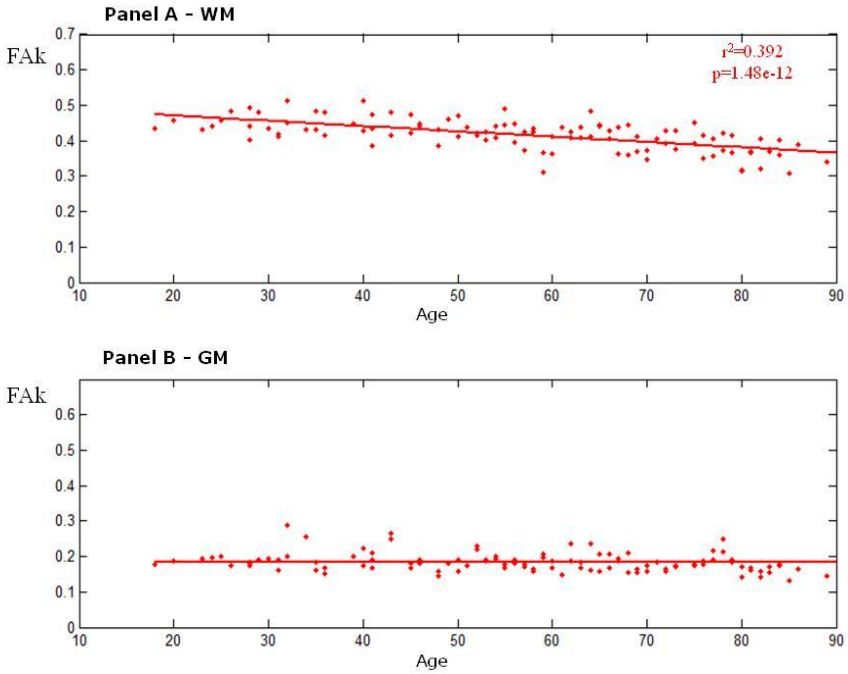



Figure 5.5 – FAk longitudinal changes for white matter (panel A) and grey matter (panel B).

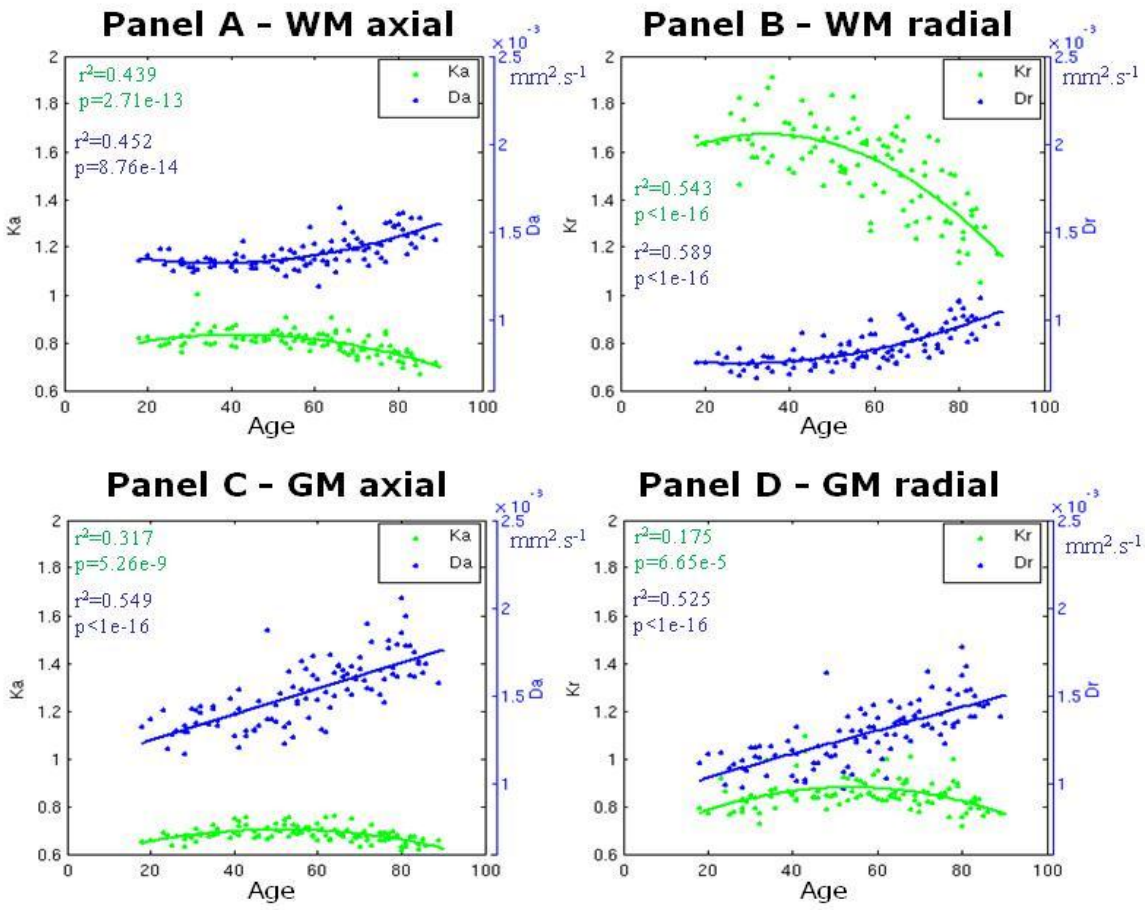



Figure 5.6 – Diffusion (blue) and kurtosis (green) axial and radial changes with age. Panel A and B corresponds to the values of axial and radial diffusivity/kurtosis for white matter, while Panel C and D the axial and radial values from grey matter.

**DKI direct linear least squares**

The longitudinal changes on the MK and MD extracted from DKI direct linear least squares method is shown in Figure 5.7. For both white and grey matter (panel A and B), MK and MD profiles are identical to the profiles of MK and MD extracted using the ordinary linear least squares (Figure 5.4). However, $r^2$ statics values are larger and p-values are smaller comparatively to the ones presented in Figure 5.4.

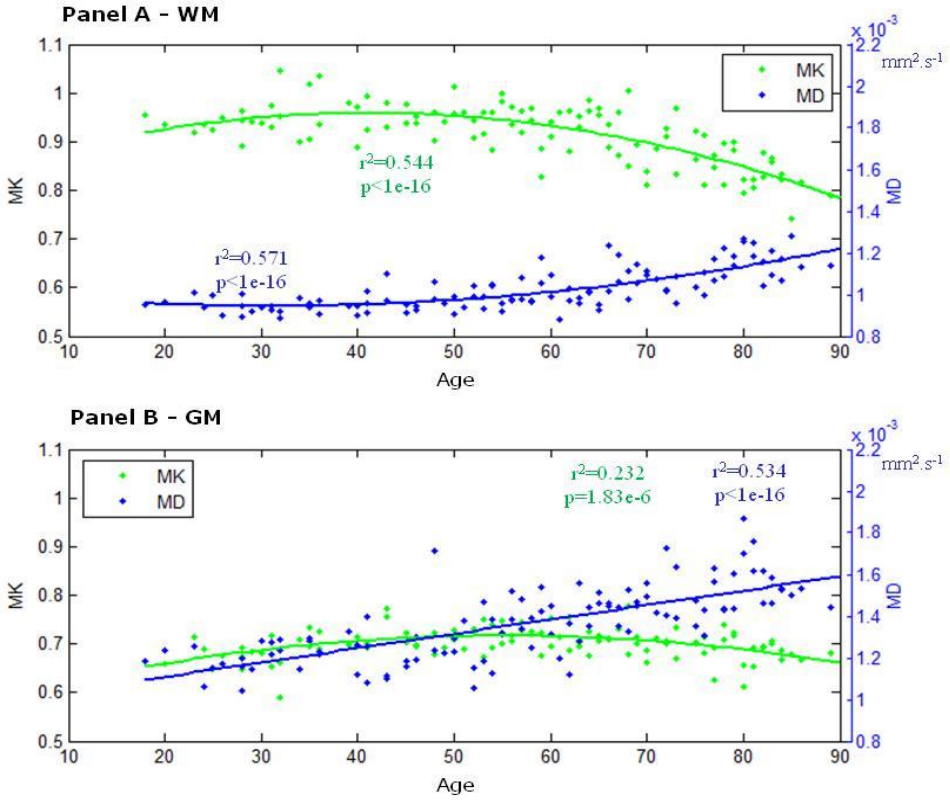



Figure 5.7 – Longitudinal changes on MK (green) and MD (blue) for white matter (panel A) and grey matter (panel B). Metrics were extracted using DKI direct linear least squares solution.

**DTI non-linear least squares**

The results for the FA and MD values extracted from non-smoothed DWIs using non-linear DTI are shown in Figure 5.8. In general FA values are larger than the ones extracted from ordinarily linear DKI (and shown on Figure 5.4), while MD values are lower. However, identical profiles for both MD and FA can be observed between these different techniques. Similar observations can be taken for the white and grey matter longitudinal profiles of $\lambda_{||}$and $\lambda_{\perp}$extracted from non-linear DTI (Figure 5.9) when compared to the directional diffusion results of DKI presented on Figure 5.6.

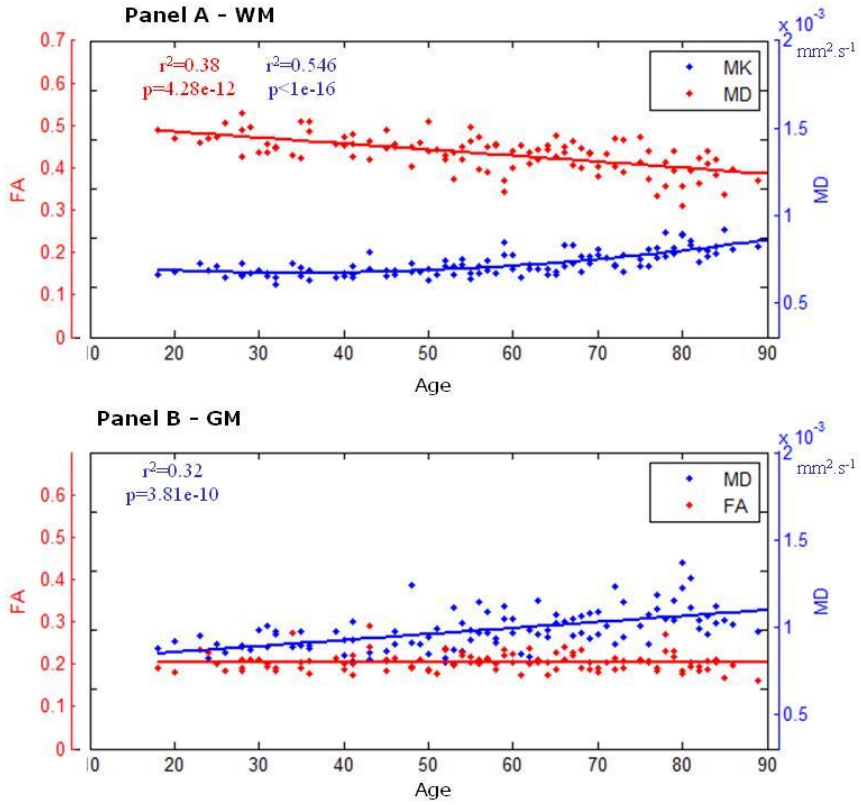



Figure 5.8 – Longitudinal changes on MD (blue) and FA (red) for white matter (panel A) and grey matter (panel B). Metrics were extracted using DTI non-linear least squares solution.

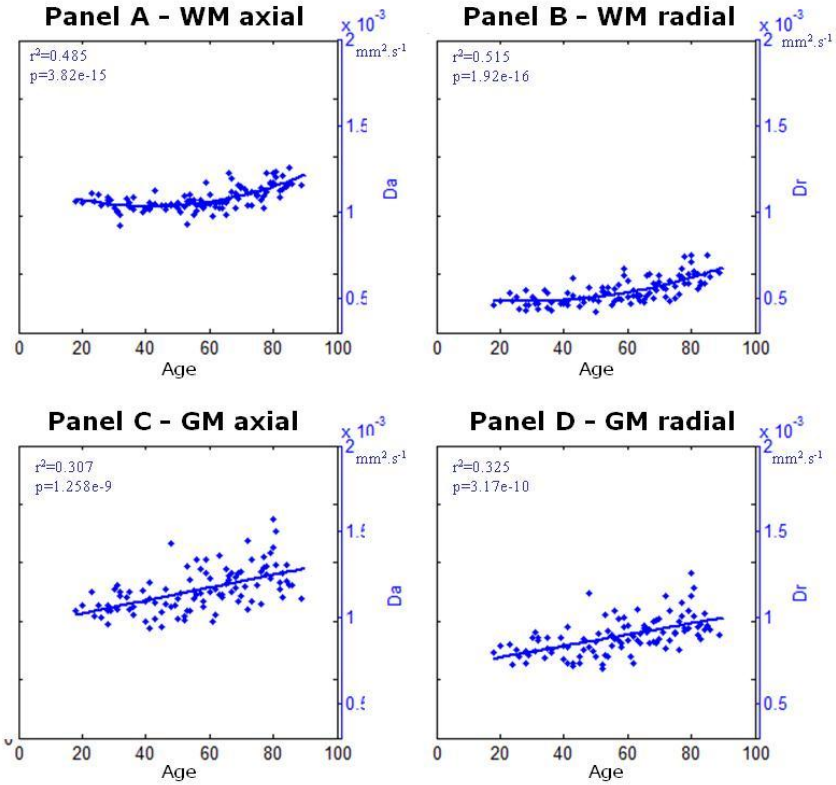



Figure 5.9 – Axial and radial diffusivities changes with age. Panel A and B corresponds to the values of $\lambda_{||}$ and $\lambda_{\perp}$ for white matter, while panel C and D the $\lambda_{||}$ and $\lambda_{\perp}$ values from grey matter.

### 5.3.2. Analysis of the MK prefrontal histogram

Histograms of all MK prefrontal voxels for all 103 Cam-CAN subjects are presented in Figure 5.10. The individual histograms were grouped into seven histograms corresponding to seven age groups. Each histogram bar is subdivided by three areas corresponding to the fraction of WM (red), GM (green), and CSF (blue) voxels. Younger decades show two peaks that seem to be related to larger amounts of WM (right peak) and GM (left peak). For middle aged groups, the GM peak seems to be less evident and a peak corresponding to CSF starts to emerge. Histograms from older decades show also two peaks; however one reflects the increased volume of CSF while the other a mixture between WM and GM voxels.

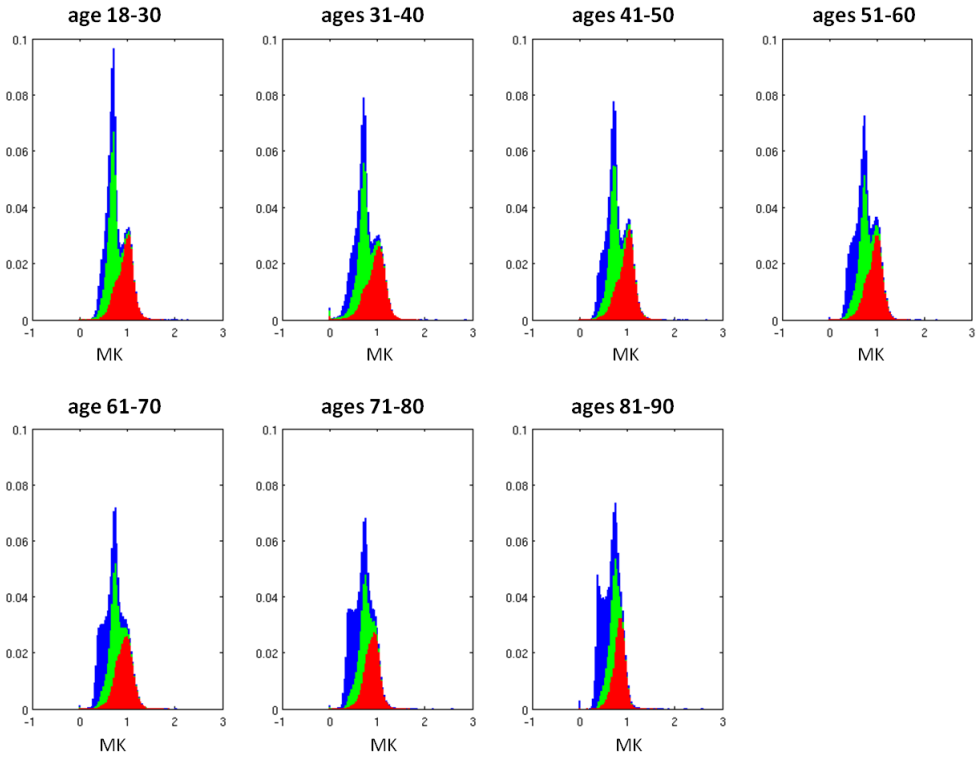



Figure 5.10 – Prefrontal MK voxels histograms for seven different age groups.

The relationship between individual histogram peak positions and age are presented in Figure 5.11. In contrast to the MK mean values of GM presented on Figure 5.4, grey matter's peak position seem to increase over all lifespan (panel A of Figure 5.11). Relative to WM, its related MK peak position shows a linear decrease over all lifespan (panel B of Figure 5.11).

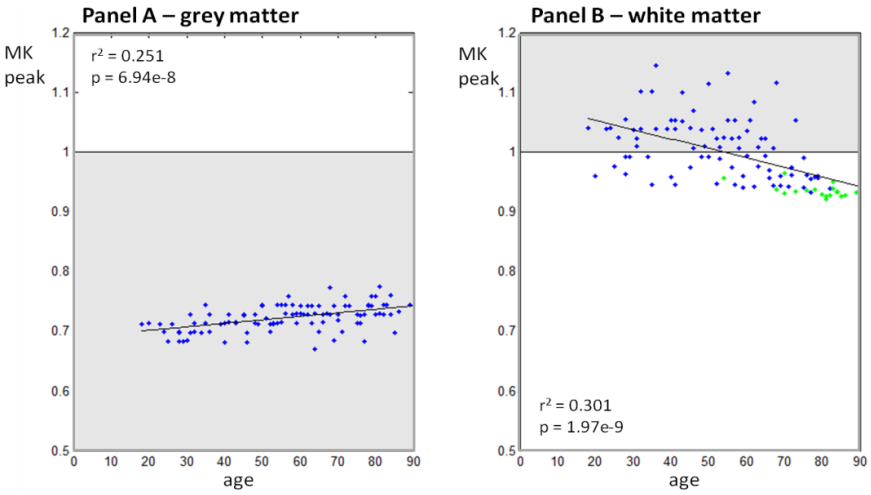



Figure 5.11 – Values of the GM (panel A) and WM (panel B) peak position in order to subject's age. Peak positions were extracted from the MK histograms of the prefrontal region of interest. The assumed typical values of WM and GM for the peak detection algorithm are presented by the grey areas. Green point on panel B correspond to values were WM peak were not evident on the prefrontal voxels histogram.

Figure 5.12 shows the individual MK histograms of WM (red), GM (green), and CSF (blue) prefrontal maps for the seven age groups. The peak positions from the individual WM and GM histograms for all individual Cam-CAN subjects are shown in Figure 5.13. Despite showing an initial increase (until early thirties), the MK peak for GM seems to reach a plateau. On the other hand, the WM peak shows longitudinal changes with similar profiles when compared to the MK mean values of WM presented on Figure 5.4.

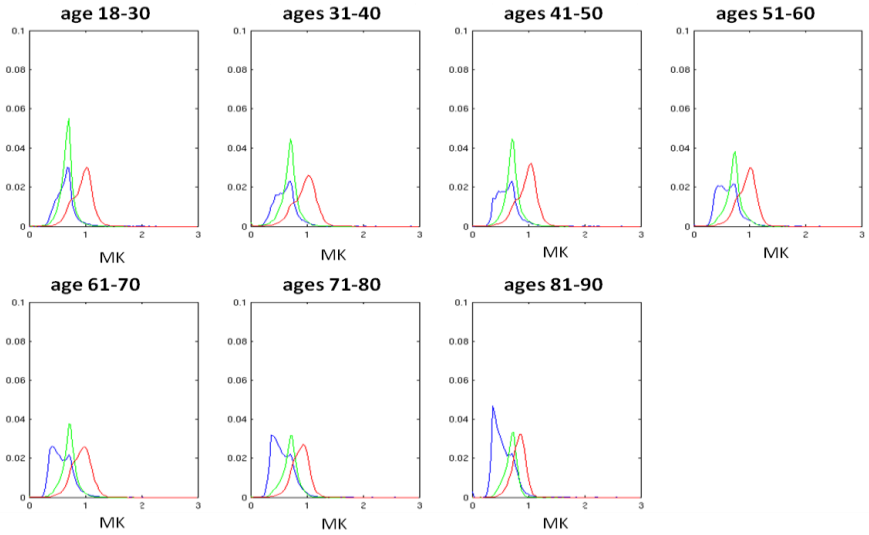



Figure 5.12 – CSF (blue), GM (green), and WM (red) histograms of voxels from seven Cam-CAN age groups.

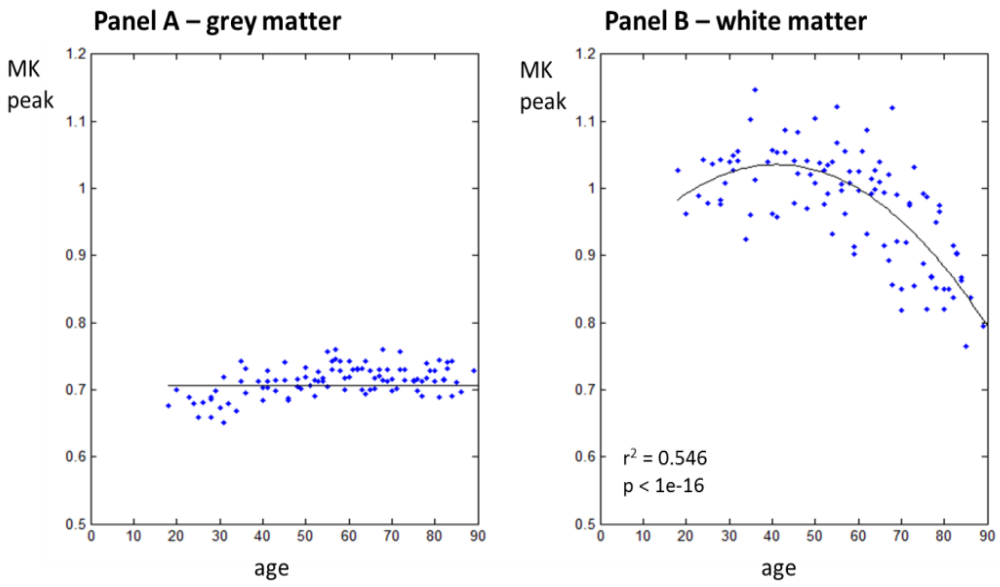



Figure 5.13 – Longitudinal changes of the peak position extracted from GM (panel A) and WM (panel B).

### 5.3.2. TBSS analysis

Figure 5.14 shows the results for track based spatial statistics for FA, MK, and MD volumes. Red regions on the FA skeleton correspond to regions where linear ageing increases were significant, while significant linear ageing decreases are shown in blue. Therefore, it can be observed that global linear increases are not significant on FA and MK skeletons', while global linear decreases are not significant on MD skeletons. Although, MD linear age related changes can be observed in all subjects' common WM tracts, linear decreases are not significant on the left and right corticospinal tracts and on the splemium of the corpus callosum. The spatial profile of the significant MK decreases is different from the spatial profile of FA decrease. In particular, MK is shown to significantly decrease on left and right corticospinal tract and on the splemium of the corpus callosum however such decreases were not visible in regions related to the external capsule.

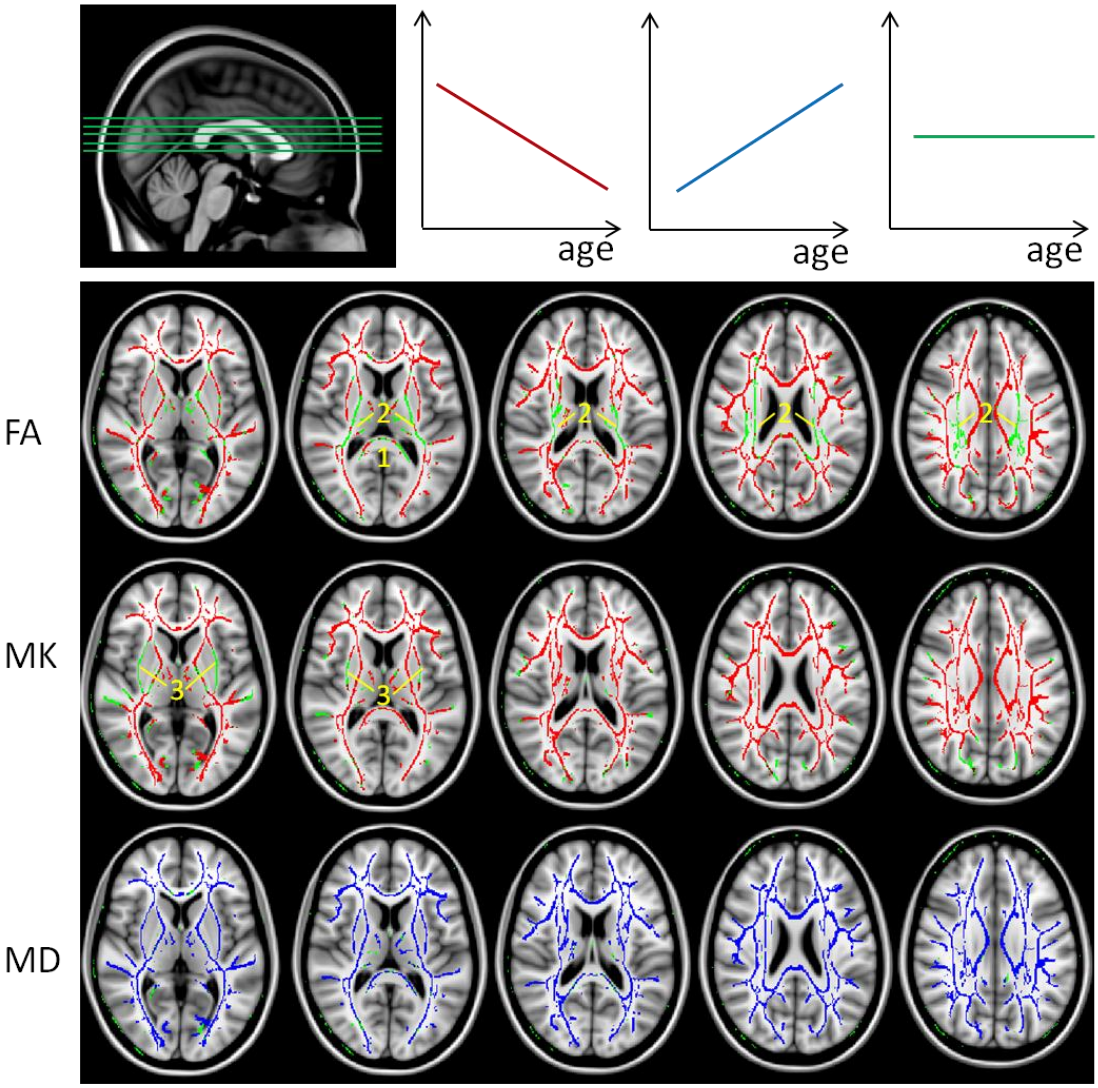



Figure 5.14 - Tract Based Spatial Statistics of age correlated linear increases (blue) and decreases (red) on brain FA, MK, and MD. MD increases are present on all white matter tracts. FA changes are not evident for the corticospinal tract (1) and the corpus callosum splemium (2), while for MK such are not visible for tracts related to the external capsule (3).

## 5.4. Discussion

### 5.4.1. Analysis based on the prefrontal region of interest

**Prefrontal age changes on tissue types percentages**

The results shown in Figure 5.3 are consistent with previous reports. For example, increase of WM until early forties can be related to late myelination processes suggested by *post mortem* studies, (e.g. Benes et al., 1994). Increase of CSF can be explained by a

global atrophy of the prefrontal brain which is more accentuated for older subjects. The decline of GM since younger ages can be only an apparent effect, i.e., even if GM volume is stable on the first decades, its percentage will apparently decrease due to the increases of WM and CSF volumes (Sowell et al., 2004). Therefore, it is possible that degeneration of GM is only significant after middle of the fifth decade when its decrease rate accentuates and when WM also starts to decrease. In addition to the apparent changes, percentage values of tissue types are also limited for not providing information of ageing microstructural changes.

**Prefrontal age changes detected by MK, MD, FA, and FAk from DKI**

Figure 5.4 show that metrics from DKI have potential to measure microstructural changes with age. For instance, decreases of FA on WM can be a consequence of degeneration of oriented microstructures that results on tissue anisotropy decrease, as suggested by previous DTI studies, (e.g. Sullivan & Pfefferbaum, 2006). The processes related to such degeneration may include demyelination of axonal fibres and axonal loss. FA is not sensitive to changes in gray matter due to the isotropic nature of this type of tissue.

MD increases occur first on GM than on WM which is consistent with previous studies suggesting that CSF-filled space increases primarily in GM, (e.g. Raz et al., 2005). For WM regions, MD seems to remain constant in the first analysed decades which can be a cause of the late myelination that can occur until the fourth decade of life, (e.g. Sowell et al., 2003).

MK data suggests that WM tissue barrier complexity increases until the fourth decade. This development of WM is consistent with the increases of WM density previously reported, (e.g. Sowell et al., 2003). Since during these earlier decades FA is shown to decrease, increases on MK can reflect maturation of microstructures of the prefrontal brain that do not have anisotropic proprieties (which in turn can be also a factor of anisotropy decrease). The later decrease in MK can be associated to the global degeneration of the prefrontal white matter. In contrast to FA, MK is sensitive to changes in GM. Increase of GM kurtosis can be explained by increases of the cortical cell-packing density (Falangola et al., 2008), while decreases can reflect axonal of cellular lost and increase of CSF-filled space (or even artefacts introduced by the DKI processing and analysis procedures as mention later).

FAk was shown to be only a noisier version of the FA extracted from the diffusion tensor (Figure 5.5). Therefore, and in contrast to MK, FAk does not show potential to

add more information for ageing studies relative to conventional diffusion tensor metrics. Such observations are consistent with a preliminary DKI study on healthy human for single group and on formalin-fixed rat brains (Hui, Cheung, Qi, & Wu, 2008).

**Age differences detected by axial and radial diffusivity/kurtosis for WM**

As expected, diffusion is larger along the axial direction when compared to the radial direction on WM, since the axial direction corresponds to the direction parallel to the fibres where the random translation of molecules of water is not restricted by barriers (Figure 5.6 panel A and panel B). Since kurtosis is related to barrier complexity, its values are larger for the direction perpendicular to the myelin barriers of the axonal fibres, and thus $k_{\perp}$values are larger than values of$k_{||}$. Interestingly, the ageing degeneration process measured by increases of diffusion and decreases of kurtosis are more significant along radial direction, i.e., for $\lambda_{\perp}$and $k_{\perp}$. This indicates that degeneration during ageing is largely responsible for loss on myelin integrity. Moreover, later increases of $\lambda_{||}$ and decreases of $k_{||}$ relative to the increases of $\lambda_{\perp}$and decreases of $k_{\perp}$ cannot be only a consequence of more stable axonal integrity than myelin integrity but also an indirect consequence of the maturation of isotropic microstructures with age.

**Axial and Radial diffusivity/kurtosis for GM and eigenvalue repulsion artefacts**

As GM does not consist on anisotropic structures, the diffusion and kurtosis values should be similar for all directions. Therefore, directional diffusivities and MD show similar profiles as well as directional kurtosis relative to the MK profiles (Figure 5.7, panel C and D, and Figure 5.4, panel B). In fact, one can argue that radial and axial metrics for diffusion and kurtosis do not have a biological and physical meaning in GM since axial and radial directions cannot be defined on tissues that do not consist of directional oriented structures as axonal fibres; however directional metrics can be used to confirm the profiles shown by MD and MK and to analyse the impacts of Rician noise artefacts (as discussed in the next paragraph).

As mentioned in section 2.3.3, eigenvector repulsion effects induce overestimations of diffusion eigenvalue $\lambda_1$ and underestimations of eigenvalues $\lambda_2$ and $\lambda_3$. Consequently, from Eq. 2.3.20 and Eq. 2.3.21, $\lambda_{||}$ is overestimated while $\lambda_{\perp}$is

underestimated. Therefore, $\lambda_{||}$ shows values that are slightly larger than the values of $\lambda_{\perp}$[6]. $k_{||}$ for GM is also negatively biased since the equations involved on its estimation include the numerical division of the overestimated value of $\lambda_1$ (Eq. 2.4.26 and 2.4.27), while $k_{\perp}$ is positively biased since its estimation includes the numerical division of the underestimated values of $\lambda_2$ and $\lambda_3$ (Eq. 2.4.26 and 2.4.28).

**Pitfalls of the analysis based on mean values of ROI**

There are some limitations of extracting parameters based on mean values of ROIs. Mainly, the process of converting each ROI to each individual space can result in miss-registration errors that will depend on each subject's anatomy. In particular, some age groups can be more prone to this effect than other groups due to typical anatomical age changes, for example the enlargement of brain ventricles and the global brain atrophy. Consequently, the profiles for the different metrics shown here can also be influenced by ROI alignments that are dependent on age. Similarly, the SPM segmentation procedure can be more influenced by miss-classification errors for some specific age groups. Moreover, WM, GM, and CSF brain tissue maps can be corrupted by diffusion-weighted intensities of another tissue type due to the blurring effects introduced by the Gaussian smoothing.

**The validation performed by direct linear DKI and non-linear DTI**

As expected, the MK and MD curves obtained by DLS (Figure 5.7), show larger $r^2$ statistics and smaller p-values compared to the DKI ordinary linear least squares. Despite the ordinary least squares showing inferior accuracy, it produces similar age profiles for MK and MD. Therefore, it has been shown that the ordinary fit is a viable method to replace the direct linear fit for the estimation of the diffusion and kurtosis tensors (bearing in mind the appropriate adequate pre-processing algorithms on DWIs).

Relative to the non-linear DTI, lower $r^2$ statistics and higher p-values for the model fitting on MD, FA, and directional diffusivities suggest poorer ability to detect age changes when compared to the results extracted from DKI. However, such metrics were obtained from data which resolution was not corrupted by the Gaussian smoothing.

[6] Notably, the eigenvalues repulsion is also responsible for the residual amount of anisotropy measured by FA on grey matter, panel B, Figure 5.4

Relative to the metrics extracted from non-linear DTI[7], one can observe that the values extracted from the ordinary DKI shows lower FA values and higher MD values, which is a direct consequence of the blurring effect of the Gaussian kernel, however it shows identical profiles. Therefore, despite the bias introduced, it has been shown that smoothing DWIs with a Gaussian kernel size of 2.5 mm does not impair DKI's ability to detect ageing changes on diffusion and kurtosis metrics.

### 5.4.2. Analysis of the MK prefrontal histogram

**Advantages and disadvantages of using analysis using histograms**

Analysing MK ageing changes on histograms can provide more details than the analysis based on single mean values of metrics over ROIs. Moreover, MK histograms provide a way to extract information for both WM and GM without requiring segmentation techniques, avoiding errors from possible miss-classifications of the different types of tissues (Falangola et al., 2008). However, for the results shown in Figure 5.10 some pitfalls on the use of histograms can be identified. Despite Falangola et al. (2008) assuming that different histogram peaks correspond to different tissue types, the results of Figure 5.10 show that peaks are contaminated with voxels from different types of tissues. For example, during the lifespan the WM peak shows a shift from the typical assumed WM range and for the older age groups the WM peak seems to not be distinguishable from the GM values. Differences between the histogram peak analysis presented on this chapter and presented on Falangola et al. (2008) can be related to methodological differences, in particular on the definition of the prefrontal ROIs, on the followed DKI procedures and on the population sample used.

The peak detection algorithm could be improved by adjusting the GM and WM range priors; however such will not resolve the problem in separating the WM peak from GM on the older subjects. Therefore, it was opted to perform the peak detection based on the MK histograms of the segmented prefrontal WM and GM maps (Figure 5.12). Although this procedure does not avoid possible miss-classifications in the segmented process, it can still provide a different characterization of brain changes with age when compared to the metrics extracted by ROI means.

[7] Interestingly, linear DTI shown a significant reduction of metrics sensitivity which can be related to the large diffusion underestimation introduced by non-Gaussianity.

**Peak ageing changes**

Comparing with the study reported by Falangola et al. (2008), similar ageing profiles for the GM peak position detected from the histograms of all prefrontal ROI voxels are presented (Figure 5.11), i.e., a monotonic increase. As mentioned in Falangola et al. (2008) such ageing profile can be related to the systematic increase of cell-packing density also observed on *post mortem* studies. However, as described in the previous paragraphs, the GM peak profile can also be influenced by the contamination of WM and CSF voxels. On the other hand, the profile of the WM peak is only sensitive to a linear decrease with age, due to the problem of detecting the WM peak.

For the peaks obtained on the histograms from the segmented WM and GM maps, the systematic GM peak shift to higher values of MK are not so evident. In particular, after the early thirties the GM peak position seems to be stable (and that should be the reason why the step down model suggests no ageing correlation with GM peak position). For WM, the profiles of the peak position changes is identical to the ones described by the MK mean on prefrontal WM (panel A of Figure 5.4), and therefore it can be also related to the later maturation process and subsequent demyelination.

Since kurtosis is a very recent modality and the ageing differences between the different features of the GM were never reported before, it is important to note that further analysis is required to validate the profiles observed and to assess which biological mechanisms are responsible for the different age profiles on different diffusion features for a specific tissue type. For example, future Cam-CAN's analysis using the 700 subjects can provide a large significance on distinguishing true biological changes from artefacts introduced by the MRI Rician noise or other possible artefacts introduced by DKI processing and analysis procedures.

### 5.4.3. Analysis based on TBSS

**Preliminary TBSS results show spatial differences in diffusion metrics**

Despite being based only on linear correlation with age, the TBSS analysis reported in this thesis shows significant age changes which depended on different spatial tracts but also on different diffusion and kurtosis metrics. Although significant increase of MD occurs on all the analysed tracts, FA and MK show specific patterns which seem to be

related to specific tracts of the human brain. Such patterns are also different between FA and MK metrics, and therefore it is once more shown that kurtosis values have the potentiality of showing additional information to conventional diffusion metrics.

**Next steps on TBSS analysis**

It is important to note that tracts which did not show linear correlations with age cannot be interpreted as not suffering changes over human lifespan. In particular, such regions can be sensitive to other correlation profiles such as quadratic or even cubic profiles. Therefore, for the next steps of Cam-CAN ageing analysis it will be useful to test the significance of ageing differences using general linear models with terms with higher degree, such as quadratic or cubic terms. After this, one can also analyse which tracts stops maturing earlier, which starts to degenerate later, or which remain mostly unaffected by age. Finally, such profiles can be related to changes in functional anatomy, (e.g. Zarei et al., 2006), by correlating the tract changes and data from functional MRI, and related to cognitive changes by correlating observed changes on tracts to the scores that each subject presented on behaviour tests.

## 5.5. Chapter Summary

In this chapter three approaches to investigate the ageing changes on metrics extracted by DTI and DKI were compared: analysis based on the means across regions of interest, based on histograms, and voxel-based. First the analysis based on the mean across regions of interest showed results that were consistent with previous studies, in particular for the percentage of WM, GM, and CSF, and the metrics extracted from the diffusion tensor. Since DKI is a very recent modality, some of the results shown here were never reported before and, therefore, they should require further analysis to validate the detected ageing changes. However it was already shown that kurtosis metrics have potential to reveal different microstructural changes when compared to conventional metrics extracted from the diffusion tensor. In this chapter, non-linear DTI seems to be useful to validate DKI results from the smoothing blurring effects on metrics extracted from the diffusion tensor, while the more accurate DKI direct fit method supports the MK and MD profiles given by the DKI ordinary linear least square. Despite some pitfalls for the histograms analysis being identified, this type of analysis seems to have potential in providing different features when compared to the mean values extracted from the

prefrontal ROI. In particular the average of MK on the prefrontal GM might show degeneration of axon body's and increase of CSF-filled space on older adults, while GM peak position might reflect a constant increase of GM cell packing. Finally, the preliminary results from tract based spatial statistics have shown ageing differences that not only vary for different white matter tracks but it also vary for the different metrics analysed (MD, MK, and FA).

In general, the indexes of tissue complexity given by MK, $k_{\perp}$, and $k_{||}$ have been shown to provide different information relative to the metrics extracted from the diffusion tensor, and thus it can be concluded that DKI has the potential to identify unique ageing changes on microstructures for the Cam-CAN project.

# Chapter 6

# Conclusions and Future Work

## 6.1. Main objectives accomplished

The results presented, analysed, and discussed in the previous chapters had two relevant contributions for the Cam-CAN project according to the two objectives proposed in the beginning of my master's thesis (which were mentioned on section 1.3).

First of all, an accurate procedure for DKI processing was suggested based on the validation tests documented in Chapters 3 and 4. All these processing steps were implemented on automatized scripts that are now available to all Cam-CAN members. In addition to the methods for DKI processing, the DKI analysis procedures are also available, which includes the methods for automatically converting ROIs from a stereotypical image to each subject space as well as the methods to extract automatically quantitative parameters for DWIs (as mean values of different metrics from diffusion and kurtosis tensors and the MK histograms' peak positions).

The second contribution to the Cam-CAN project was showing the potential of DKI in providing unique information regarding the microstructural changes in the healthy brain ageing. This was performed with three different approaches (Chapter 5): analysis based on ROI means; histogram based analysis; and TBSS.

## 6.2. Summary of the results of this thesis and impacts

The results presented in this thesis are not only relevant for the purposes of the Cam-CAN project but they also show relevant and general findings on the development of DKI

estimation framework and on human healthy ageing. Therefore in the next two sub-sections the results of the thesis will be summarized and their impacts discussed.

### 6.2.1. Results on DKI methodological improvements and their impacts

In Chapter 3, it was shown for the first time that linear approaches for extracting the diffusion and kurtosis metrics from DKI have similar accuracy when compared to non-linear DKI approaches. As discussed on section 3.4, this could not be intuitively expected since this was not observed on previous DTI studies. Moreover, results from the diffusion tensor estimated by linear DKI were shown to be similar to the ones extracted from the diffusion tensor estimated by non-linear DTI (section 5.4). Consequently, the findings of this thesis show that even simple DKI linear approaches can be used to produce results similar to the more accurate procedures of DTI. Moreover, since linear approaches can be implemented with high computational efficiency, DKI shows potential in providing faster algorithms for DWI analysis than non-linear DTI algorithms.

Despite being one of the approaches most followed in recent papers (e.g. Fieremans, Jensen, & Helpern, 2011; Gao et al., 2012), the results of Chapter 3 show for the first time pitfalls of using the constrained linear least square method based on the approach suggested by Tabesh et al. (2011). In particular, constrains were shown to induce large amounts of overestimations and underestimations and thus constrained linear least squares might not provide a viable alternative fitting method relative to the simplest unconstrained linear least solutions. Therefore, further developments on diffusion and kurtosis tensors estimation framework are still required.

Chapter 3 showed advances on the estimation of MK and MD by using the proposed direct linear least squares fit. Despite not providing estimates of the diffusion and kurtosis tensors, the values estimated by this method can provide accurate measures for studies that only required values of MK and MD.

Some findings regarding the effect of noise and artefacts on the DKI estimation framework were also reported in Chapter 3. Previous studies reported some consequences of noise and artefacts on DKI such as the observed implausible values of kurtosis (e.g. Tabesh et al. 2011); however the analysis presented in Chapter 3 of this thesis goes deeply on this issue, pointing some specific characteristics of DKI noise effects. For instance, it was shown that: 1) MK is more sensitive to noise and artefacts than MD; 2) kurtosis is largely underestimated for voxels with low values of diffusion; 3)

FA values are largely overestimated from voxels with low anisotropy; and 4) overestimation of kurtosis due to rectified noise are not significant for the b-values used by Cam-CAN. Such information can be useful for future work on distinguishing real biological changes from artefacts on DKI and for the development of more robust DKI procedures.

Despite previous studies having already argued that registration-based algorithms are inadequate to correct motion artefacts on diffusion-weighted data (e.g. Rohde et al., 2004), they are still used currently since more sophisticated alternative are not yet provided on the current toolboxes available. However, with the results of Chapter 4 it was shown that motion correction procedures based only on the registration relative a single DWI volume can corrupt the images with additional translations and rotations instead of reducing the head motion effects. Therefore, from the results of this thesis, it is highly recommended not to use such approach on diffusion MRI. If alternatives are not available, one should verify the head motion correction results from the image registration algorithm since the use of DWI with no correction can be more suitable.

For the first time the FWHM value of 2.5 mm for the Gaussian kernel to smoothed DWIs which resolution of 2x2x2 $mm^3$ was experimentally justified in Chapter 4. Moreover, apart from faster acquisition times, data acquisition at a lower resolution does not show advantages for DKI since they were not sufficient to overcome the artefact of implausible kurtosis values. Finally, the results of Chapter 4 have shown that data acquired with larger voxel size does not necessary require larger filter size. Notably, these results provide useful information for the selection of the filter FWHM size for future studies.

### 6.2.2. Results on the preliminary study of ageing using DKI and impacts

The preliminary ageing study using DKI presented in Chapter 5 shows results that are consistent with previous DTI studies. Briefly, mean diffusivity increases and white matter anisotropy was shown to decrease from the early twenties. Since DKI is a very recent technique, kurtosis values can require additional test to validate its accuracy as an index of microstructural barrier complexity. However, from the results of this thesis, kurtosis already shows potential to reveal unique information and to allow the formulation of new hypothesis for the healthy ageing of the brain. For example, due to

the increase of white matter kurtosis until the forties, it can be hypothesized that the initial FA decrease cannot be only related to a degenerative process but might be due to maturation of white matter structures with low degrees of anisotropy, such as the maturation of regions with crossing fibres.

Relative to the previous preliminary ageing study (Falangola et al. 2008) the results of Chapter 5 present a more detailed description of the ageing changes, since it uses a larger sample of subjects. Moreover, it presents for the first time the age variation of kurtosis and diffusion values from radial and axial direction relative to axonal fibres. Chapter 5 also reveals some pitfalls on analyses based on peak positions of MK histograms. For instance, grey matter peak position values were shown to be inconsistent with the profiles extracted from mean values of MK within the prefrontal grey matter ROI. Therefore, the association of MK increases on grey matter to the increase of cortical cell packing density (as hypothesized by Falangola et al., 2008) should be further investigated for its validation.

Finally, TBSS shows that age related changes occur differently on specific tracts, which suggests that changes in cognitive abilities can be related to changes on the white matter connectivity between regions of the human brain.

## 6.3. Future Steps

### 6.3.1. Future steps on DKI methodological improvements

Despite an accurate DKI procedure for Cam-CAN having been suggested from the results of Chapters 3 and 4, further improvements can still be performed on DKI procedures. For instance, although the results of Chapter 4 show that image registration based methods are not adequate for correcting DWIs, no alternative correction for head motion artefacts was proposed. Instead, for the preliminary ageing study presented in Chapter 5, data from subjects that largely move during the acquisition of diffusion-weighted data (section 5.2) was removed. To avoid this exclusion of data on future Cam-CAN studies, the development of procedures for DKI should be followed by introducing an adequate and accurate algorithm for head motion correction, for example based on the recent suggestion of Ben-Amitay, Jones, and Assaf (2012) which proposed a correction algorithm based on pre-estimations of DTI and simulations using a composite hindered

and restricted model of diffusion framework. After reducing the impact of motion artefacts on the DWIs, more robust methods for image noise-removal can be introduced, such as approaches that directly account for the diffusion Rician noise bias, e.g. techniques based on a Bayesian framework (Basu et al., 2006) or on non-local means algorithms (Wiest-Daesslé et al., 2008). More sophisticated algorithms to extract the diffusion and kurtosis metrics from pre-treated DWIs can also be studied, for example the fitting methods that account directly for the Rician noise nature (e.g. Veraart et al., 2011b), or tensor fitting methods that incorporate the MK and MD information extracted from the more accurate DLS method proposed in Chapter 3 of this thesis. Finally, additional metrics to extract information from diffusion and kurtosis tensors can be developed and implemented in the DKI processing pipelines, for example the DKI metrics suggested by Fieremans, Jensen, and Helpern (2011), which have the potential to distinguish different types of white matter pathology, such as demyelination and axonal loss.

### 6.3.2. Future steps on DKI on human brain ageing

After acquisition of data for the 700 healthy subjects, the DKI procedures will be applyed to the remaining data to increase the power of the analysis based on regions of interest, histograms, and TBSS. Regarding the analysis based on mean values and histogram peak positions, other regions of interest could be tested to study the brain spatial changes on both white and grey matter. As mentioned at the end of Chapter 5, TBSS analysis can be modified to test different ageing models, such as general linear models with quadratic or cubic profiles. Other voxel based analysis can also be performed on the diffusion-weighted Cam-CAN data, for example standard voxel-based morphometry techniques to identify spatial ageing patterns on grey matter. The biological interpretation of diffusion and kurtosis changes with age can be explored by relating the given results to other *post mortem* studies, by simulations based on *monte carlo* algorithms, and by biological models (Fieremans et al., 2011).

Finally, the relationship between the microstructural changes measured by DKI and the brain cognition changes quantified by the behavioural measures of cognitive abilities (memory, attention, emotion, language, and action) will be studied, as well as the relationship between changes on DKI metrics and the functional anatomy information

provided by the Cam-CAN's functional MRI datasets. These studies will aim to provide unique information for the Cam-CAN project in determining the extent of neural flexibility across the lifespan and the potential for neural re-organisation to sustain mental cognitive, being fundamental for the future understanding on how can we preserve cognitive capabilities and increase our quality of life along our lifespan.

Rafael Neto Henriques
October 2012

## References


Aja-Fernandez, S., Alberola-Lopez, C., & Westin, C.-F. (2008). Noise and Signal Estimation in Magnitude MRI and Rician Distributed Images: A LMMSE Approach. *IEEE Transactions on Image Processing*, *17*(8), 1383 -1398. doi:10.1109/TIP.2008.925382

Anderson, M. J., & Robinson, J. (2001). Permutation Tests for Linear Models. *Australian & New Zealand Journal of Statistics*, *43*(1), 75–88. doi:10.1111/1467-842X.00156

Assaf, Y., & Cohen, Y. (1998). Non-Mono-Exponential Attenuation of Water andN-Acetyl Aspartate Signals Due to Diffusion in Brain Tissue. *Journal of Magnetic* Resonance, 131(1), 69-85. doi:10.1006/jmre.1997.1313

Barmpoutis, A., & Zhuo, J. (2011). Diffusion kurtosis imaging: Robust estimation from DW-MRI using homogeneous polynomials. *ISBI* (p. 262-265). IEEE. Recuperado de http://dblp.uni-trier.de/db/conf/isbi/isbi2011.html#BarmpoutisZ11

Bartzokis, G., Beckson, M., Lu, P. H., Nuechterlein, K. H., Edwards, N., & Mintz, J. (2001). Age-related changes in frontal and temporal lobe volumes in men: a magnetic resonance imaging study. *Archives of general psychiatry*, *58*(5), 461-465.

Basser, P. J. (1995). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in biomedicine*, *8*(7-8), 333-344.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994a). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of magnetic resonance. Series B*, *103*(3), 247-254.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994b). MR diffusion tensor spectroscopy and imaging. Biophysical Journal, 66(1), 259-267. doi:10.1016/S0006-3495(94)80775-1

Bastin, M. E. (1999). Correction of eddy current-induced artefacts in diffusion tensor imaging using iterative cross-correlation. *Magnetic Resonance Imaging*, *17*(7), 1011-1024. doi:10.1016/S0730-725X(99)00026-0

Basu, S., Fletcher, T., & Whitaker, R. (2006). Rician noise removal in diffusion tensor MRI. *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, *9*(Pt 1), 117-125

Behrens, T. e. j., Woolrich, M. w., Jenkinson, M., Johansen-Berg, H., Nunes, R. g., Clare, S., Matthews, P. m., et al. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, *50*(5), 1077–1088. doi:10.1002/mrm.10609

Benes, F. M., Turtle, M., Khan, Y., & Farol, P. (1994). Myelination of a key relay zone in the hippocampal formation occurs in the human brain during childhood, adolescence, and adulthood. *Archives of general psychiatry*, *51*(6), 477-484.

Berg, H. C. (1993). *Random Walks in Biology* (Revised.). Princeton University Press.

Brody, H. (1992). The aging brain. *Acta neurologica Scandinavica. Supplementum*, *137*, 40-44.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging Gracefully: Compensatory Brain Activity in High-Performing Older Adults. *NeuroImage*, *17*(3), 1394-1402. doi:10.1006/nimg.2002.1280

Carr, H. Y., & Purcell, E. M. (1954). Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical Review*, *94*(3), 630-638.

Cheung, J. S., Wang, E., Lo, E. H., & Sun, P. Z. (2012). Stratification of heterogeneous diffusion MRI ischemic lesion with kurtosis imaging: evaluation of mean diffusion and kurtosis MRI mismatch in an animal model of transient focal ischemia. *Stroke; a journal of cerebral circulation*, *43*(8), 2252-2254. doi:10.1161/STROKEAHA.112.661926

Cheung, M. M., Hui, E. S., Chan, K. C., Helpern, J. A., Qi, L., & Wu, E. X. (2009). Does diffusion kurtosis imaging lead to better neural tissue characterization? A rodent brain maturation study. *NeuroImage*, *45*(2), 386-392. doi:10.1016/j.neuroimage.2008.12.018

Clark, C. A., Hedehus, M., & Moseley, M. E. (2002). In vivo mapping of the fast and slow diffusion tensors in human brain. *Magnetic Resonance in Medicine*, *47*(4), 623–628. doi:10.1002/mrm.10118

Conolly, S., Macovski, A., Schenck, J., Kwong, K. K., Chesler, D. A., Hu, X., Chen, W., et al. (2000). Magnetic Resonance Imaging. *The Biomedical Engineering Handbook* (second., Vol. I, p. 63-1). Springer.

Correia, M. M. (2009). *Development of Methods for the Acquisition and Analysis of Diffusion Weighted MRI Data*. University of Cambridge.

Dixon, M. (2010). Cam-CAN. *Cambridge Centre for Ageing and Neuroscience*. Recuperado agosto 2, 2012, de http://www.cam-can.org/

Einstein, A. (1956). *Investigations on the Theory of the Brownian Movement*. Courier Dover Publications.

Evans, A. C., Marrett, S., Neelin, P., Collins, L., Worsley, K., Dai, W., Milot, S., et al. (1992). Anatomical mapping of functional activation in stereotactic coordinate space. *NeuroImage*, *1*(1), 43-53.

Falangola, M. F., Jensen, J. H., Babb, J. S., Hu, C., Castellanos, F. X., Di Martino, A., Ferris, S. H., et al. (2008). Age-related non-Gaussian diffusion patterns in the prefrontal brain. *Journal of Magnetic Resonance Imaging*, *28*(6), 1345–1350. doi:10.1002/jmri.21604

Fick, A. (1855). Concerns diffusion and concentration gradient. *Annals of Physics*, 170:59.

Fieremans, E., Jensen, J. H., & Helpern, J. A. (2011). White matter characterization with diffusional kurtosis imaging. *NeuroImage*, *58*(1), 177-188. doi:10.1016/j.neuroimage.2011.06.006

Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., et al. (2011). Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. *NeuroImage*, *56*(1), 220-234. doi:10.1016/j.neuroimage.2011.01.032

Gao, Y., Zhang, Y., Wong, C.-S., Wu, P.-M., Zhang, Z., Gao, J., Qiu, D., et al. (2012). Diffusion abnormalities in temporal lobes of children with temporal lobe epilepsy: a preliminary diffusional kurtosis imaging study and comparison with diffusion tensor imaging. NMR in Biomedicine, n/a–n/a. doi:10.1002/nbm.2809

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, *2*(10), 861-863. doi:10.1038/13158

Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., Pietrini, P., et al. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *The Journal of Neuroscience*, *14*(3), 1450-1462.

Hahn, E. L. (1950). Spin Echoes. *Physical Review*, *80*(4), 580-594. doi:10.1103/PhysRev.80.580

Helpern, J. A., Adisetiyo, V., Falangola, M. F., Hu, C., Di Martino, A., Williams, K., Castellanos, F. X., et al. (2011). Preliminary evidence of altered gray and white matter microstructural development in the frontal lobe of adolescents with

attention-deficit hyperactivity disorder: A diffusional kurtosis imaging study. *Journal of Magnetic Resonance Imaging*, *33*(1), 17–23. doi:10.1002/jmri.22397

Hui, E. S., Cheung, M. M., Qi, L., & Wu, E. X. (2008). Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. *NeuroImage*, *42*(1), 122-134. doi:10.1016/j.neuroimage.2008.04.237

Jaynes, E. T. (1989). *Clearing Up Mysteries - The Original Goal*.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. (G. L. Bretthorst, Org.). Cambridge University Press.

Jensen, J. H., Falangola, M. F., Hu, C., Tabesh, A., Rapalino, O., Lo, C., & Helpern, J. A. (2011). Preliminary observations of increased diffusional kurtosis in human brain following recent cerebral infarction. *NMR in Biomedicine*, *24*(5), 452–457. doi:10.1002/nbm.1610

Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., & Kaczynski, K. (2005). Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *53*(6), 1432-1440. doi:10.1002/mrm.20508

Johansen-Berg, H., & Behrens, T. E. J. (2009). *Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy*. Academic Press.

Jones, D. K., & Basser, P. J. (2004). "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine*, *52*(5), 979–993. doi:10.1002/mrm.20283

Jones, D. K., & Cercignani, M. (2010). Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR in Biomedicine*, *23*(7), 803–820. doi:10.1002/nbm.1543

Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory* (1º ed.). Prentice Hall.

Kingsley, P. B., & Monahan, W. G. (2005). Contrast-to-noise ratios of diffusion anisotropy indices. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *53*(4), 911-918. doi:10.1002/mrm.20433

Koay, C. G., Carew, J. D., Alexander, A. L., Basser, P. J., & Meyerand, M. E. (2006). Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging. *Magnetic Resonance in Medicine*, *55*(4), 930–936. doi:10.1002/mrm.20832

Le Bihan, D, Breton, E., Lallemand, D., Grenier, P., Cabanis, E., & Laval-Jeantet, M. (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, *161*(2), 401-407.

Le Bihan, Denis, Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, *13*(4), 534–546. doi:10.1002/jmri.1076

Lu, H., Jensen, J. H., Ramani, A., & Helpern, J. A. (2006). Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, *19*(2), 236–247. doi:10.1002/nbm.1020

Mattiello, J., Basser, P. J., & LeBihan, D. (1994). Analytical Expressions for the b Matrix in NMR Diffusion Imaging and Spectroscopy. *Journal of Magnetic Resonance*, A 108, 131-141.

McRobbie, D. W., Moore, E. A., Graves, M. J., & Prince, M. R. (2003). *MRI from Picture to Proton*. Cambridge University Press.

Miller, D. H., Grossman, R. I., Reingold, S. C., & McFarland, H. F. (1998). The role of magnetic resonance techniques in understanding and managing multiple sclerosis. *Brain*, *121*(1), 3-24. doi:10.1093/brain/121.1.3

Mori, S., & Barker, P. B. (1999). Diffusion magnetic resonance imaging: Its principle and applications. *The Anatomical Record*, *257*(3), 102–109. doi:10.1002/(SICI)1097-0185(19990615)257:3<102::AID-AR7>3.0.CO;2-6

Moseley, M. (2002). Diffusion tensor imaging and aging – a review. *NMR in Biomedicine*, *15*(7-8), 553–560. doi:10.1002/nbm.785

Moseley, M. E., Cohen, Y., Kucharczyk, J., Mintorovitch, J., Asgari, H. S., Wendland, M. F., Tsuruda, J., et al. (1990). Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology*, *176*(2), 439-445.

Niendorf, T, Dijkhuizen, R. M., Norris, D. G., van Lookeren Campagne, M., & Nicolay, K. (1996). Biexponential diffusion attenuation in various states of brain tissue: implications for diffusion-weighted imaging. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *36*(6), 847-857.

Niendorf, Thoralf, Dijkhuizen, R. M., Norris, D. G., van Lookeren Campagne, M., & Nicolay, K. (1996). Biexponential diffusion attenuation in various states of brain tissue: Implications for diffusion-weighted imaging. *Magnetic Resonance in Medicine*, *36*(6), 847–857. doi:10.1002/mrm.1910360607

Özarslan, E., & Mareci, T. H. (2003). Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magn. Reson. Med*, 955–965.

Pierpaoli, C., & Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine*, *36*(6), 893–906. doi:10.1002/mrm.1910360612

Qi, L., Wang, Y., & Wu, E. X. (2008). D-eigenvalues of diffusion kurtosis tensors. *Journal of Computational and Applied Mathematics*, *221*(1), 150-157. doi:10.1016/j.cam.2007.10.012

Raab, P., Hattingen, E., Franz, K., Zanella, F. E., & Lanfermann, H. (2010). Cerebral Gliomas: Diffusional Kurtosis Imaging Analysis of Microstructural Differences1. *Radiology*, *254*(3), 876-881. doi:10.1148/radiol.09090819

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., et al. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*, *15*(11), 1676-1689. doi:10.1093/cercor/bhi044

Reese, T. g., Heid, O., Weisskoff, R. m., & Wedeen, V. j. (2003). Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magnetic Resonance in Medicine*, *49*(1), 177–182. doi:10.1002/mrm.10308

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science*, *17*(3), 177-182. doi:10.1111/j.1467-8721.2008.00570.x

Reuter-Lorenz, P. A., & Park, D. C. (2010). Human neuroscience and the aging mind: a new look at old problems. *The journals of gerontology. Series B, Psychological sciences and social sciences*, *65*(4), 405-415. doi:10.1093/geronb/gbq035

Rohde, G. k., Barnett, A. s., Basser, P. j., Marenco, S., & Pierpaoli, C. (2004). Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magnetic Resonance in Medicine*, *51*(1), 103–114. doi:10.1002/mrm.10677

Rosenkrantz, A. B., Sigmund, E. E., Winnick, A., Niver, B. E., Spieler, B., Morgan, G. R., & Hajdu, C. H. ([s.d.]). Assessment of hepatocellular carcinoma using apparent diffusion coefficient and diffusion kurtosis indices: preliminary experience in fresh liver explants. *Magnetic Resonance Imaging*, (0). doi:10.1016/j.mri.2012.04.020

Sandson, T. A., Felician, O., Edelman, R. R., & Warach, S. (1999). Diffusion-weighted magnetic resonance imaging in Alzheimer's disease. *Dementia and geriatric cognitive disorders, 10*(2), 166-171.

Sevick, R. J., Kucharczyk, J., Mintorovitch, J., Moseley, M. E., Derugin, N., & Norman, D. (1990). Diffusion-weighted MR imaging and T2-weighted MR imaging in acute cerebral ischaemia: comparison and correlation with histopathology. *Acta neurochirurgica. Supplementum, 51*, 210-212.

Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., et al. (2008). Neurodevelopmental Trajectories of the Human Cerebral Cortex. *The Journal of Neuroscience, 28*(14), 3586-3594. doi:10.1523/JNEUROSCI.5309-07.2008

Sijbers, J., den Dekker, A. J., Scheunders, P., & Van Dyck, D. (1998). Maximum-likelihood estimation of Rician distribution parameters. *IEEE Transactions on Medical Imaging, 17*(3), 357 -361. doi:10.1109/42.712125

Simon, T. J., Ding, L., Bish, J. P., McDonald-McGinn, D. M., Zackai, E. H., & Gee, J. (2005). Volumetric, connective, and morphologic changes in the brains of children with chromosome 22q11.2 deletion syndrome: an integrative study. *NeuroImage, 25*(1), 169-180. doi:10.1016/j.neuroimage.2004.11.018

Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping, 17*(3), 143-155. doi:10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage, 31*(4), 1487-1505. doi:10.1016/j.neuroimage.2006.02.024

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage, 44*(1), 83-98. doi:10.1016/j.neuroimage.2008.03.061

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience, 6*(3), 309-315. doi:10.1038/nn1008

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience, 2*(10), 859-861. doi:10.1038/13154

Sowell, E. R., Thompson, P. M., & Toga, A. W. (2004). Mapping changes in the human cortex throughout the span of life. *The Neuroscientist: a review journal bringing neurobiology, neurology and psychiatry*, *10*(4), 372-392. doi:10.1177/1073858404263960

Stejskal, E. O. (1965). Use of Spin Echoes in a Pulsed Magnetic-Field Gradient to Study Anisotropic, Restricted Diffusion and Flow. *The Journal of Chemical Physics*, *43*(10), 3597-3603. doi:doi:10.1063/1.1696526

Stejskal, E. O., & Tanner, J. E. (1965). Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *The Journal of Chemical Physics*, *42*(1), 288-292. doi:doi:10.1063/1.1695690

Sundgren, P. C., Dong, Q., Gómez-Hassan, D., Mukherji, S. K., Maly, P., & Welsh, R. (2004). Diffusion tensor imaging of the brain: review of clinical applications. *Neuroradiology*, *46*(5), 339-350. doi:10.1007/s00234-003-1114-x

Sullivan, E. V., & Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neuroscience & Biobehavioral Reviews*, *30*(6), 749-761. doi:10.1016/j.neubiorev.2006.06.002

Tabesh, A., Jensen, J. H., Ardekani, B. A., & Helpern, J. A. (2011). Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. *Magnetic Resonance in Medicine*, *65*(3), 823–836. doi:10.1002/mrm.22655

Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine*, *52*(6), 1358–1372. doi:10.1002/mrm.20279

Vallesi, A., McIntosh, A. R., & Stuss, D. T. (2011). Overrecruitment in the aging brain as a function of task demands: evidence for a compensatory view. *Journal of cognitive neuroscience*, *23*(4), 801-815. doi:10.1162/jocn.2010.21490

Veraart, J., Poot, D. H. J., Van Hecke, W., Blockx, I., Van der Linden, A., Verhoye, M., & Sijbers, J. (2011a). More accurate estimation of diffusion tensor parameters using diffusion Kurtosis imaging. *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *65*(1), 138-145. doi:10.1002/mrm.22603

Veraart, J., Van Henke, W., Sijbers, J. (2011b). Constrained Maximum Likelihood Estimation. *Magnetic Resonance in Medicine, 66*, 678-686. doi: 10.1002/mrm.22835

Wiest-Daesslé, N., Prima, S., Coupé, P., Morrissey, S. P., & Barillot, C. (2008). Rician noise removal by non-Local Means filtering for low signal-to-noise ratio MRI:

applications to DT-MRI. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, *11*(Pt 2), 171-17

Wilkinson, D. S. (2000). *Mass Transport in Solids and Fluids*. Cambridge University Press.

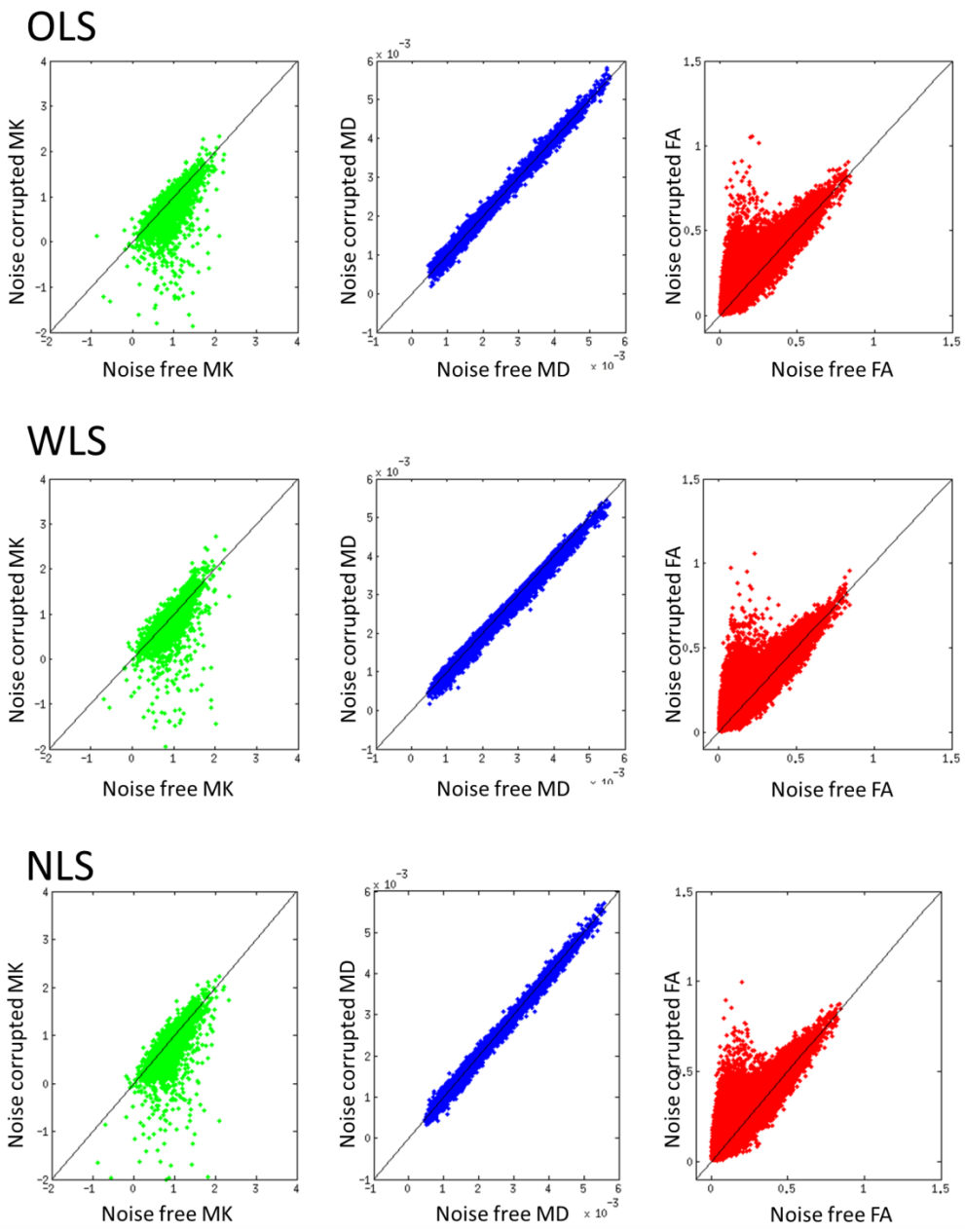
Wu, E. X., & Cheung, M. M. (2010). MR diffusion kurtosis imaging for neural tissue characterization. *NMR in Biomedicine*, *23*(7), 836–848. doi:10.1002/nbm.1506

Yablonskiy, D. A., Bretthorst, G. L., & Ackerman, J. J. H. (2003). Statistical Model for Diffusion Attenuated MR Signal. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *50*(4), 664-669. doi:10.1002/mrm.10578

Zarei, M., Johansen-Berg, H., Smith, S., Ciccarelli, O., Thompson, A. J., & Matthews, P. M. (2006). Functional anatomy of interhemispheric cortical connections in the human brain. *Journal of Anatomy*, *209*(3), 311–320. doi:10.1111/j.1469-7580.2006.00615.x

# Appendix A

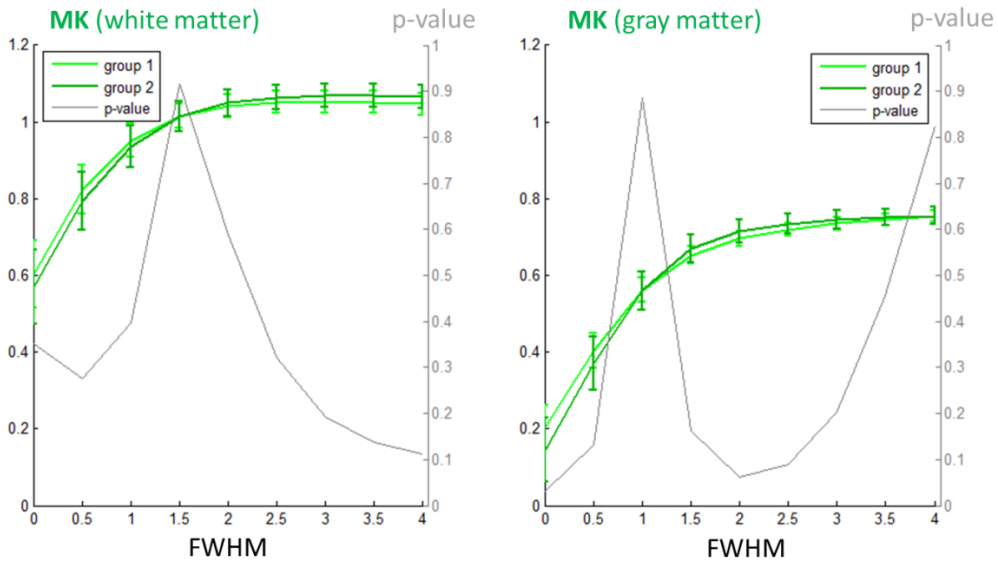
# Scatter plots between noise free and Rician corrupted



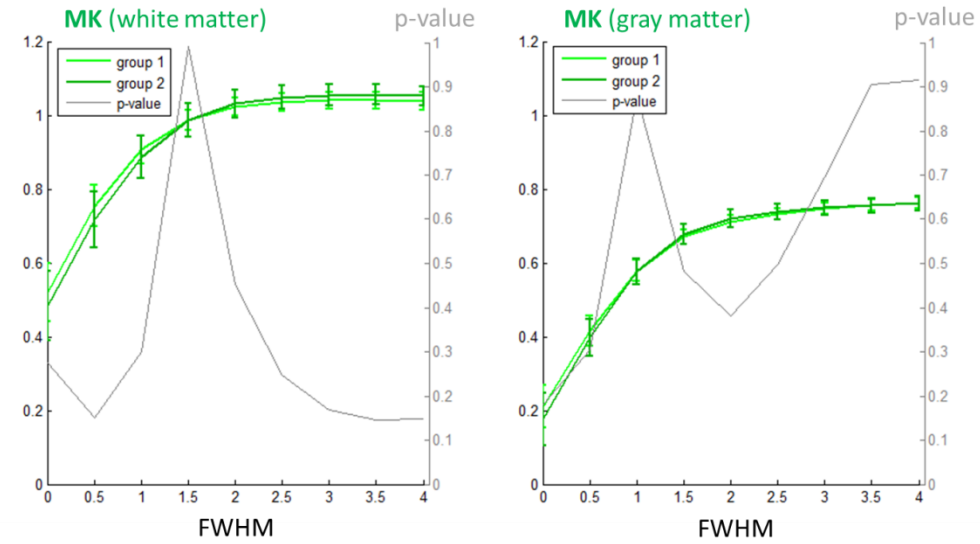

**A.1. Scatter plots between noise free and Rician corrupted values for WLS and NLS**

# Appendix B

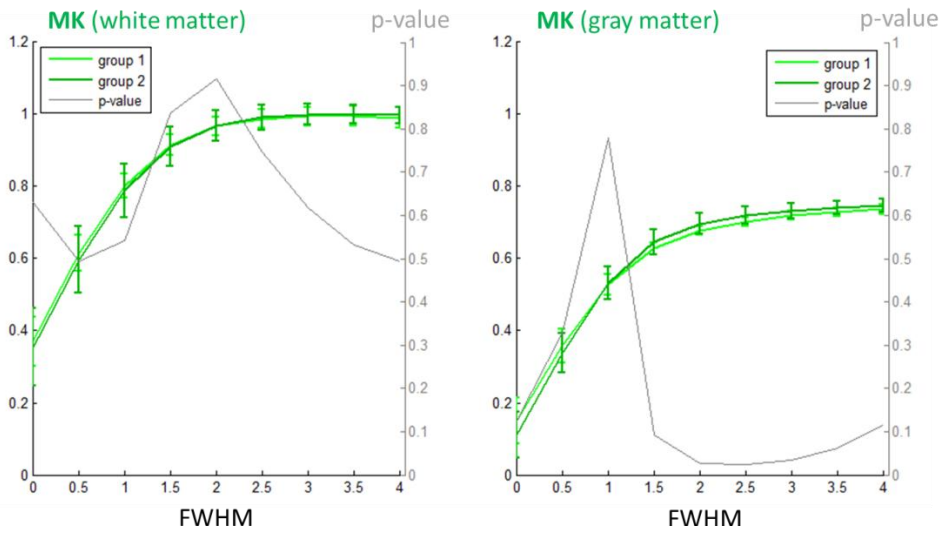
# Age differences of white and grey matter between two ageing groups using OLS



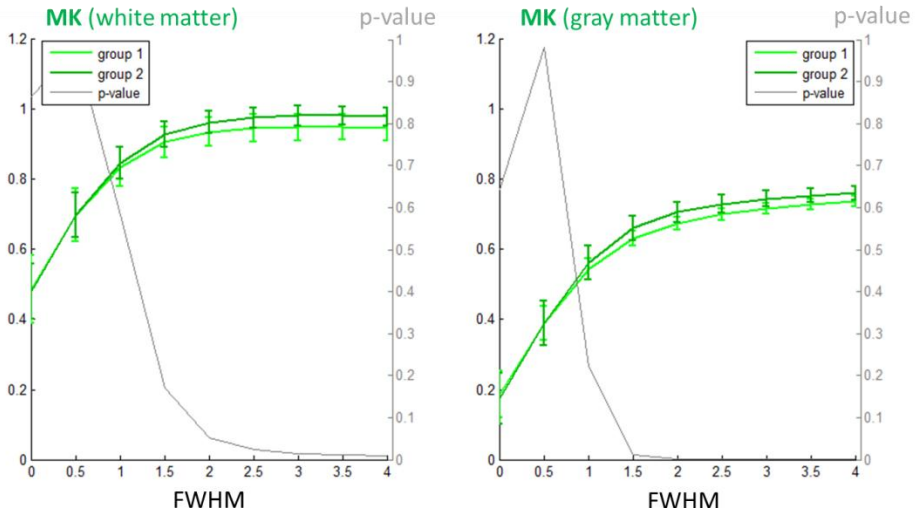

**B.1. Sub-region 3**



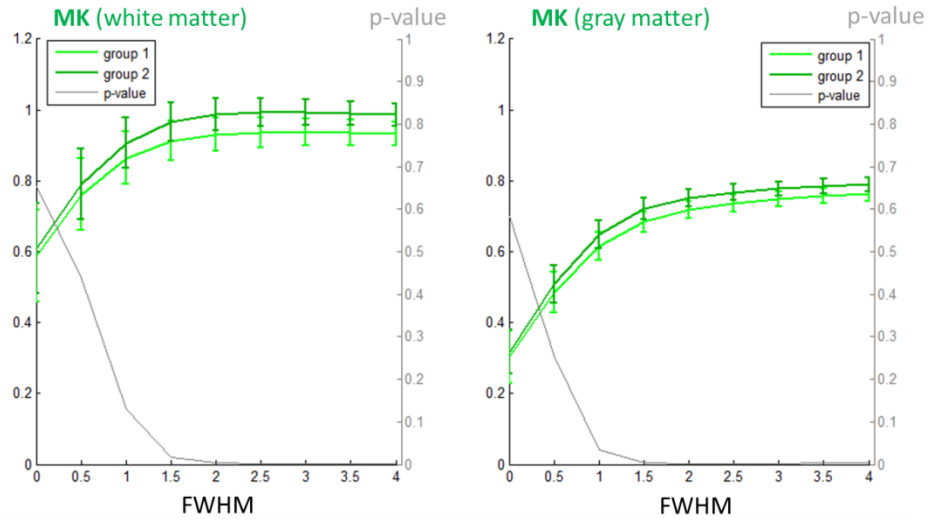

**B.2. Sub-region 4**




## B.3. Sub-region 5



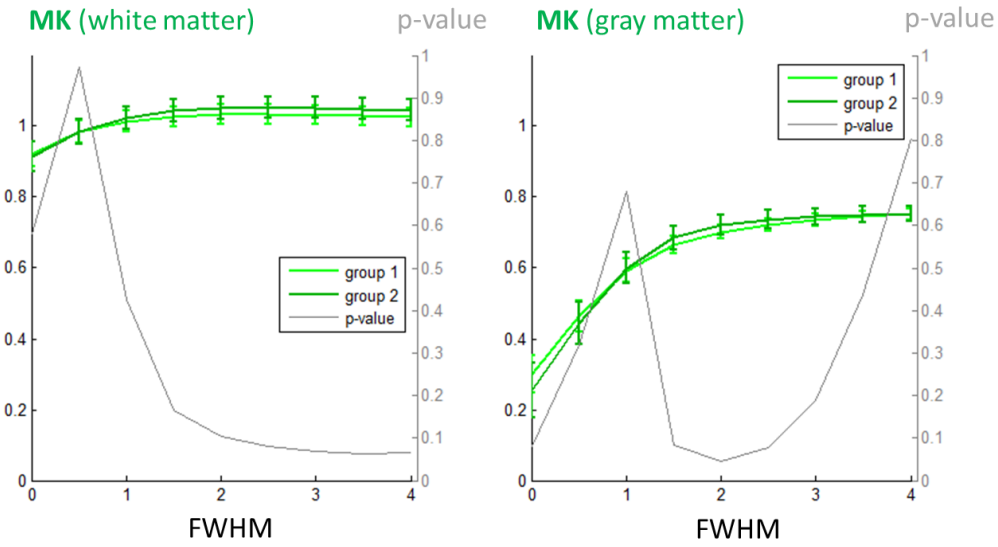

## B.4. Sub-region 6



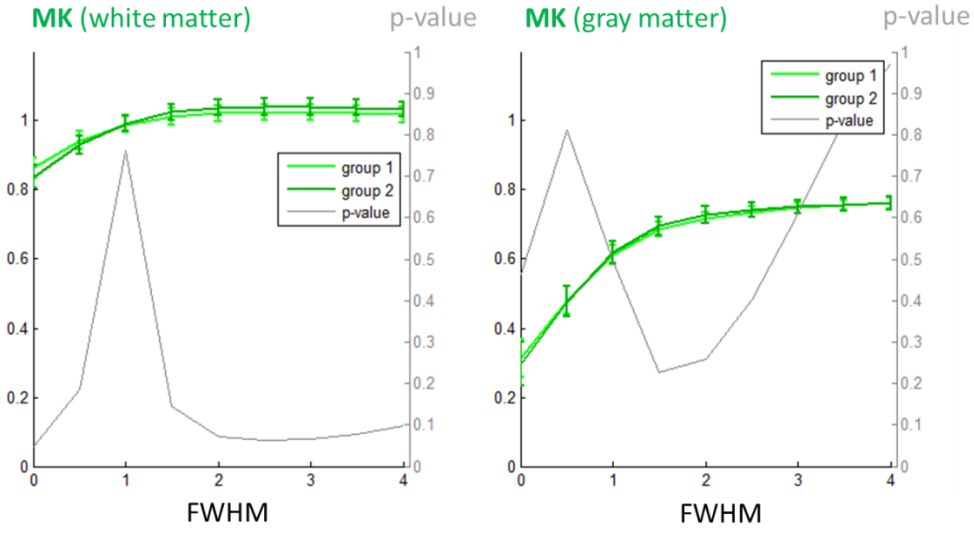

## B.5. Sub-region 7

# Appendix C

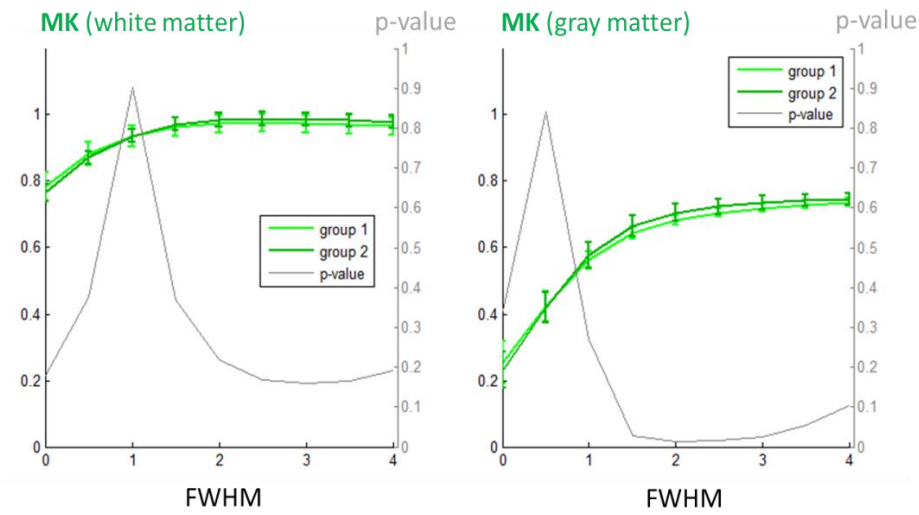
# Age differences of white and grey matter between two ageing groups using DLS



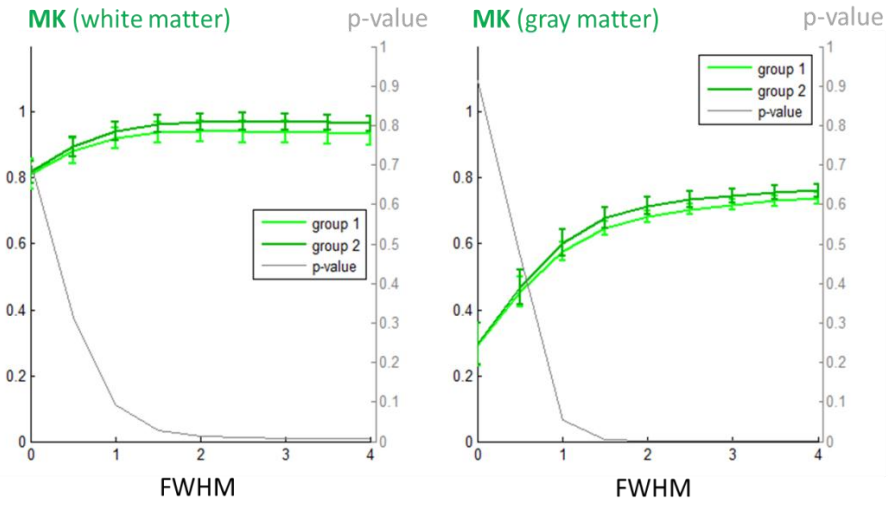

**C.1. Sub-region 3**



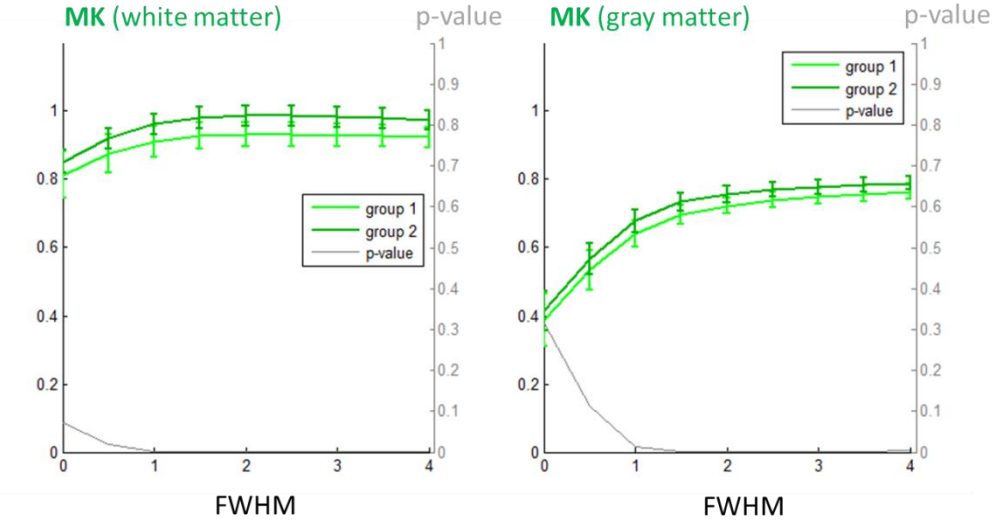

**C.2 .Sub-region 4**




## C.3. Sub-region 5



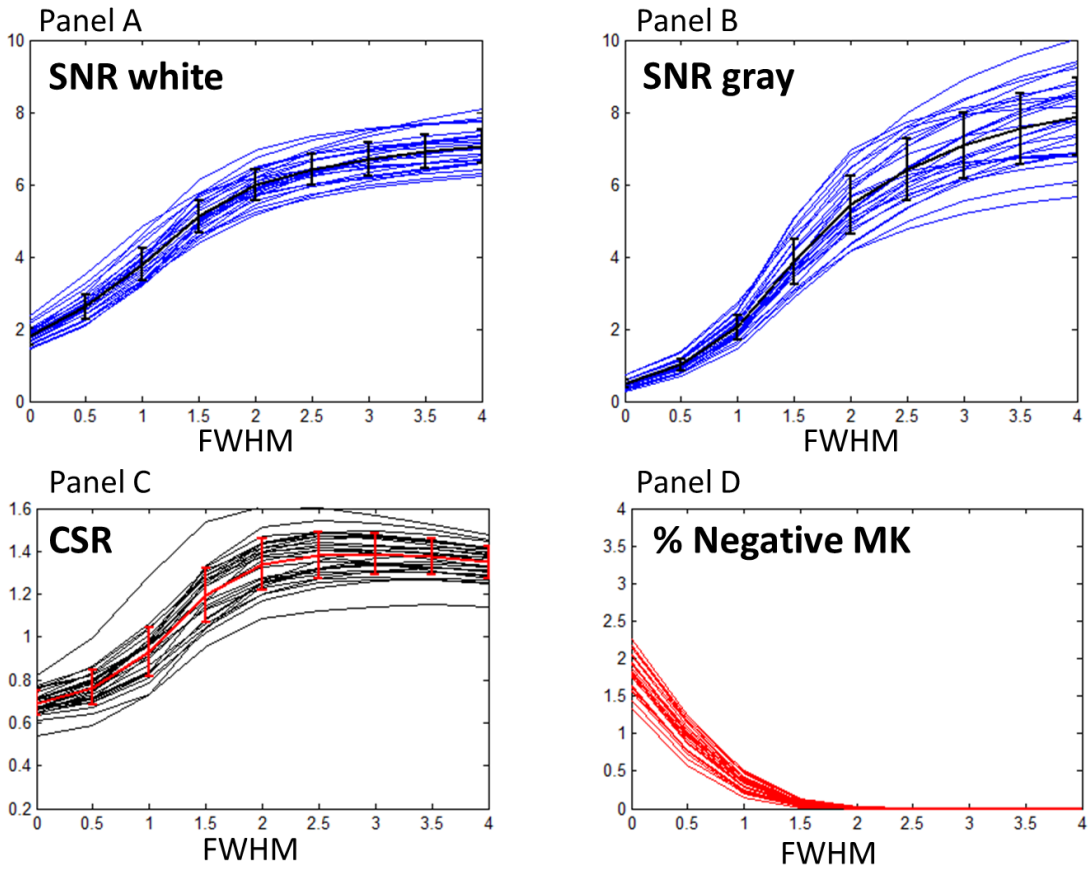

## C.4.Sub-region 6




## C.5.Sub-region 7

# Appendix D

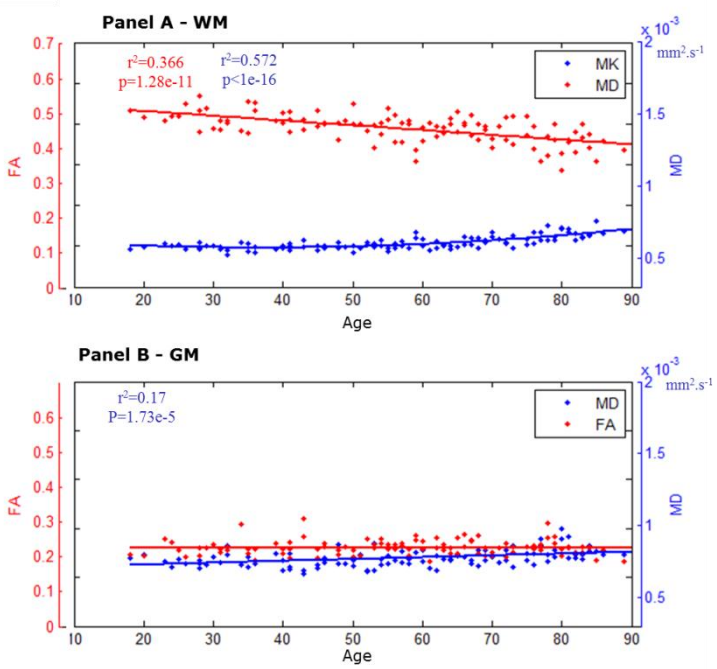
# Image quality metrics



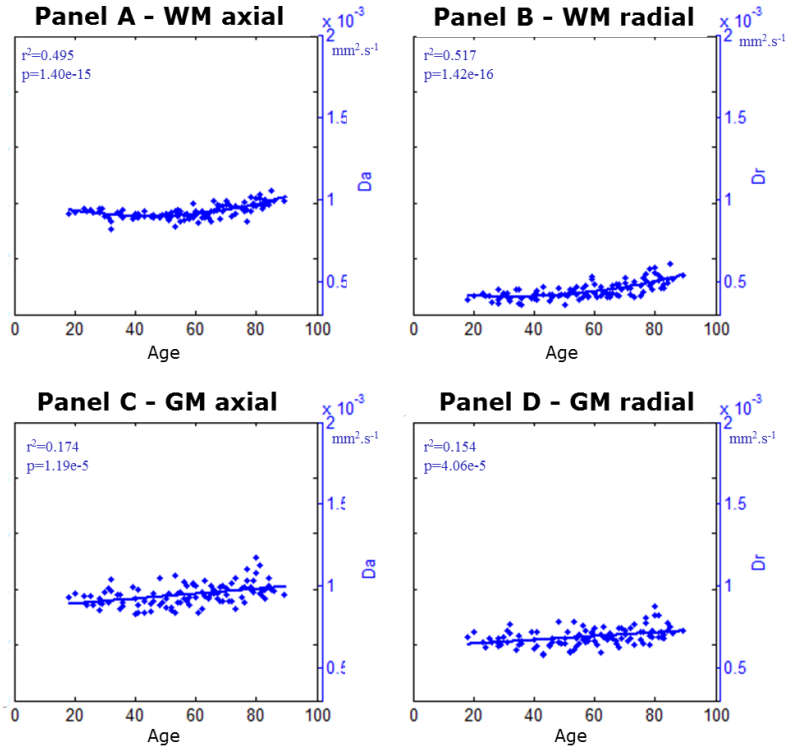

**D.1. DLS Image quality metrics of the neocortex ROI**

# Appendix E

# Longitudinal Changes on Linear DTI




**E.1. Longitudinal changes on MD (blue) and FA (red) extracted from linear DTI**




**E.2. Longitudinal changes on directional diffusivities extracted from linear DTI**