

ESTIMATING CEREBRAL WATER DIFFUSION
METRICS FROM MRI USING DIFFERENT MODEL
ASSUMPTIONS AND SEQUENCE
CONFIGURATIONS

A SIMULATION STUDY

by

Live Wang Jensen

THESIS

for the degree of

MASTER OF SCIENCE



Department of Physics
Faculty of Mathematics and Natural Sciences
University of Oslo

December 2019

Abstract

Diffusion-weighted imaging (DWI or DW-MRI) is a widely applied and clinically important MRI technique probing the diffusion displacement of water molecules in biological tissue on a micrometer length scale. DWI is however an indirect probe, because the extraction of quantitative diffusion metrics requires modeling of the diffusion signal. A plethora of diffusion models have been adapted trying to accurately characterize and quantify the true biological microstructure. Many of the established diffusion models provide the same type of information, with a varying degree of additional information, and there is no gold standard for when to use a given model of a certain complexity.

Two well-established diffusion models of different complexity were investigated in this study; diffusion tensor imaging (DTI) and diffusion kurtosis imaging (DKI). DTI is effective in measuring the dominant direction of water diffusion, however, the model is based upon the assumption that water molecules follow a Gaussian diffusion distribution. Real tissue contains complex cellular structures causing the water molecules to diffuse through highly heterogeneous environments, which leads to a deviation from the Gaussian distribution. DKI is an expansion of the DTI model, including an excess kurtosis term that quantifies the degree of non-Gaussian diffusion. Both models provide information about the standard diffusion parameters; fractional anisotropy (FA), mean diffusion (MD), radial diffusion (RD), axial diffusion (AD). In addition, DKI provides information about mean kurtosis (MK), radial kurtosis (RK) and axial kurtosis (AK).

DWI measures diffusion with the use of magnetic gradients, and the precision of the diffusion parameter estimations depends on the signal-to-noise ratio (SNR), number of signal averages (NSA), number of gradient directions, as well as the degree of diffusion-weighting (b -value) used in the image acquisition.

In this study, DTI and DKI signals simulating white matter, gray matter and cerebrospinal fluid were generated based on DWI signals extracted from real DW-MRI acquisitions in healthy volunteers. Monte Carlo simulations were performed to investigate the effect of SNR, NSA, number of gradient directions and b -values on the parameter estimations of DTI and DKI, with the aim of optimizing the parameter estimation while keeping the acquisition time at a minimum.

The results showed that DTI is more sensitive to noise than DKI in the white matter regions of the brain. In contrast, DKI was more sensitive to noise in the gray matter and cerebrospinal fluid. Increasing the NSA resulted in a general improvement in the parameter estimations for both models. The number of b_0 -images also had a remarkable influence on the parameter estimations. Using 6 b_0 -images instead of 1 b_0 resulted in a noteworthy increase in precision. The analyses showed that the optimal gradient set for DTI was 6 b_0 -images and 30 $b = 1000$ s/mm² measurements, while the optimal gradient set for DKI was 6 b_0 -images, 12 $b = 500$, 30 $b = 1000$ and 50 $b = 3000$ s/mm² measurements.

Acknowledgements

First and foremost, I would like to offer my special thanks to my three supervisors. I would like to thank Atle Bjørnerud for his valuable and constructive suggestions during the planning and development of this thesis, and for providing me with helpful comments on my draft. I would also like to thank Inge Groote for his encouragement and for providing me with the necessary image data for this thesis. I would like to thank Tuva Hope for her guidance through each stage of the process, for always being willing to share her knowledge and immediately answering all my questions, for all the hours spent helping me with the simulations and for the valuable inputs and comments on my draft. Your assistance in this matter has been greatly appreciated.

I would like to thank Eli Olaus for the inspiring and motivational pep-talks and for always keeping her door open.

I also wish to thank my fellow master students and friends. I would like to thank Lise Linn for being an invaluable support throughout the entire process and always being willing to discuss the concepts of diffusion-MRI with me. Thank you, Martine, for providing me with coffee and encouraging words when I needed it the most. I would also like to thank Veljko for the helpful comments on my draft and for taking the time to proofread my thesis.

Finally, I would like to thank my sister and my parents, for their support and encouragement throughout my years of study and through the process of researching and writing this thesis.

Contents

Abstract	iii
List of abbreviations	I
I Introduction	1
1 Introduction	3
1.1 Motivation	4
1.2 Research question and outline	5
II Theory	7
2 Fundamental MR physics	9
2.1 Spin, energy states and magnetic fields	9
2.2 RF excitation	16
2.3 Proton relaxation processes	19
2.3.1 T_1 relaxaton	19
2.3.2 T_2 relaxaton	20
2.4 Image formation	21
2.5 Pulse sequences	24
2.5.1 Image contrast	26
3 Diffusion-weighted MRI	27
3.1 Diffusion	28
3.2 Imaging diffusion	31
3.2.1 Apparent Diffusion Coefficient	33
3.3 Artefacts	34
3.3.1 Motion artefacts	34
3.3.2 Susceptibility artefacts	34
3.3.3 Eddy currents	35
4 Diffusion Tensor Imaging	37

4.1	DTI mathematics	37
4.2	Ordinary least squares	43
4.3	Weighted least squares	44
4.4	Obtaining the diffusion tensor	46
4.5	Diffusion anisotropy in the brain	47
4.6	DTI parameters	48
4.6.1	Mean Diffusivity	49
4.6.2	Axial Diffusivity	50
4.6.3	Radial Diffusivity	50
4.6.4	Fractional Anisotropy	50
5	Diffusion Kurtosis Imaging	53
5.1	Kurtosis	54
5.2	The DKI signal	54
5.3	The diffusion kurtosis tensor	56
5.4	DKI parameters	59
5.4.1	Mean kurtosis	59
5.4.2	Axial kurtosis	59
5.4.3	Radial kurtosis	60
5.4.4	Fractional anisotropy of kurtosis	60
III	Materials and methods	63
6	Experimental data	65
6.1	Extraction of baseline MRI diffusion signals	65
6.2	Diffusion MRI data	66
6.3	Data preprocessing	67
6.4	Slice selection	69
6.5	Segmentation	69
6.6	Voxel selection	73
7	Simulations	75
7.1	Tensor estimation	75
7.2	Signal generation	76
7.3	Establishing ground truth	77
7.4	Adding synthetic noise	77
7.5	Calculation of diffusion parameters	79
7.6	Statistical analysis	80
7.6.1	Outlier removal	80
7.6.2	Relative error	80
7.6.3	Box plots	80
7.7	Noise images	81
7.8	Analytic DTI signal generation	81

8	Simulation experiments	83
8.1	Parameter estimation as a function of SNR	83
8.2	Parameter estimation as a function of NSA	83
8.3	Parameter estimation as a function of N_{grad} and b -values	84
IV	Results	85
9	Results	87
9.1	Signal generation	87
9.1.1	Noise images	89
9.1.2	Analytic DTI signal generation	92
9.2	Parameter estimation as a function of SNR	93
9.2.1	White matter	93
9.2.2	Gray matter	97
9.2.3	CSF	100
9.2.4	MK, RK and AK	103
9.3	Parameter estimation as a function of NSA	104
9.3.1	FA	104
9.3.2	MK	107
9.4	Parameter estimation as a function of N_{grad} and b -values	107
9.4.1	DTI analysis	108
9.4.2	DKI analysis	112
9.5	Summary	116
V	Discussion and conclusion	119
10	Discussion and conclusion	121
10.1	Parameter estimation as a function of SNR	121
10.2	Parameter estimation as a function of NSA	124
10.3	Parameter estimation as a function of N_{grad} and b -values	125
10.3.1	DTI analysis	125
10.3.2	DKI analysis	127
10.4	Study limitations	129
10.5	Conclusion	130
	Bibliography	I
	Appendices	III
A	Additional results	V
A.1	Parameter estimation as a function of NSA	V
A.2	Parameter estimation as a function of N_{grad} and b -values	XI

B	Matlab code	XV
B.1	dki_fit.m	XV
B.2	dti_fit.m	XVIII
B.3	dki_parameters.m	XX
B.4	dti_parameters.m	XXI
B.5	percentile.m	XXI
B.6	sim_exp1_DKI.m	XXIII
B.7	sim_exp1_DTI.m	XXIV
B.8	sim_exp3_DKI.m	XXV

List of abbreviations

- AD** Axial diffusivity
ADC Apparent diffusion coefficient
AK Axial kurtosis
CSF Cerebrospinal fluid
DKI Diffusion kurtosis imaging
DT Diffusion tensor
DTI Diffusion tensor imaging
DWI Diffusion weighted imaging
FA Fractional anisotropy
FID Free induction decay
GM Gray matter
GRE Gradient echo
K Kurtosis
KT Kurtosis tensor
MD Mean diffusivity
MK Mean kurtosis
MRI Magnetic resonance imaging
OLS Ordinary least squares
RD Radial diffusivity
RF Radio frequency
RK Radial kurtosis
ROI Region of interest

SE Spin echo

TE Echo time

TR Repetition time

WLS Weighted least squares

WM White matter

Part I

Introduction

Chapter 1

Introduction

Magnetic resonance imaging (MRI) is a widely applied medical imaging technique used in research and in the clinic as a tool for disease detection, diagnostics and treatment monitoring. The technique produces detailed anatomical pictures of the human body, revealing both structural and functional information, and is especially suitable for imaging of soft tissue. The concept of MRI is based on the discovery made by Felix Block and Edward Purcell, namely that atomic nuclei in possession of spin angular momentum can interact with magnetic fields. They were later awarded with the Nobel Price in Physics in 1952, “for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith”[1]. Subsequently, their discovery lead to the development of the MR-scanner, where the combination of strong magnetic fields, radiofrequency waves and magnetic gradients are used to generate images of the human body. The technique is based on the magnetic properties of the hydrogen nuclei. When placed in a strong magnetic field, the hydrogen nuclei align in the direction of the magnetic field. Radiofrequency waves then stimulate the hydrogen nuclei to precess in an orderly fashion, which results in an induced current in a receiver coil.

The adult human body consists of approximately 60% water[2]. It is here, as part of the water molecules, the vast amount of hydrogen nuclei are located. The water molecules in the tissue constantly move around, both inside and in between cells. The random, thermal movement of water molecules is referred to as *diffusion*. The diffusion pattern of the water molecules is very sensitive to the microstructure of the tissue. Various biological barriers may restrict the diffusion, like cell membranes, organelles and macromolecules. The neural tissue of the brain mainly consists of neurons and glial cells. The axons, which is a part of the neurons, are organized in bundles forming fiber tracts that connect the different parts of the brain. Diffusion is more likely to occur *along* the axons than across. Diffusion measurements might therefore reveal important information about the underlying microstructural characteristics of the tissue. It is possible to measure the diffusion displacement with a special MRI technique called diffusion-weighted MRI (DWI or DW-MRI). DWI

is a widely applied and clinically important imaging technique that measures the diffusion of water molecules on a microscopic length scale[3]. Particularly, the technique offers a method for non-invasively displaying both healthy and pathological anatomy of the brain. DWI has proven to be a unique and powerful tool in both diagnostics and in research.

1.1 Motivation

DWI uses magnetic gradients to measure the diffusion displacement of water molecules. Diffusion of water molecules results in a signal loss, and the greater the diffusion displacement, the greater the signal loss. A magnetic gradient measures the signal loss, and hence the diffusion occurring in *one* specific direction. Because diffusion is direction-dependent, the signal loss must be measured in several non-collinear directions to get the full picture of the diffusion process.

Water diffusion in biological tissue is a complex process, due to multiple tissue compartments with different diffusion properties usually separated by semi-permeable membranes, allowing transfer of water molecules between the different compartments. Parameters describing the diffusion characteristics are estimated by fitting the measured signal to a suitable *model*. However, to account for all possible diffusion processes using DWI techniques is challenging, and practical models make certain assumptions and simplifications to enable implementation with acceptable scan times.

Several DWI-based models have been established. One of the challenges today is that many of the established DWI-based models provide the same type of information, with a varying degree of additional information. One of the most frequently employed diffusion models is diffusion tensor imaging (DTI). The DTI model assumes a mono-exponential signal decay, hence it only takes into account so-called non-restricted diffusion, where water molecules are diffusing freely and unrestricted during the acquisition time[4]. This implies a Gaussian probability distribution of the water diffusion over time.

However, biological tissue contains hindrances that may restrict the diffusion of water molecules. In this case, the assumption of free diffusion is no longer valid, and more complex models may be employed. One such non-Gaussian diffusion model is diffusion kurtosis imaging (DKI). The DKI model is an extension of the DTI model and may provide complementary information about the tissue microstructure[5, 6].

The challenge with using more complex models is that more model parameters are needed, with the potential risk of overfitting in cases where the SNR (signal-to-noise ratio) is low. There is no gold standard for when to use a given model of a certain complexity. In some cases, applying a more complex model results in overfitting and a less complex model would be sufficient. In other cases however, a less com-

plex model may not be able to detect subtle differences in the underlying tissue microstructure that a more complex model would have. More complex models also require larger data sets to be acquired. Diffusion imaging of the brain can be time-consuming, as it requires measurements in many directions. The acquisition time increases with the number of gradient directions, as well as the number of b -values, defined by the strength, duration and spacing of the gradients. It is therefore beneficial to reduce the number of gradient directions and b -values in order to improve the acquisition time. However, reducing these parameters will also reduce information collected from the acquisition. Hence, finding a balance between improving the scan time and preserving enough relevant information to enable proper modeling is of great importance. A short acquisition time in general is preferred, as it may reduce the stress of the patient lying in the scanner and consequent motion induced artefacts, as well as shorten the wait for patients in line. With the ageing population, the demand will only increase. Reducing the acquisition time is both cost-effective and beneficial as it results in more scans per day and earlier diagnoses.

1.2 Research question and outline

The main objective of this thesis is to compare two DWI models which differ in complexity and, based on computer simulations, study the effect of number of gradient directions, b -values and NSA (number of signal averages) on SNR and model complexity. Based on the axiom that the simplest model able to properly describe a given phenomenon should be preferred, the main research question of this thesis is thus:

What effect does the number of gradient directions and b -values have on the estimation of diffusion parameters, and are there any significant differences between the DTI and the DKI model?

The research question is investigated with computer simulations in Matlab by generating synthetic data based on a given DWI model, and simulating the signals from different regions of the brain based on DWI signals extracted from DW-MRI acquisitions in healthy volunteers. A number of simulation experiments using different number of gradient directions and b -values are then executed and the results analyzed by investigating the accuracy and precision of the parameter estimation. Three simulation experiments form the basis of the analysis:

1. Analyzing the parameter estimation as a function of SNR.
2. Analyzing the parameter estimation as a function of NSA.
3. Analyzing the parameter estimation as a function of number of gradient directions and b -values.

The simulations and analyses were performed in Matlab version R2018b, a compre-

hensive numerical computing environment especially suitable for matrix manipulation, plotting of functions and data, and implementation of algorithms.

Finally, the famous Occam's razor principle deserves to be mentioned: "*Entia non sunt multiplicanda praeter necessitatem*" (William of Ockham, 14th century)[7], which roughly translates to "*More things should not be used than are necessary*". The statement was later rephrased by Isaac Newton: "*We are to admit no more causes of natural things other than such as are both true and sufficient to explain their appearances. Therefore, to the same natural effects we must, so far as possible, assign the same causes*"[8].

Part II

Theory

Chapter 2

Fundamental MR physics

The MR scanner uses a combination of magnetic fields and radiofrequency waves to generate images of the inside of the human body. In order to better understand how an image can be formed by these means, it is necessary to take a closer look at how an atomic nucleus behaves under the influence of an external magnetic field.

The following sections are written with the intention of giving the reader a brief introduction to the large and complex field of MRI physics. Unless stated otherwise, the theory in this chapter is collected from *The Physics of Magnetic Resonance Imaging* (A. Bjørnerud)[9], and *Methods in Molecular Biophysics: Structure, Dynamics, Function* (I. N. Serdyuk *et al.*)[10].

2.1 Spin, energy states and magnetic fields

In quantum mechanics, atoms are characterized by their mass, electric charge and spin. The following analogy might be helpful to better understand the behavior of a spinning nucleus. Imagine that the nucleus can be modeled as a spherical body with the nuclear charge distributed uniformly over its surface. Spin angular momentum, or just spin for short, is then often described as the particles' rotation around its own axis. According to this model, a non-spinning nucleus will have a nuclear spin of zero. A spinning nucleus, however, will give rise to a circulating charge which generates a magnetic field, analogous to the field produced when electric current flows through a coil. The induced magnetic field results in a magnetic dipole moment $\boldsymbol{\mu}^1$ oriented along the axis of spin, as illustrated in figure 2.1. This nucleus has a non-zero spin value. The magnetic dipole moment can be compared to a bar magnet with a north and a south pole. Nuclei with a magnetic dipole moment can interact with external magnetic fields, just like when a compass needle interacts with the magnetic field of the Earth.

¹Vectors will be denoted by bold typeface.

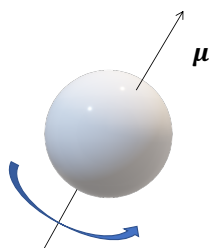


Figure 2.1: Schematic view of a spinning nucleus and its corresponding magnetic dipole moment, μ .

Remember that the image of spin as a nucleus literally spinning around its own axis can be misleading. Spin is an intrinsic physical property, meaning it is a property that the particle has of itself, unrelated to its motion in space, just like a particles' mass or electric charge. Spin is carried by most elementary particles and has both magnitude and direction. It is therefore characterized by a vector \mathbf{I}^2 . Unlike angular momentum in classical physics, the spin angular momentum is quantized. It can only take on discrete values that are integer multiples of \hbar , and all particles with spin follow the quantization condition

$$|\mathbf{I}| = \sqrt{I(I + 1)}\hbar \quad (2.1)$$

where $|\mathbf{I}|$ is magnitude of the spin angular momentum, I is the spin quantum number and \hbar is the reduced Planck constant where $\hbar = \frac{h}{2\pi}$. The spin quantum number I is determined by the number of unpaired protons and neutrons in the nucleus, where each unpaired nucleon contributes $\frac{1}{2}$ to the spin quantum number I . Generally the different elementary particles takes on either integer or half-integer values of I ,

$$I = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots \quad (2.2)$$

In a nucleus with an even atomic mass number and an even nuclear charge, each proton pairs with a proton of opposite sign, and so does the neutrons, leaving zero unpaired nucleons and hence a spin quantum number of zero. In contrast, a nucleus with odd numbers of both protons and neutrons usually has an integer non-zero spin quantum number, due to the even number of unpaired nucleons, which each contributes $\frac{1}{2}$ to the spin quantum number I . Finally, nuclei with an odd number of unpaired nucleons generally have half-integer spin quantum numbers. The proton,

²Although they have exactly the same properties, nuclear spin is commonly denoted \mathbf{I} while electron spin is denoted \mathbf{S} .

electron, and neutron are all examples of elementary particles having $I = \frac{1}{2}$. A number of nuclei of particular interest in structural biology (^1H , ^{15}N , ^{13}C , ^{19}F and ^{31}P) also have a spin quantum value of $\frac{1}{2}$. Keep in mind that it is the value of the spin quantum number I , and not the spin angular momentum $|\mathbf{I}|$ that is referred to when particles with “spin $\frac{1}{2}$ ” are mentioned, because the spin angular momentum $|\mathbf{I}|$ bears the value

$$|\mathbf{I}| = \sqrt{\frac{1}{2} \left(\frac{1}{2} + 1 \right)} \hbar = \frac{\sqrt{3}}{2} \hbar$$

when $I = \frac{1}{2}$. The z -component of the spin angular momentum vector is given as

$$I_z = m\hbar \quad (2.3)$$

where

$$m = -I, -I + 1, \dots, I - 1, I \quad (2.4)$$

is the magnetic quantum number. An atomic nucleus can take on a set of $2I + 1$ magnetic quantum numbers, or energy states, in integral steps between $-I$ and $+I$. Spin $\frac{1}{2}$ nuclei therefore have two energy states related to $m = \pm\frac{1}{2}$, and the corresponding z -component can take on two permitted directions, $I_z = \pm\frac{1}{2}\hbar$.

Nuclei with spin induce a magnetic field with a corresponding magnetic dipole moment $\boldsymbol{\mu}$. The size of $\boldsymbol{\mu}$ is directly proportional to the spin angular momentum \mathbf{I} :

$$\boldsymbol{\mu} = \gamma\mathbf{I} \quad (2.5)$$

where γ is a proportionality constant commonly called the *gyromagnetic ratio*. This constant is unique for each type of nucleus in possession of spin. The magnetic dipole moment is a vector quantity oriented parallel (or antiparallel, for negative γ values) to the spin angular momentum. The ^1H nucleus³ has a gyromagnetic ratio of $\gamma/2\pi = 42.6 \text{ MHz/T}$, which is relatively high compared to other spin $\frac{1}{2}$ nuclei commonly used in NMR. In fact, ^1H has the highest gyromagnetic ratio of all isotopes present in vivo. It is known that the adult human body consists of

³Remember that ^1H is a hydrogen isotope with a single proton as nucleus, meaning that a ^1H nucleus is simply a proton.

approximately 60% water[2]. It is here, in the water molecules, the vast amount of ^1H atoms can be found. The high abundance of ^1H in the human body combined with its relatively large gyromagnetic ratio makes the ^1H nucleus the ideal imaging isotope in clinical MRI. However, a single dipole moment is not large enough to induce a measurable signal. In order to detect a signal, it is necessary to look at a *large sample* of magnetic dipole moments. The sum of all the individual magnetic dipole moments,

$$\mathbf{M} = \sum_{i=1}^n \boldsymbol{\mu}_i \quad (2.6)$$

is referred to as the *macroscopic magnetization*. A sample of living tissue typically contains 10^{22} protons per cm^3 . In the absence of an external magnetic field, the dipole moments are pointing in all possible directions, so the spin's z -axis is oriented arbitrary in space. The net magnetization \mathbf{M} is therefore 0 and not measurable.

A strong magnetic field is able to manipulate the orientation of a spinning nucleus by interacting with its magnetic dipole moment. When placed in an external magnetic field \mathbf{B} , the magnetic dipole moment experiences a torque,

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}. \quad (2.7)$$

This torque causes the spins to line up parallel or antiparallel to the magnetic field (see figure 2.2), just like a tiny compass needle aligns with the magnetic field of the Earth[11].

If the magnetic field \mathbf{B} is defined to be directed along the z -axis, \mathbf{B} can be expressed as

$$\mathbf{B} = B_0 \hat{k} \quad (2.8)$$

where B_0 is the strength of the field and \hat{k} is the unit vector pointing in the z -direction. The magnetic field exerts a torque which not only causes the spins to align but also to *precess*. When the rotational axis of a rotating body is itself rotating about a second axis, the body is said to precess about the second axis. In our case the rotating body is the spinning nucleus, the rotational axis is the magnetic dipole moment, and the second axis is the external magnetic field, as shown in figure 2.3. The torque exerted by the magnetic field causes the spins to precess with an angular frequency of

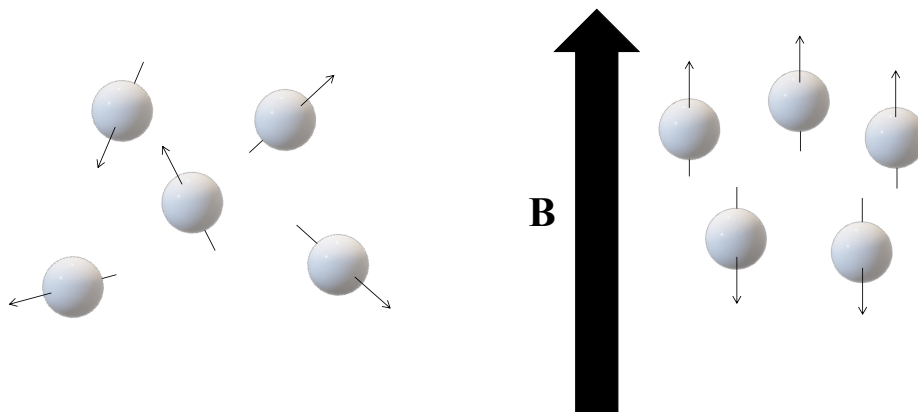


Figure 2.2: Spinning nuclei are oriented in arbitrary directions in the absence of a magnetic field, resulting in a net magnetization of zero. When placed in an external magnetic field however, the spinning nuclei align either parallel or antiparallel with the magnetic field.

$$\omega_0 = \gamma B_0. \quad (2.9)$$

This angular frequency is often referred to as the Larmor frequency or as Larmor precession. The spin's dipole moment precesses about the z -axis of the magnetic field with an angular frequency proportional to the strength of the magnetic field B_0 .

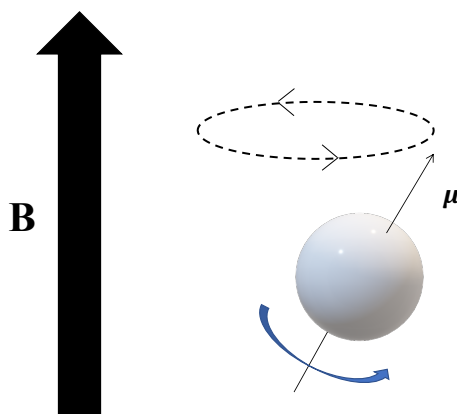


Figure 2.3: When a magnetic field is applied, the magnetic dipole moment of a spinning nucleus starts to precess about the z -axis of the external magnetic field. The precession frequency is proportional to the strength of the magnetic field.

The $2I + 1$ orientations of a spin I nucleus have different energy states. In the

presence of a magnetic field, these energy states are made evident by the Zeeman effect (see figure 2.4). The Zeeman effect is caused by interactions between the magnetic dipole moment of the spins and the external magnetic field. When a proton is placed in a magnetic field, it can only take on certain energy states. These are the *quantum states* of the proton, and each quantum state corresponds to a different energy level. When a proton goes from a higher energy state to a lower energy state, a photon is emitted. The frequency of the emitted photon depends on the strength of the magnetic field. A stronger magnetic field results in a larger energy difference, ΔE , between the energy states[12]. The energy of a proton when placed in a magnetic field is

$$E = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad (2.10)$$

or, when inserting equation (2.8) and (2.5),

$$E = -\mu_z B_0 = -\gamma B_0 I_z = -m\gamma B_0 \hbar \quad (2.11)$$

where $\mu_z = \gamma I_z = \gamma m \hbar$ is the z -component of the magnetic dipole moment. Spin $\frac{1}{2}$ nuclei give rise to two energy states corresponding to $m = -\frac{1}{2}$ and $m = \frac{1}{2}$. This gives us the energies

$$E_{\pm} = \mp \frac{\gamma B_0 \hbar}{2} \quad (2.12)$$

where E_+ is the energy level of the $m = \frac{1}{2}$ state, and E_- is the energy level of the $m = -\frac{1}{2}$ state. If γ is a positive value, which is the case for ^1H nuclei, then the $m = \frac{1}{2}$ state has the lower energy. Traditionally, the energy state with $m = \frac{1}{2}$ is often referred to as “spin up”, while $m = -\frac{1}{2}$ is described as “spin down”[13]. Protons with spin up are oriented parallel to B_0 and are in the lower energy state, while protons with spin down are oriented antiparallel to B_0 and are in a higher energy state. The energy difference between the two states is

$$\Delta E = \gamma B_0 \hbar. \quad (2.13)$$

In thermal equilibrium, there will always be a tiny proportion of spins in the lower energy state. Therefore, ^1H nuclei tend to assume the $m = \frac{1}{2}$ state, which is parallel to the magnetic field. A population of spins can be represented by

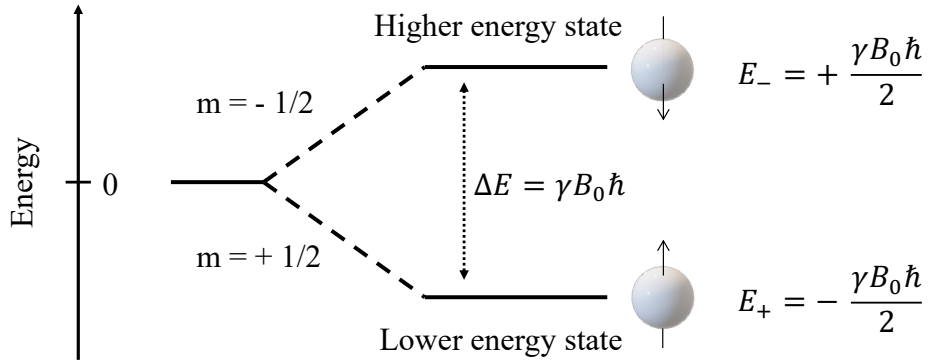


Figure 2.4: The energy states of a spin $\frac{1}{2}$ nucleus. When placed in a magnetic field, the spin can assume one of two energy states. A spin in the lower energy state is oriented parallel to the direction of the magnetic field, while a spin in the higher energy state is oriented antiparallel to the direction of the field.

$$N = N^+ + N^- \quad (2.14)$$

where N is the total number of spins, N^+ is the number of spins in the lower energy state, and N^- the number of spins in the higher energy state. In thermal equilibrium, the distribution of spins in each energy state is governed by the Boltzmann law by the following relationship:

$$\frac{N_m}{N} = \frac{\exp(-E_m/k_B T)}{\sum_{n=-I}^I \exp(-E_n/k_B T)} \quad (2.15)$$

where N_m is the number of spins in energy state m , E_m is the energy of state m , E_n is the energy of state n , k_B is the Boltzmann constant and T is the absolute temperature. The relationship between the two spin populations N^+ and N^- is then

$$\frac{N^+}{N^-} = \exp(\Delta E/k_B T) = \exp(\gamma \hbar B_0/k_B T). \quad (2.16)$$

Equation (2.16) states that there is just a small excess of spins in the lower energy state.

The Larmor precession of the spins in a sample is not coherent. When the magnetic dipole moment vector is decomposed into μ_x , μ_y and μ_z , the components in the positive x - and y -direction are canceled out by components in the opposite, negative x - and y -direction. This results in a net magnetization of zero in the xy -plane, $\mathbf{M}_{xy} = 0$. However, as seen by equation (2.16), there is a tiny excess of spins in the lower energy state, meaning that the positive z -components are not completely canceled out by the negative z -components. This results in a *net magnetization* (illustrated in figure 2.5) in the positive z -direction,

$$\mathbf{M}_z = \frac{N\gamma^2\hbar^2 I(I+1)}{3k_B T} \mathbf{B}_0 \quad (2.17)$$

where N is the total number of spins. This equation is known as Curie's law. It is made evident by equation (2.17) that a stronger magnetic field will result in a larger \mathbf{M}_z , and thus a stronger MR signal. In a magnetic field of 1.5 T at room temperature, only about 1 out of every 10^5 spins will contribute to the macroscopic magnetization. This emphasizes the fact MRI signals are rather weak, which is why strong magnetic fields are used in order to increase \mathbf{M}_z . The former equation also states that $\mathbf{M}_z \propto \gamma^2$. The fact that ^1H exists in great numbers in living tissue and has a large gyromagnetic ratio comes in handy when it comes to signal detection. Nevertheless, since \mathbf{M}_z points in the same direction as the external magnetic field, it is difficult to detect it. In order to detect a signal from \mathbf{M}_z , it has to be flipped an angle α away from the z -axis. This is done by applying a radio frequency (RF) pulse. When \mathbf{M}_z is oriented in another direction than \mathbf{B}_0 , its signal can be detected.

2.2 RF excitation

The macroscopic magnetization vector originates from a large ensemble of spins. A single spin nucleus belongs to the quantum realm and thus follows quantum mechanical laws. It is therefore not possible to predict the behavior of a single spin with the laws of classical mechanics. However, the signal observed in MRI derives from a large population of spins, and hence the magnetization vector obeys classical mechanics. Therefore, the behavior of \mathbf{M} under the influence of \mathbf{B} can be described classically by the Bloch equation as

$$\frac{d\mathbf{M}}{dt} = \gamma(\mathbf{M} \times \mathbf{B}), \quad (2.18)$$

which states that the change in time of the magnetization vector \mathbf{M} is orthogonal to both \mathbf{B} and \mathbf{M} . The magnetization vector \mathbf{M} is therefore rotating, or precessing,

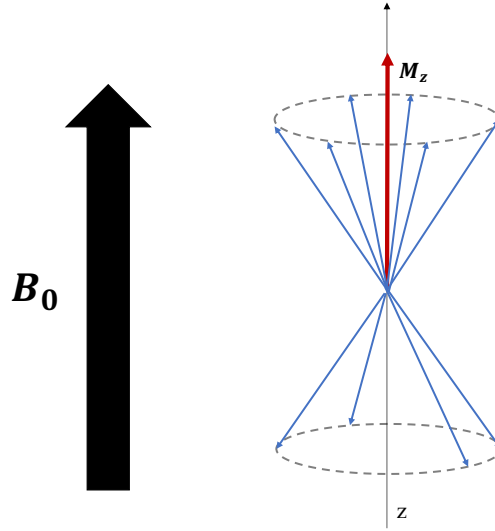


Figure 2.5: The blue arrows represent the magnetic dipole moments, while the red arrow is the net magnetization in the positive z -direction. When protons are under the influence of a magnetic field \mathbf{B}_0 , they align either parallel or antiparallel to the field direction. The protons will also start to precess about the z -axis, as indicated by the dashed, circular line. A small excess of spins assume the lower energy state, parallel to the direction of \mathbf{B}_0 . This results in a net magnetization \mathbf{M}_z in the positive z -direction.

about the direction of \mathbf{B} at the Larmor frequency given in equation (2.9). When the system is in equilibrium, the net magnetization is oriented in the z -direction of the field, so $\mathbf{M} = \mathbf{M}_z$. In order to detect \mathbf{M} , it has to be flipped, or tilted away from its equilibrium orientation. This is done by applying a second magnetic field, \mathbf{B}_1 , perpendicular to the main field \mathbf{B}_0 . The second magnetic field \mathbf{B}_1 is oscillating at the same frequency Ω as the precessing spins. In this way, energy can be exchanged between the spins and \mathbf{B}_1 . The exchange of energy between spins and an oscillating magnetic field is called magnetic resonance. The behavior of \mathbf{M} under the influence of \mathbf{B}_0 and \mathbf{B}_1 can be expressed as

$$\frac{d\mathbf{M}}{dt} = \gamma\mathbf{M} \times (\mathbf{B}_0 + \mathbf{B}_1). \quad (2.19)$$

In order to describe the motion of \mathbf{M} , it is helpful to introduce a new Cartesian coordinate system, (x', y', z') , that rotates around the z -axis of the fixed coordinate system (x, y, z) , with the same angular velocity Ω as \mathbf{B}_1 . These coordinate systems are referred to as the rotating frame and the laboratory frame, respectively. In the rotating frame, both \mathbf{B}_0 and \mathbf{B}_1 are constant, and the effective magnetic field experienced by \mathbf{M} can be expressed as

$$\mathbf{B}_{\text{eff}} = \mathbf{B}_0 + \mathbf{B}_1 - \frac{\Omega}{\gamma} \quad (2.20)$$

where \mathbf{B}_{eff} is the effective magnetic field. The behavior of \mathbf{M} in the rotating frame can then be expressed as

$$\frac{d\mathbf{M}'}{dt} = \gamma \mathbf{M}' \times \mathbf{B}_{\text{eff}}. \quad (2.21)$$

When the \mathbf{B}_1 field oscillates at the same frequency as the spins, $\Omega = \omega_0 = \gamma B_0$. The effective field then becomes $\mathbf{B}_{\text{eff}} = \mathbf{B}_1$, and equation (2.21) is reduced to

$$\frac{d\mathbf{M}'}{dt} = \gamma (\mathbf{M}' \times \mathbf{B}_1). \quad (2.22)$$

According to the rotating frame, \mathbf{M}' rotates about \mathbf{B}_1 with an angular velocity of $\omega_1 = -\gamma B_1$. When \mathbf{B}_1 has a frequency equal to the frequency of the protons, resonance is created and there will be an exchange of energy. The protons absorb energy and are excited from a lower energy state to a higher energy state. The exchange of energy flips the macroscopic magnetization vector away from equilibrium. As long as \mathbf{B}_1 is turned on, \mathbf{M}' moves away from its equilibrium orientation and is rotated an angle $\alpha = \gamma B_1 t$ down towards the xy -plane. This angle of rotation is called the *flip angle*, and the size of it depends on the duration of t and the strength of \mathbf{B}_1 . The oscillating magnetic field is switched on for only a few milliseconds, which is why the application of \mathbf{B}_1 is often referred to as an RF-pulse; it is short and it oscillates in the radio frequency range.

During, and a short time after the application of \mathbf{B}_1 , a component of \mathbf{M} will be present in the xy -plane, \mathbf{M}_{xy} .⁴ Seen from the laboratory frame, this vector component rotates around the z -axis of \mathbf{B}_0 (see figure 2.6). If the sample is surrounded by a coil, the rotation of \mathbf{M}_{xy} will induce a current in the coil. The induced current makes it possible to detect the transverse component \mathbf{M}_{xy} of the magnetization vector, which is the MR-signal. It is not possible to detect a signal from the longitudinal component \mathbf{M}_z , because it does not induce a current in a coil. The strength of the signal is proportional to the size of \mathbf{M}_{xy} . Due to proton relaxation processes, the size of \mathbf{M}_{xy} will decrease over time, and so will the signal intensity.

⁴Other sources may denote the transverse magnetization vector as \mathbf{M}_T or \mathbf{M}_\perp .

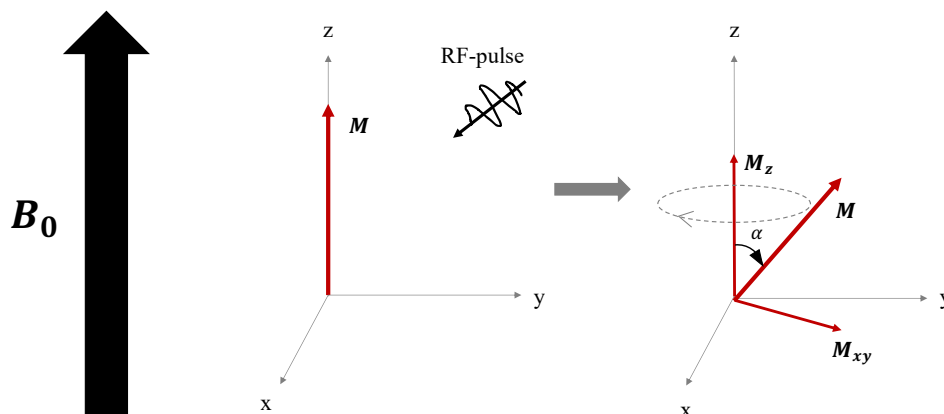


Figure 2.6: The application of an RF-pulse causes the magnetization vector \mathbf{M} to move an angle α from the z -axis, down towards the xy -plane. \mathbf{M} can be decomposed into the longitudinal magnetization vector \mathbf{M}_z and the transverse magnetization vector \mathbf{M}_{xy} . The transverse component rotates around the z -axis and gives rise to a measurable MR-signal.

2.3 Proton relaxation processes

When the RF-pulse is switched off, the strength of the signal rapidly decays due to proton relaxation processes, and \mathbf{M} returns to its equilibrium state. In MRI, the term relaxation is associated with how a signal changes over time. After the RF-pulse, the signal deteriorates by getting weaker and broader with time. The deterioration of the MRI signal can be explained by two separate relaxation processes, T_1 relaxation and T_2 relaxation.

2.3.1 T_1 relaxaton

T_1 relaxation⁵ is related to the recovery of the longitudinal magnetization \mathbf{M}_z . This process is characterized by the relaxation time T_1 . More precisely, T_1 is the time it takes for \mathbf{M}_z to reach $1 - 1/e$, or roughly 63% of its initial value. After the RF-pulse is switched off, \mathbf{M}_z goes back to be in thermodynamic equilibrium with its environment, the lattice. The rate at which \mathbf{M}_z reaches its equilibrium state is dependent on the surrounding lattice. In T_1 relaxation, the excited protons leave their high energy state and enter the low energy state by stimulated emission.⁶ Stimulated emission

⁵Often referred to as longitudinal relaxation or spin-lattice relaxation.

⁶It is also possible for a proton to go from a high energy state to a low energy state by *spontaneous emission*, and as the name suggests, this process occurs at random intervals without regard to the external environment. In MRI, stimulated emission is of most relevance, as spontaneous emission is highly unlikely in the RF range[9].

involves the release of a photon and occurs when the excited protons interact with their external environment at or near the Larmor frequency. The interaction may be induced through the magnetic fields generated by other molecules that are located in the surrounding lattice. As time goes by, more and more of the excited protons will undergo stimulated emission, causing the longitudinal component of the macroscopic magnetization to return to its equilibrium state.

2.3.2 T_2 relaxation

T_2 relaxation⁷ is the process where the transverse component of the macroscopic magnetization vector, \mathbf{M}_{xy} , decays towards its equilibrium state, and is associated with a loss of phase coherence among the spinning protons. When the RF-pulse is switched on, the protons are precessing in an orderly fashion. Due to the RF-pulse, the protons precess in phase, and their magnetic dipole moments can be summed up. \mathbf{M}_{xy} is then at its maximum value.

When the RF-pulse is switched off, the protons will start to lose their phase coherence. Biological tissue is made up of many kinds of atoms and molecules with different local magnetic fields, and individual protons are influenced by the local magnetic fields from neighboring nuclei. As illustrated by the Larmor equation (2.9), the precession frequency depends on the strength of the magnetic field sensed by the proton. When there are local variations in the magnetic field, the protons are affected by the local field, \mathbf{B}_{loc} as well as the external field, \mathbf{B}_0 . As a consequence, different protons start to precess at different frequencies. This results in a loss of phase coherence, and \mathbf{M}_{xy} starts to decay. The rate at which \mathbf{M}_{xy} decays is characterized by the T_2 relaxation time, which denotes when \mathbf{M}_{xy} has decayed to $1/e$, or 37% of its maximum value. However, there are not only local variations in the magnetic field that causes the protons to dephase. The external magnetic field \mathbf{B}_0 is not completely homogeneous, it has minor irregularities in magnetic field strength. The inhomogeneities of \mathbf{B}_0 causes the protons to precess at different frequencies and lose phase coherence. The relaxation time T_2^* takes the irregularities of \mathbf{B}_0 into account. It is the T_2^* relaxation that is measured when the signal decays. This is referred to as the free induction decay, or FID for short. T_2^* includes both local inhomogeneities and \mathbf{B}_0 inhomogeneities, and is usually much shorter than T_2 ,

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma\Delta B_0 \quad (2.23)$$

where ΔB_0 is the difference in field strength of the locally varying field. Their relation shows that the T_2^* relaxation time always is shorter than the T_2 relaxation time. Another relation that also holds true in all situations is that $T_2 \leq T_1$. It

⁷Also called transverse relaxation or spin-spin relaxation.

is important to notice that T_1 and T_2 relaxation are independent processes, even though their relaxation times are related by the previous expression.

The effect of the relaxation processes can be included in the Bloch equation,

$$\frac{d\mathbf{M}}{dt} = \gamma\mathbf{M} \times \mathbf{B}_{\text{eff}} + \frac{M_{xy}}{T_2} - \frac{M_z - M_0}{T_1} \quad (2.24)$$

where M_0 is the equilibrium magnetization. The effects of the relaxation processes on the magnetization components can be obtained by solving the differential equation above. The solution yields

$$M_z(t) = M_0 \left[1 - e^{-\frac{t}{T_1}} \right] + M_z(0)e^{-\frac{t}{T_1}} \quad (2.25)$$

$$M_{xy}(t) = M_{xy}(0)e^{-\frac{t}{T_2}}. \quad (2.26)$$

Equations (2.25) and (2.26) make it possible to predict the signal behavior of both the longitudinal and transverse component of the magnetization vector following an RF-pulse (illustrated in figure 2.7).

2.4 Image formation

The RF-pulse excites the protons in the sample and tilts the macroscopic magnetization vector down into the xy -plane. In clinical use however, it is preferable to only look at a *specific* part of the sample (read: patient), like a slice through the head or the abdomen. In order to excite only the protons located in a specific slice, a *slice selective magnetic gradient* is applied.

A magnetic gradient is a magnetic field with position-dependent field strength and is measured in mT/m. A slice selective gradient is, in other words, a magnetic field that causes the protons to precess with a position-dependent frequency. The Larmor frequency of the protons now becomes a function of position along this gradient. The slice selective gradient is commonly referred to as $\mathbf{G}_s(t)$. It is possible to place $\mathbf{G}_s(t)$ along any arbitrary direction in order to select the slice of interest. However, it is common to place the slice selective gradient along the z -axis, and in that case, $\mathbf{G}_s = \mathbf{G}_z$. The z -axis is usually oriented in the same direction as the patient when the patient is lying down inside the MR-scanner, from feet to head. This creates a slice through the transverse plane (also known as an axial or horizontal plane) of the patients' body. Protons at different positions along the z -axis will have different

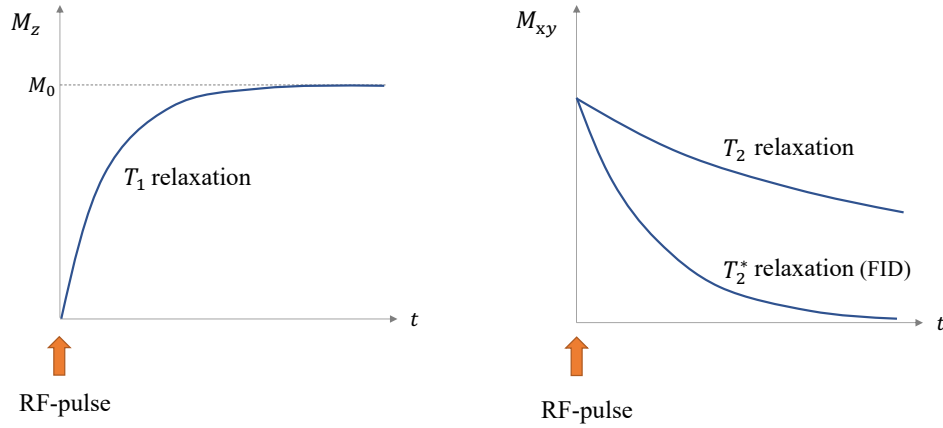


Figure 2.7: T_1 , T_2 and T_2^* relaxation processes. The RF-pulse is applied at $t = 0$. In T_1 relaxation, excited protons leave their high energy state (spin down) and enter the low energy state (spin up), leading to the recovery of the longitudinal magnetization \mathbf{M}_z . \mathbf{M}_0 is the size of \mathbf{M}_z in thermodynamic equilibrium. In T_2 relaxation, local variations in the magnetic field result in a loss of phase coherence among the protons, leading to the decay of the transverse magnetization \mathbf{M}_{xy} . T_2^* relaxation include local variations in the magnetic field as well as inhomogeneities of the external magnetic field, leading to a much more rapid decay of \mathbf{M}_{xy} .

resonance frequencies, while protons at different positions in the xy -plane will have the same resonance frequency. The effective field strength along the z -axis is now

$$\mathbf{B}_z(t) = \mathbf{B}_0 + \mathbf{G}_s(t) \cdot \mathbf{r} \quad (2.27)$$

where \mathbf{r} is the position vector. The effective field strength can now be written as

$$\mathbf{B}_{\text{eff}}(t) = \mathbf{B}_0 + \mathbf{G}_s(t) \cdot \mathbf{r} + \mathbf{B}_1 + \frac{\Omega}{\gamma}. \quad (2.28)$$

The resonance frequency of the protons depends on the magnetic field they experience. The varying magnetic field makes it possible to select and excite the protons located in a desired slice along the z -axis. Only the protons with the same resonance frequency as the RF-pulse will be excited. The position of the slice can be adjusted by adjusting the frequency Ω of the RF-pulse, or by adjusting the field strength of the magnetic gradient.

So far, the assumption has been that the RF-pulse consists of only a single frequency, Ω . In practice, however, the RF-pulse contains a range of frequencies within a small

interval. This is the bandwidth $\Delta\omega$ of the RF-pulse. The result of this is that the RF-pulse excites protons located in a slice with a thickness Δz along the z -axis,

$$\Delta z = \frac{\Delta\omega}{\gamma G_z}. \quad (2.29)$$

According to (2.29), the thickness of the slice can be adjusted by either increasing or decreasing the bandwidth of the RF-pulse, or by adjusting the gradient G_z . If a thinner slice is desired, the bandwidth can be made smaller by increasing the duration of the RF-pulse. The cost of this is a longer scan-time.

The slice selective gradient makes it possible to select a slice within the volume at hand. All the protons in the slice are precessing with the same frequency and giving off signals. In order to encode the signals, two new magnetic field gradients are introduced; the *phase encoding gradient* and the *frequency encoding gradient*. The phase encoding gradient G_y is usually placed along the y -axis, while the frequency encoding gradient G_x is placed along the x -axis (see figure 2.8). The phase encoding gradient gives the protons a position-dependent phase shift along the y -axis, and the frequency encoding gradient gives the protons a position-dependent frequency along the x -axis, making the frequency a function of x .

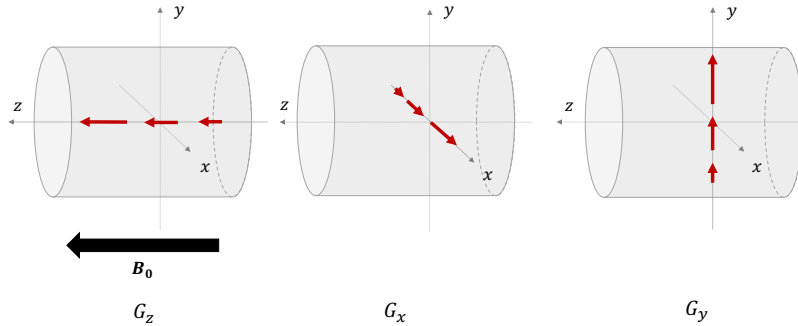


Figure 2.8: Magnetic field gradients inside an MR-scanner. The red arrows are normal vectors, and their size indicates the magnetic field strength in a plane perpendicular to the vector. The slice selective gradient G_z is directed along the z -axis and gives the protons a position-dependent frequency along the z -axis. The frequency encoding gradient G_x gives the protons a position-dependent frequency along the x -axis. The phase encoding gradient G_y gives the protons a position-dependent phase shift along the y -axis. Adapted from[14].

The slice can be divided into smaller subvolumes called voxels, where each voxel contains protons with a unique combination of frequency and phase. During an MR scan, all the frequency and phase information is collected into the Fourier domain

called k-space. The k-space represents the spatial frequency distribution of the MR image, and is the Fourier transform of the MR image. By simply doing an inverse Fourier transform of the k-space, the MR image is obtained.⁸

2.5 Pulse sequences

The spatial frequency distribution collected in k-space forms the basis for the quality of the resulting MR image. It is necessary to sample the k-space representation of the object evenly in order to get a good quality image. There are several different ways to properly sample information in k-space, depending on how the gradients and RF-pulses are applied. A *pulse sequence* is when gradients and RF-pulses are applied in a certain sequence and with certain time intervals. Two of the most common pulse sequences are the gradient echo (GRE) sequence and the spin echo (SE) sequence.

In diffusion weighted imaging, the technique in focus in this thesis, the SE sequence is most commonly used. In the SE sequence, a 90° RF-pulse first flips the magnetization vector towards the xy -plane. The protons will start to dephase under the influence of the frequency encoding gradient G_x , which results in a signal loss. A 180° RF-pulse then flips the proton spins 180° about the xy -plane. This inverts the phase of the protons. The dephasing effect of G_x will then be reverted, and a signal peak, called a spin echo, is observed after a time TE (echo time). The process of dephasing and rephasing spins is illustrated in figure 2.9. The 180° pulse also reverts the dephasing effect caused by inhomogeneities in the main magnetic field. The signal therefore decays with T_2 , and not with T_2^* . The time interval between two 90° pulses is called the repetition time (TR).

The GRE sequence uses a frequency encoding gradient G'_x of opposite polarity instead of a 180° pulse to rephase the spins. This results in a gradient echo. Note that a bipolar frequency encoding gradient does not revert the dephasing effect due to field inhomogeneities. The signal therefore decays with T_2^* in a GRE sequence.

The process of rephasing the spins by either gradient reversal or a 180° RF-pulse can be illustrated by analogy to a running competition. The fast runners represent spins with a high precessing frequency, while the slow runners represent spins with a slower precessing frequency. After a time t , the fast runners have traveled a greater distance than the slower ones. The fast precessing spins will accumulate a bigger phase shift than the slower ones. After a time $t = \text{TE}/2$, a flute is blown, and the competitors have to turn around and run back. The flute represents the 180° pulse or the change in polarity of G_x . Assuming that they all run at a constant speed throughout the race, the fast runner will eventually reach the slow runner, and the

⁸For a thorough discussion of k-space, please seek out Atle Bjørnerud's *The Physics of Magnetic Resonance Imaging*[9].

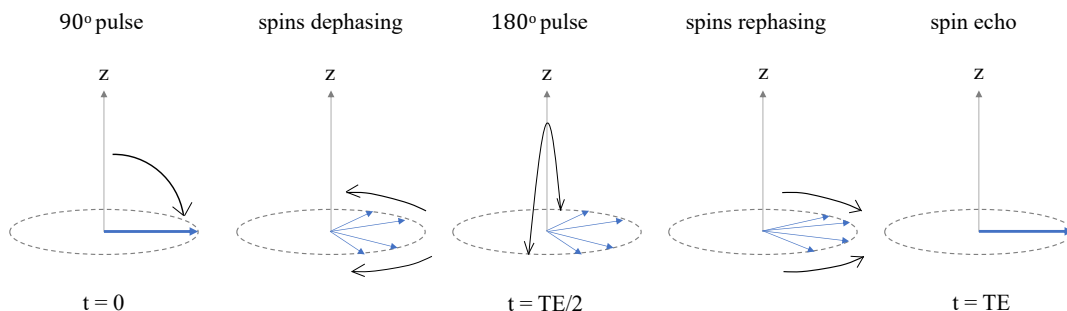


Figure 2.9: Principle of dephasing and rephasing of spins in the SE sequence. A 90° pulse flips the magnetization vector into the xy -plane. The spins will start to dephase with time due to relaxation processes. A 180° pulse the flips the spins 180° about the xy -plane, inverting the phase-shift. The spins start to rephase, and the refocusing of the spins generates a spin echo with a peak amplitude occurring at the echo time, TE .

net phase shift among the spins gets smaller and smaller. A signal echo starts to form. In the end, they will all arrive at the finish line at the same time $t = TE$. The peak amplitude of the echo occurs at the finish line when the phase shift among the spins is zero.

The conventional SE and GRE sequence reads out a single line in k-space per TR interval. This is a time demanding process, especially when larger volumes are imaged. The acquisition time can be significantly reduced by a so-called echo-planar imaging (EPI) sequence. In an EPI sequence, multiple lines in k-space are acquired per TR interval. By applying short ‘blips’ of the phase encoding gradient, a train of echoes is generated following the 90° excitation pulse. The EPI sequence provides faster acquisition time and reduces the effects of patient motion, but the downside is an extreme sensitivity to field inhomogeneities and a general sensitivity to many artefacts[9].

An EPI sequence can be combined with a GRE sequence or a SE sequence. In a GRE-EPI sequence, the spins are rephased between each echo by applying a frequency encoding gradient of opposite polarity. A SE-EPI sequence uses a 180° pulse between each echo to rephase the spins[15].

The most common sequence in DWI is the single shot SE-EPI sequence. In a single shot EPI sequence, data from all of k-space is acquired following a single 90° excitation pulse. The amplitudes of the last echoes in the echo train are therefore going to be very low, thus a single shot sequence requires large system stability and field homogeneity.

2.5.1 Image contrast

By careful manipulation of the echo time TE, and the repetition time TR, is it possible to achieve different types of image contrasts. Different biological tissues have different T_1 -and T_2 time constants, and this difference is what gives the final image its contrast.

The differences in T_1 can be demonstrated in a T_1 -weighted image. A T_1 -weighted image can be achieved by using a short TR and a short TE, where $TR \approx T_1$. Because the signal intensity decays with T_2 in a spin echo sequence, a short TE will minimize the T_2 -effects. A short TR highlights the differences in T_1 recovery times among the tissue types. It is the signal from the \mathbf{M}_z component of the magnetization vector that is measured in a T_1 -weighted image. When the RF-pulse is switched off, \mathbf{M}_z starts to recover, and after a time t , tissue types with a short T_1 will have a larger \mathbf{M}_z than tissue types with a long T_1 . The signal intensity is proportional to the size of \mathbf{M}_z . Tissues with short T_1 , like fat, therefore appears bright in a T_1 -weighted image, while tissues with longer T_1 , like the cerebrospinal fluid, appear darker.

If the time between consecutive 90° pulses is too long (when TR is longer than T_1 of most tissue types), \mathbf{M}_z will completely recover and the differences in T_1 -times are minimized. Therefore, a T_2 -weighted image has a long TR and TE in order to minimize the T_1 effects and emphasize the T_2 effects. The T_2 relaxation time is connected to time it takes for \mathbf{M}_{xy} to decay. Tissues with a long T_2 appear bright in a T_2 -weighted image, while tissues with a short T_2 appear dark. However, the T_1 -effects can not be completely eliminated, even with a long TR, because some tissue types with a long T_1 are still able to create some T_1 -weighting in the image.

Chapter 3

Diffusion-weighted MRI

Diffusion is the random movement of molecules in a fluid or gas. Biological tissue contains large amounts of water molecules. The water molecules do not stay at the same place inside the tissue, but tend to move about in a random manner due to what is called Brownian motion. If the diffusion is free and unrestricted, the diffusion of molecules is equal in all directions. This is referred to as *isotropic* diffusion. Biological tissues often contain barriers, such as macromolecules, membranes and nerve fibers. When diffusion is directionally dependent, the diffusion is said to be *anisotropic*. Certain tissue types, like muscle and brain, consist of coherently oriented fibers, and the water molecules tend to diffuse along these fibers[14]. Hence, the degree of diffusion depends on the direction of the measurement, and the diffusion pattern may reveal the underlying structure of the tissue. One clear example is the anatomy of the white matter in the brain. The structure of white matter is highly complicated and not always understood. Measurement of the diffusion anisotropy makes it possible to uncover the intricate fiber structures of the brain.

Diffusion-weighted magnetic resonance imaging (DWI or DW-MRI) is an MRI technique that measures the diffusion of water molecules by applying diffusion sensitizing gradients. A diffusion-weighted image can then be generated based on the degree of diffusion in each voxel. Diffusion-weighted MRI has several advantages. It maps the diffusion process in vivo non-invasively and without the help of contrast agents. The diffusion can be measured along any arbitrary axis. A diffusion-weighted image gives access to new information about the underlying tissue structure that can not be provided by conventional MRI. It is known that a number of pathological conditions may alter the morphology of the tissue, and a diffusion-weighted image makes it possible to visualize and hopefully detect these changes.

The theory in this chapter is collected from *The Physics of Magnetic Resonance Imaging* (A. Bjørnerud)[9], *Introduction to Diffusion Tensor Imaging* (S. Mori)[14] and *Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors* (P. B. Kingsley)[16], unless otherwise specified.

3.1 Diffusion

Molecular diffusion, or simply just diffusion, is the movement of molecules in a gas or fluid from a region of high concentration to a region of lower concentration. The change in concentration from one region to another is called a concentration gradient. The diffusion process leads to a net flux of molecules along the concentration gradient, and the result is a gradual mixing of the different molecules. The gradual mixing eventually leads to a homogeneous distribution of molecules in the mixture. To exemplify: Imagine that a drop of ink is released into a glass of water. Even though the water is completely still, the ink molecules will start to spread out and mix with the water molecules. After some time, the ink molecules will be completely mixed with the water molecules. The distribution of water and ink has become homogeneous, due to diffusion. Even after the mixture has become homogeneous and the concentration gradient is no longer present, the molecules continue to move about randomly.

The random movement of particles suspended in a fluid is known as Brownian motion. The phenomenon is named after Robert Brown, who in 1827 observed the continuous jittery motion of pollen grains suspended in water. The random movement of a particle in a homogeneous fluid is caused by a large number of collisions between the particle and the fast-moving molecules of the fluid. The collisions happen at random and cause the particle to move in arbitrary directions.

The driving force in diffusion is the concentration gradient, and molecules are transported along this gradient by random collision motions. The diffusion process is slower in fluids than in gases, because the density of molecules is higher in a fluid, resulting in more collisions. The process of diffusion and Brownian motion is illustrated in figure 3.1.

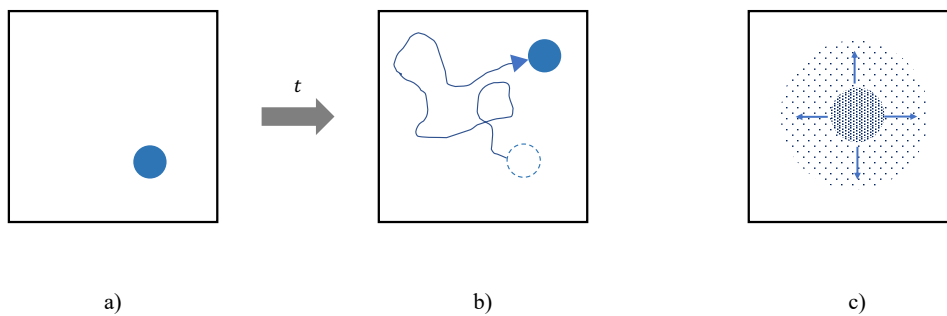


Figure 3.1: a) A particle suspended in a fluid at time t_0 . No concentration gradient is present. b) After a time t , the particle has moved around due to Brownian motion. c) A group of particles suspended in a fluid. The concentration gradient is equal in all directions and the diffusion is isotropic. Adapted from[14].

The diffusion flux is the amount of particles per unit area per unit time, and can in one dimension be described by Fick's first law of diffusion,

$$J = -D \frac{\partial \phi}{\partial x} \quad (3.1)$$

where J is the diffusion flux ($\text{mol m}^{-2} \text{s}^{-1}$), D is the diffusion coefficient or diffusivity (m^2/s), ϕ is the particle concentration per unit volume (mol/liter) and x is the position. $\frac{\partial \phi}{\partial x}$ is the change in concentration with respect to x , and is the driving force of diffusion. The negative sign (-) indicates that the diffusion flux goes from regions of high concentration to regions of low concentration. The diffusion coefficient D can be expressed as

$$D = \frac{k_B T}{6\pi\eta R} \quad (3.2)$$

where k_B is the Boltzmann constant, T is the absolute temperature measured in Kelvin, η is the viscosity and R is the particle radius. In the three-dimensional case, a small modification of equation (3.1) is made, and Fick's first law can be expressed as

$$J = -D \nabla \phi \quad (3.3)$$

where $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$ is the del or gradient operator in \mathfrak{R}^3 . Fick's first law does not take into account that the concentration gradient changes in time. Fick's second law of diffusion was therefore developed, and can in three dimensions be expressed as

$$\frac{\partial \phi}{\partial t} = D \Delta \phi \quad (3.4)$$

where $\Delta = \nabla^2$ is the Laplace operator, which generalizes the second derivative. If the diffusion is unrestricted and one-dimensional, equation (3.4) has the solution

$$\phi(x, t) \propto \frac{1}{\sqrt{Dt}} \exp\left(\frac{-x^2}{4Dt}\right). \quad (3.5)$$

Equation (3.5) shows that the particle concentration at a given position x after a certain time t is given by a Gaussian distribution. A particle that is diffusing freely, will move a certain distance from the starting point after a given time t . The diffusion probability $P(x, t)$ that a particle is located at a position x after a time t is equivalent to the concentration distribution, giving

$$P(x, t) \propto \frac{1}{\sqrt{Dt}} \exp\left(\frac{-x^2}{4Dt}\right). \quad (3.6)$$

The mean square displacement $\langle r^2(t) \rangle$ of a particle, as derived by Albert Einstein in 1905, is given by

$$\langle r^2 \rangle = q_i Dt \quad (3.7)$$

where $\langle r^2 \rangle$ is the mean square displacement, and r is the mean distance from the starting point that a particle will have diffused after time t . q_i is a numerical constant which depends on dimensionality ($q_i = 2, 4$, or 6 for 1, 2 or 3 dimensional diffusion, respectively), and D is the diffusion coefficient. In three-dimensional space, the mean-square displacement is

$$\langle r^2 \rangle = x^2 + y^2 + z^2 = 6Dt. \quad (3.8)$$

When the diffusion is isotropic, the mean diffusion distance is the same in all directions. This can be illustrated by a sphere, where the diffusion probability is highest at the origin and lower closer to the bounding circle of the sphere. The radius of the sphere is given by the mean diffusion distance, which can be found by taking the square root of equation (3.8)

$$r = \sqrt{6Dt}. \quad (3.9)$$

The typical water diffusion distance during MR measurements is approximately 5-10 μm , while the voxel size is usually around 2 mm.

3.2 Imaging diffusion

Diffusion can be measured by introducing a diffusion sensitizing gradient \mathbf{G} ¹. Diffusion gradients can be added to any type of pulse sequence, but the majority of diffusion-weighted images are done with an SE or SE-EPI sequence. A pulse sequence where a pair of diffusion sensitizing gradients are added is called a Stejskal-Tanner sequence. Figure 3.2 illustrates the diffusion sensitization of a SE sequence. The diffusion gradient \mathbf{G} is applied before and after the 180° refocusing pulse. The gradient induces a position-dependent phase shift, $\gamma\delta Gr$, to the protons along the gradient direction, where δ is the duration of the gradient, G is the gradient strength, and r is the position of the proton. The protons are now tagged with a certain phase shift. After a time $t = \tau$, the 180° refocusing pulse is applied. The second diffusion gradient then reverses the phase shift with $-\gamma\delta Gr'$, where r' is the position of the proton at the time of the second gradient application. The net phase shift of the proton is therefore

$$\Delta\phi = \gamma\delta G(r - r') = \gamma\delta GR \quad (3.10)$$

where R is the net displacement of the proton. If the proton has remained stationary ($r = r'$) during the gradient separation time Δ , the net phase shift is zero. Protons that have diffused a distance ($r \neq r'$) have a non-zero phase shift which is proportional to the diffusion distance R . These protons have a different phase than the neighboring protons, and they will not realign perfectly during the rephasing. The signal from the diffused protons is lost, as made evident by the attenuated signal echo amplitude. The diffusion gradient pair thus induces a signal loss due to the diffusion of water molecules. The signal attenuation is proportional to the strength G and duration δ of the diffusion gradients, as well as the diffusivity D . The larger the diffusion distance, the more the signal is attenuated (assuming the diffusion took place in the gradient direction).

A single voxel contains a large number of water molecules with completely random diffusion paths. The diffusion gradient can only measure diffusion along a single, predetermined axis, and diffusion occurring in other directions will not be registered by the gradient. In anisotropic diffusion, the degree of diffusion varies with the direction. Diffusion gradients are therefore applied in several noncollinear directions to get the full picture of the diffusion process.

Diffusion only affects the transverse component of the magnetization vector, which can be written as $M_T = M_x + jM_y$ in complex notation. The behavior of M_T under both relaxation and diffusion is described by the Bloch-Torrey equation,

¹Not to be confused with the slice selecting gradient G_z or the phase- and frequency encoding gradients G_y and G_x .

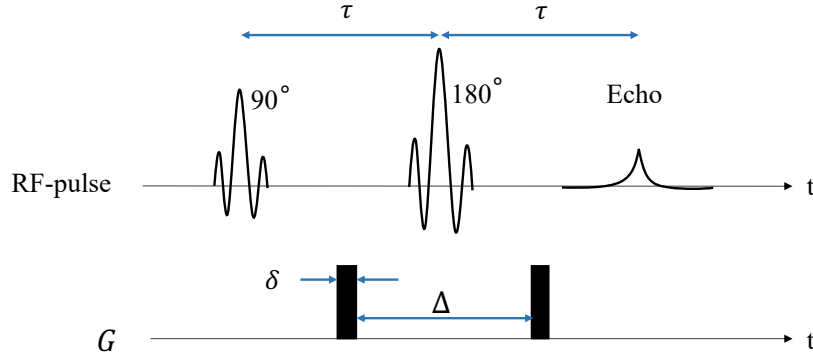


Figure 3.2: Illustration of a Stejskal-Tanner sequence. Here diffusion gradients have been added to a SE sequence with echo time $TE = 2\tau$. The diffusion gradients induce a measurable signal loss which is proportional to the degree of diffusion. The duration of G is given by δ and the gradient separation time is Δ . The gradient pair have the same polarity due to the 180° refocusing pulse.

$$\frac{dM_T}{dt} = -\frac{M_T}{T_2} - j\gamma(\mathbf{G}_s(t) \cdot \mathbf{r})M_T + D\Delta M_T \quad (3.11)$$

where Δ is the Laplace operator. The last term in equation (3.11) is the diffusion term, which describes the behavior of M_T under the effects of diffusion. The diffusion term can be compared to Fick's second law in equation (3.4). Fick's second law, $\frac{\partial \phi}{\partial t} = D\Delta\phi$, describes the temporal change in particle concentration as a result of particle diffusion along the direction of the concentration gradient. The diffusion term, $D\Delta M_T$, describes the temporal change in M_T as a result of water diffusion along the direction of the diffusion gradient. Diffusion that occurs along the gradient direction causes decay in M_T , resulting in a signal loss. When diffusion is isotropic, the signal intensity can be expressed as

$$S(b) = S_0 \exp(-bD) \quad (3.12)$$

where S_0 is the signal intensity without diffusion gradients, and

$$b = \gamma^2 G^2 [\delta^2 (\Delta - \delta/3) + \epsilon^2/30 - \delta\epsilon^2/6] \quad (3.13)$$

is the b -factor (s/m²) or b -value, which determines the amount of diffusion weighting in the sequence. ϵ is the rise and fall time of the gradient. If the rise and fall time is neglected ($\epsilon \approx 0$), the result is perfectly rectangular pulses. The b -factor is then reduced to

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3). \quad (3.14)$$

The signal intensity S decays mono-exponentially with the diffusivity D and the amount of diffusion weighting, b . A high b -factor gives the sequence a high diffusion weighting, whereas a low b -factor gives low diffusion weighting. The amount of diffusion weighting is determined by the gradient strength G , the gradient separation time Δ , and the duration time δ of the gradients. S_0 is the signal intensity without any diffusion sensitizing gradients, and only decays with T_2 . The S_0 signal forms a T_2 -weighted image, and because no diffusion weighting is applied ($b = 0$), it is often called the b_0 image.

3.2.1 Apparent Diffusion Coefficient

The diffusion coefficient D of a voxel can be estimated after the signal intensity S is obtained. The estimated diffusivity D_{app} is called the apparent diffusion coefficient (ADC), and is not the true diffusivity of the tissue, but the observed diffusivity. The apparent diffusion coefficient can be found by taking signal measurements at two different b -factors, a low b -factor b_1 and a higher b -factor b_2 . When all other pulse sequence parameters are held constant, the two signals can be expressed as

$$S_1 = S_0 \exp(-b_1 D_{app}) \quad (3.15)$$

and

$$S_2 = S_0 \exp(-b_2 D_{app}). \quad (3.16)$$

D_{app} can be found by dividing equation (3.15) by equation (3.16)

$$S_1/S_2 = \exp[(b_2 - b_1)D_{app}]. \quad (3.17)$$

Solving equation (3.17) with respect to D_{app} gives

$$\text{ADC} = -\frac{1}{b_2 - b_1} \ln \left(\frac{S_2}{S_1} \right) \quad (3.18)$$

where $D_{app} = \text{ADC}$. It is also possible to estimate D with more than two b -factors. This is done by a linear least-squares fit of the log-transformed signal intensities. The ADC values of each voxel can be illustrated in an ADC map, where the pixel intensity is proportional to the ADC value. A high ADC value indicates high diffusivity and gives a bright pixel. A low ADC value indicates reduced diffusivity, and the pixel becomes dark. Note that the intensities of an ADC map are opposite of that of a diffusion-weighted image, where a high degree of diffusion results in a low signal intensity S , and hence a low pixel intensity.

3.3 Artefacts

An MRI artefact is a feature appearing in the final image that is not present in the original object. The artefacts are typically classified as either patient-related, signal-processing dependent or hardware-related[17]. A brief overview of the most common artefacts in DWI are presented in the following sections.

3.3.1 Motion artefacts

Motion artefacts are one of the most common artefacts in MRI. The DW-MRI technique is sensitive to stochastic motions in the micrometer range. As a consequence, the technique is particularly sensitive to macroscopic motions of the object being imaged. The two main components of motion in brain imaging are bulk movements of the head (translational and rotational) and periodic local movements of brain tissue due to cardiac-induced pulsations of the cerebrospinal fluid[18]. Motions during the image acquisition may give rise to a ‘ghost image’, where the object being imaged appears shifted over a distance, commonly referred to as ghosting.

3.3.2 Susceptibility artefacts

Susceptibility artefacts may arise due to differences in magnetic susceptibilities of the tissues and materials being imaged. In transitions where the differences are large, typically around tissue/air boundaries (such as the frontal lobe) or metal implants, local magnetic field distortions may arise in the B_0 field. The field distortions causes a change in the frequency precession of the spins. The EPI-sequences are very sensitive to B_0 -variations due to the single-shot nature of the k-space read-out. The off-resonance frequency of the spins produces an additive phase during

the EPI trajectory, resulting in a geometric distortion, or warping, of the image. Rapid changes in the magnetic field also give rise to signal variations, seen as signal brightening and signal voids, due to the displacement of signal registration during image reconstruction relative to true location[19].

3.3.3 Eddy currents

DWI uses rapidly changing magnetic gradients during the diffusion measurements. The rapid change in magnetic fields may induce eddy currents in the metallic components of the MR scanner, as well as in wires and devices within the patient, or in the patient as a whole. The eddy currents generate off-resonance magnetic fields, causing the spins to dephase, which results in warping of the image. The warping caused by eddy currents is different than the warping caused by constant magnetic field inhomogeneities, as the eddy currents change with each diffusion-encoding gradient direction. The eddy currents induced by diffusion gradients can be corrected for by either fitting a small constant gradient to each diffusion-weighted image, or by reducing the eddy currents by hardware revisions[19].

Chapter 4

Diffusion Tensor Imaging

So far the discussion has been about unrestricted, *isotropic* diffusion. Measurements of isotropic diffusion along any arbitrary axis give the same result, hence only one measurement is needed to describe isotropic diffusion. However, the diffusion of water molecules in biological tissues is often restricted by various barriers, like cell membranes, muscle fibers and nerve fibers. Water molecules tend to diffuse along these fibers, which results in *anisotropic* diffusion. This means that the diffusion is more likely to occur in one direction than another. Anisotropic diffusion can not be described by measurements in just one direction, simply because the diffusivity is direction specific. In order to deduce the underlying tissue anatomy, diffusion has to be measured in a number of non-collinear directions. The measurement orientation with the highest diffusivity is most likely aligned with the fiber, while the measurement orientation with the lowest diffusivity is aligned perpendicular to the fiber. In theory, the diffusivity could be measured in thousands of different directions to find the primary axis of diffusion. In a clinical situation however, the scanning time is often limited, and so is the number of diffusion measurements. Mathematical models are used to describe anisotropic diffusion when the number of gradient directions are limited.

Unless stated otherwise, the theory in this chapter is collected from *The Physics of Magnetic Resonance Imaging* (A. Bjørnerud)[9] and *Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors* (P. B. Kingsley)[16].

4.1 DTI mathematics

Diffusion tensor imaging (DTI) is a mathematical model that uses tensor calculations to describe the diffusion anisotropy in a voxel. In DTI, the diffusion anisotropy is represented by a three-dimensional tensor called the diffusion tensor, \mathbf{D} . The

eigenvectors and eigenvalues of the diffusion tensor describe the size and orientation of the diffusion in a voxel, and can be illustrated as an ellipsoid (see figure 4.1 b). The eigenvectors ϵ_1 , ϵ_2 and ϵ_3 determine the orientation of the principal axes of the ellipsoid, and the eigenvalues λ_1 , λ_2 and λ_3 represent the length of the principal axes. An ellipsoid with center at the origin of a Cartesian coordinate system can be described mathematically by the equation

$$\frac{(x')^2}{6\lambda_1 t} + \frac{(y')^2}{6\lambda_2 t} + \frac{(z')^2}{6\lambda_3 t} = 1 \quad (4.1)$$

where λ_1 , λ_2 and λ_3 are the eigenvalues of the corresponding eigenvectors ϵ_1 , ϵ_2 and ϵ_3 of the ellipsoid. More often than not, the eigenvectors of the ellipsoid are not aligned with the x , y - and z -axes of the laboratory frame. The marked coordinates (x', y', z') represent the coordinate system of the ellipsoid, defined by the direction of the eigenvectors. The eigenvalues symbolize the diffusivity along the principal axes. The principal axes of the ellipsoid are axes of symmetry, which means that the diffusivity along one direction of the axis is the same in the other direction. If the diffusivity is equal in all three directions ($\lambda_1 = \lambda_2 = \lambda_3$), equation (4.1) reduces to that of a sphere, and the diffusion is isotropic (see figure 4.1 a).

In DTI, the diffusion in a voxel is described by the three-dimensional diffusion tensor \mathbf{D} ,

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}. \quad (4.2)$$

The diffusion tensor is symmetric, $D_{ij} = D_{ji}$ where $i, j = \{x, y, z\}$, as a result of the symmetric properties of the diffusion ellipsoid. The diagonal elements, D_{xx} , D_{yy} and D_{zz} , represent the diffusivity along the primary axes of the ellipsoid. The off-diagonal elements are non-zero when the diffusion ellipsoid is rotated relative to the laboratory reference frame.

The primary objective in DTI is to measure the diffusion in a sufficient number of directions to determine \mathbf{D} . When the elements of \mathbf{D} are known, the eigenvectors and eigenvalues can be obtained by linear algebra.

If the orientation of the diffusion ellipsoid was known beforehand, the diffusion gradients could be placed in the same direction as the eigenvectors of the ellipsoid straight away. The off-diagonal elements in \mathbf{D} would then be zero. In this case, exactly three measurements would be sufficient in order to determine the diffusion tensor. However, the orientation of the diffusion ellipsoid is never known beforehand,

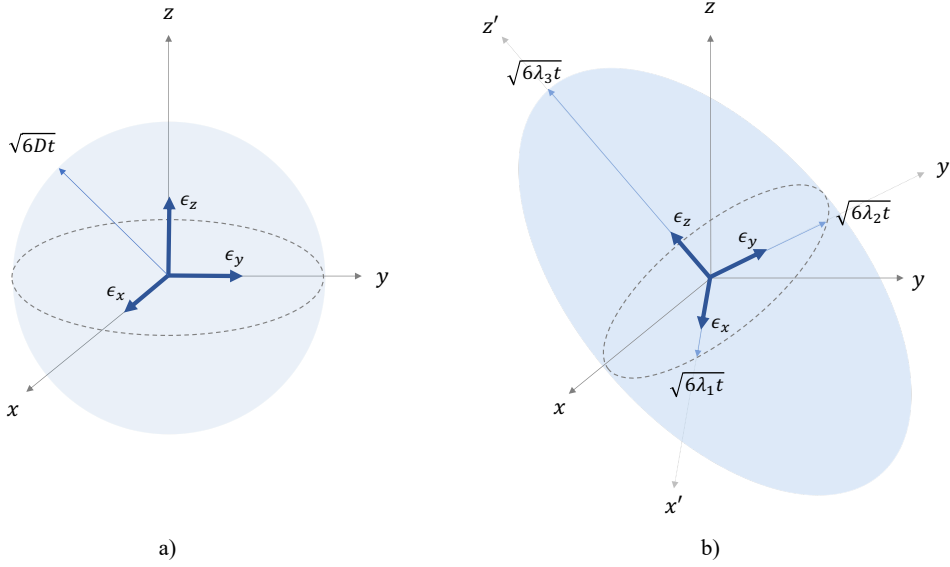


Figure 4.1: a) In isotropic diffusion, the diffusion probability is the same in all directions, resulting in a sphere where the radius is given by the mean diffusion distance $\sqrt{6Dt}$. b) Anisotropic diffusion can be illustrated by an ellipsoid, where the eigenvectors determine the axes of the ellipsoid, and the eigenvalues determine the mean diffusion distance along the axes. The principal axes are axes of symmetry. Adapted from [9, 15].

and different voxels have different tensor orientations. It is therefore necessary to measure the diffusion in more than three directions. Due to its symmetric properties, there are only six unique elements in the diffusion tensor; D_{xx} , D_{yy} , D_{zz} , D_{xy} , D_{xz} and D_{yz} . The minimum number of gradient directions needed in order to uniquely define \mathbf{D} is therefore six, provided that the gradient directions are optimally distributed in space. A sample DTI acquisition is shown in figure 4.2.

The diffusion matrix \mathbf{b} describes the diffusion weighting along the different directions and is given by

$$\mathbf{b} = \begin{bmatrix} b_{xx} & b_{xy} & b_{xz} \\ b_{xy} & b_{yy} & b_{yz} \\ b_{xz} & b_{yz} & b_{zz} \end{bmatrix}. \quad (4.3)$$

The diffusion gradient \mathbf{G} can be expressed as a vector with coordinates g_x , g_y and g_z

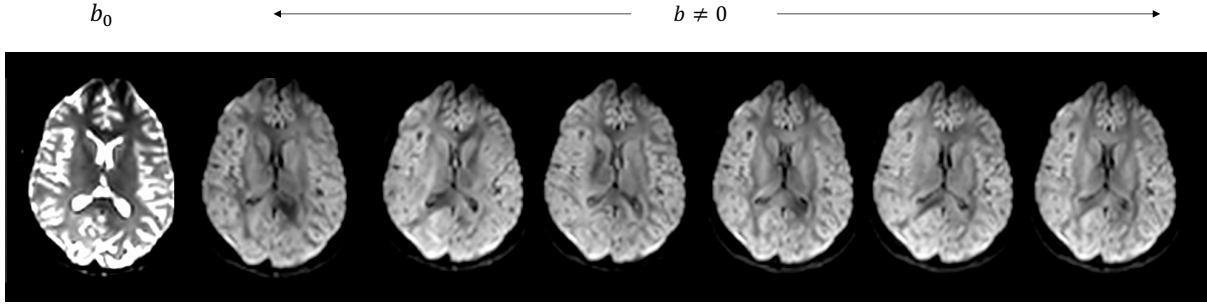


Figure 4.2: Diffusion-weighted images of the brain. The figure shows DW images for six different diffusion directions ($b \neq 0$), together with the non diffusion-weighted image ($b = 0$). Notice that the white matter signal intensity varies with the direction of the diffusion measurement. Diffusion must be measured in a minimum of six different directions to determine the diffusion tensor.

$$\mathbf{G} = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} \quad (4.4)$$

where the length of \mathbf{G} is normalized, $g_x^2 + g_y^2 + g_z^2 = 1$, such that the gradient amplitude is the same for all measurements in all directions. The diffusion matrix \mathbf{b} can be expressed as the product of the diffusion weighting constant b and the normalized diffusion vector \mathbf{G}

$$\mathbf{b} = b\mathbf{g} \quad (4.5)$$

where \mathbf{g} is the outer product of \mathbf{G} , given by

$$\mathbf{g} = \mathbf{G}_n \mathbf{G}_n^T = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} \begin{bmatrix} g_x & g_y & g_z \end{bmatrix} = \begin{bmatrix} g_x^2 & g_x g_y & g_x g_z \\ g_y g_x & g_y^2 & g_y g_z \\ g_z g_x & g_z g_y & g_z^2 \end{bmatrix}. \quad (4.6)$$

\mathbf{G}_n is the n th gradient vector, and \mathbf{G}_n^T is the transpose of \mathbf{G}_n . The signal intensity is a mono-exponentially decreasing function, where the signal attenuation is related

to the dot product¹ of the diffusion tensor \mathbf{D} and the diffusion matrix \mathbf{b} ,

$$S = S_0 e^{-\mathbf{b} \cdot \mathbf{D}} = S_0 e^{-\text{Tr}(\mathbf{bD})} \quad (4.7)$$

where $\mathbf{b} \cdot \mathbf{D}$ is the dot product and \mathbf{bD} is the matrix product of the two matrices. $\text{Tr}(\mathbf{bD})$ is the *trace* of the square matrix \mathbf{bD} , which is the sum of the elements on the main diagonal of \mathbf{bD} . The dot product $\mathbf{b} \cdot \mathbf{D}$ is equal to the trace of \mathbf{bD} , with the solution

$$\mathbf{b} \cdot \mathbf{D} = \text{Tr}(\mathbf{bD}) = b_{xx}D_{xx} + b_{yy}D_{yy} + b_{zz}D_{zz} + 2b_{xy}D_{xy} + 2b_{xz}D_{xz} + 2b_{yz}D_{yz}. \quad (4.8)$$

By inserting equation (4.5), which states that $\mathbf{b} = b\mathbf{g}$, into equation (4.7), the signal attenuation now becomes

$$S = S_0 e^{-\text{Tr}(\mathbf{bD})} = S_0 e^{-b \cdot \text{Tr}(\mathbf{gD})} \quad (4.9)$$

where the exponent can be expressed as

$$-b \cdot \text{Tr}(\mathbf{gD}) = -b(D_{xx}g_x^2 + D_{yy}g_y^2 + D_{zz}g_z^2 + 2D_{xy}g_xg_y + 2D_{xz}g_xg_z + 2D_{yz}g_yg_z). \quad (4.10)$$

When the diffusion is measured along N different directions, where $N \geq 6$, the signal attenuation in each direction can be expressed as a vector with N elements on the form

$$\mathbf{Y}_i = \frac{1}{b} \ln \left(\frac{S_0}{S_i} \right), \quad i = \{1, \dots, N\}. \quad (4.11)$$

The signal attenuation vector containing all the gradient directions is therefore

$$\mathbf{Y} = \frac{1}{b} \left[\ln \left(\frac{S_0}{S_1} \right) \quad \ln \left(\frac{S_0}{S_2} \right) \quad \dots \quad \ln \left(\frac{S_0}{S_N} \right) \right]. \quad (4.12)$$

Rearrangement of equation (4.9) gives the expression

¹also called the scalar product

$$\text{Tr}(\mathbf{gD}) = \frac{1}{b} \ln \left(\frac{S_0}{S_i} \right) \quad (4.13)$$

where S_i is the signal amplitude in the i th gradient direction. Combining equation (4.13) with equation (4.11) gives

$$\mathbf{Y}_i = \text{Tr}(\mathbf{gD}) = D_{xx}g_{xi}^2 + D_{yy}g_{yi}^2 + D_{zz}g_{zi}^2 + 2D_{xy}g_{xi}g_{yi} + 2D_{xz}g_{xi}g_{zi} + 2D_{yz}g_{yi}g_{zi} \quad (4.14)$$

which can be expressed in matrix form as

$$\mathbf{Y}_i = \begin{bmatrix} g_{xi}^2 & g_{yi}^2 & g_{zi}^2 & 2g_{xi}g_{yi} & 2g_{xi}g_{zi} & 2g_{yi}g_{zi} \end{bmatrix} \begin{bmatrix} D_{xx} \\ D_{yy} \\ D_{zz} \\ D_{xy} \\ D_{xz} \\ D_{yz} \end{bmatrix}. \quad (4.15)$$

The first part of equation (4.15) describes the i th gradient direction, and is defined as

$$\mathbf{H}_i = \begin{bmatrix} g_{xi}^2 & g_{yi}^2 & g_{zi}^2 & 2g_{xi}g_{yi} & 2g_{xi}g_{zi} & 2g_{yi}g_{zi} \end{bmatrix}, \quad i = \{1, \dots, N\}. \quad (4.16)$$

The second part in equation (4.15) is defined as \mathbf{d} , and contains the six unique elements of the diffusion tensor,

$$\mathbf{d} = \begin{bmatrix} D_{xx} & D_{yy} & D_{zz} & D_{xy} & D_{xz} & D_{yz} \end{bmatrix}^T. \quad (4.17)$$

The N gradient directions are collected into one big matrix \mathbf{H} , where the i th row represents the i th gradient direction

$$\mathbf{H} = \begin{bmatrix} g_{x1}^2 & g_{y1}^2 & g_{z1}^2 & 2g_{x1}g_{y1} & 2g_{x1}g_{z1} & 2g_{y1}g_{z1} \\ g_{x2}^2 & g_{y2}^2 & g_{z2}^2 & 2g_{x2}g_{y2} & 2g_{x2}g_{z2} & 2g_{y2}g_{z2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{xN}^2 & g_{yN}^2 & g_{zN}^2 & 2g_{xN}g_{yN} & 2g_{xN}g_{zN} & 2g_{yN}g_{zN} \end{bmatrix}. \quad (4.18)$$

The signal attenuation in all the gradient directions can now be expressed in matrix form. This results in the following matrix equation

$$\mathbf{Y} = \mathbf{H}\mathbf{d}. \quad (4.19)$$

In order to determine the elements of \mathbf{d} , equation (4.19) must be solved. If the diffusion is measured in the minimum number of required directions, then $N = 6$. \mathbf{H} is then a square 6×6 matrix, and an exact analytical solution to equation (4.19) exists,

$$\mathbf{d} = \mathbf{H}^{-1}\mathbf{Y} \quad (4.20)$$

where \mathbf{H}^{-1} is the inverse of \mathbf{H} . The solution of equation (4.20) returns the elements of \mathbf{d} . In practice, diffusion is measured along more than six directions. When $N > 6$, \mathbf{H} is no longer a square 6×6 matrix, but an $N \times 6$ matrix. In order to solve equation (4.20) one can use linear regression methods like ordinary least squares (OLS) or weighted least squares (WLS).

4.2 Ordinary least squares

The signal attenuation \mathbf{Y}_i in the i th gradient direction is linearly related to the signal amplitude S_i and \mathbf{H}_i , and can be expressed as

$$\mathbf{Y}_i = \frac{1}{b} \ln \left(\frac{S_0}{S_i} \right) = \mathbf{H}_i \mathbf{d} = \sum_j^N H_{ij} d_j \quad (4.21)$$

where N is the number of directions. When $N > 6$, the OLS method can be used to solve equation (4.21) by finding the lowest sum of squared residuals,

$$R = \sum_{i=1}^n (r_i)^2 \quad (4.22)$$

where the residual r_i is the difference between the observed data and the model,

$$\begin{aligned} R &= (\mathbf{Y} - \mathbf{H}\mathbf{d})^2 \\ &= \mathbf{Y}^T \mathbf{Y} - 2\mathbf{d}^T \mathbf{H}^T \mathbf{Y} + \mathbf{d}^T \mathbf{H}^T \mathbf{H} \mathbf{d}. \end{aligned} \quad (4.23)$$

The residual R is minimized by deriving equation (4.23) with respect to \mathbf{d} and setting the equation equal to zero. This results in the expression

$$\mathbf{H}^T \mathbf{H} \mathbf{d} = \mathbf{H}^T \mathbf{Y} \quad (4.24)$$

where \mathbf{d} can be expressed as

$$\mathbf{d} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{Y}. \quad (4.25)$$

The term $(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ is called the *pseudoinverse* of \mathbf{H} ,

$$\mathbf{H}^\Psi = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \quad (4.26)$$

and the solution to equation (4.20) can now be obtained by solving the expression

$$\mathbf{d} = \mathbf{H}^\Psi \mathbf{Y}. \quad (4.27)$$

4.3 Weighted least squares

The WLS method is a generalization of the ordinary least squares method, and uses non-negative constants, or *weights*, on the measurements. Instead of minimizing the sum of squared residuals, as seen in OLS, WLS minimizes the sum of weighted squared residuals,

$$R = \sum_{i=1}^n w_i (r_i)^2. \quad (4.28)$$

The weights give more precise measurements a bigger influence than less precise measurements. With OLS, all the measurements are assumed to be equally precise,

thus all of the weights are equal to 1.

Conventional MR images are not completely noise-free. For a set of MRI-signals, the natural logarithm of the signals $\tilde{\mathbf{S}}$ can be modeled as

$$\mathbf{Y} = \mathbf{H}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (4.29)$$

where

$$\mathbf{Y} = \left[\ln \tilde{S}(b_1, \mathbf{g}_1), \dots, \ln \tilde{S}(b_N, \mathbf{g}_N) \right]^T \quad (4.30)$$

is the vector containing all the signals, \mathbf{H} is the design matrix with all the diffusion gradient directions and gradient strengths, $\boldsymbol{\beta}$ is the parameter vector with all the tensor elements and the noise-free non-diffusion-weighted signal, and $\boldsymbol{\epsilon}$ is the vector containing the error terms[20].

When the error vector $\boldsymbol{\epsilon}$ is zero, the OLS estimation can be used to solve equation 4.29, where

$$\hat{\boldsymbol{\beta}}_{OLS} = \mathbf{d} = \left(\mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{Y}. \quad (4.31)$$

When $\boldsymbol{\epsilon}$ is non-zero, the best linear estimator of $\boldsymbol{\beta}$ can be designed by including a weight matrix \mathbf{W} into equation (4.31),

$$\hat{\boldsymbol{\beta}}_{WLS} = \left(\mathbf{H}^T \mathbf{W} \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{W} \mathbf{Y}, \quad (4.32)$$

where the weights are expressed as

$$\mathbf{W} = \text{diag} \left(\mathbf{S}^2(\mathbf{b}, \mathbf{g}) \right). \quad (4.33)$$

Because the noise-free signals \mathbf{S} are unknown, the weight matrix must be estimated. There are several ways of estimating \mathbf{W} . This thesis uses a method where the weights are given by

$$\tilde{\mathbf{W}} = \text{diag}\left(\exp(2\mathbf{H}\hat{\boldsymbol{\beta}}_{OLS})\right). \quad (4.34)$$

4.4 Obtaining the diffusion tensor

The eigenvectors and eigenvalues of \mathbf{D} are obtained by solving the equation

$$\mathbf{D} \cdot \boldsymbol{\epsilon} = \lambda \cdot \mathbf{I}_3 \cdot \boldsymbol{\epsilon} \quad (4.35)$$

where $\boldsymbol{\epsilon}$ is the matrix containing the eigenvectors, λ is the eigenvalues and \mathbf{I}_3 is the identity matrix. In matrix from the equation becomes

$$\begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} \cdot \begin{bmatrix} \epsilon_{1x} & \epsilon_{2x} & \epsilon_{3x} \\ \epsilon_{1y} & \epsilon_{2y} & \epsilon_{3y} \\ \epsilon_{1z} & \epsilon_{2z} & \epsilon_{3z} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} \epsilon_{1x} & \epsilon_{2x} & \epsilon_{3x} \\ \epsilon_{1y} & \epsilon_{2y} & \epsilon_{3y} \\ \epsilon_{1z} & \epsilon_{2z} & \epsilon_{3z} \end{bmatrix} \quad (4.36)$$

where $\boldsymbol{\epsilon}_i = [\epsilon_{ix}, \epsilon_{iy}, \epsilon_{iz}]$ is eigenvector number i , and λ_i the corresponding eigenvalue, for $i = \{1, 2, 3\}$. The solution of equation (4.35) with respect to λ is obtained by solving

$$\det(\mathbf{D} - \lambda\mathbf{I}) = 0. \quad (4.37)$$

In the special case, where the orientation of the ellipsoid is known, a calculation of equation (4.37) is unnecessary. When the diffusion gradients are placed in the direction of the eigenvectors, the off-diagonal elements of \mathbf{D} are reduced to zero. The diffusion tensor becomes diagonal, where the diagonal elements simply are the eigenvalues of the ellipsoid,

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad (4.38)$$

with the eigenvectors $\boldsymbol{\epsilon}_1 = [1 \ 0 \ 0]$, $\boldsymbol{\epsilon}_2 = [0 \ 1 \ 0]$ and $\boldsymbol{\epsilon}_3 = [0 \ 0 \ 1]$.

4.5 Diffusion anisotropy in the brain

Diffusion of water molecules in the brain can be either isotropic or anisotropic, depending on the region of interest. The tissue of the brain can be divided into two types; white matter and gray matter. Gray matter is made up of neuron cell bodies, dendrites and axon terminals, and white matter is made up of myelinated axons, or nerve fibers (see figure 4.3). The myelin coating of the axon speeds up the electrical signals sent from one cell body to another. Myelin is a lipid-rich substance, and white matter gets its bright appearance from the lipid content of myelin. The axons are organized in bundles, forming nerve tracts that connect the different regions of gray matter to each other. In addition to gray and white matter, there is cerebrospinal fluid (CSF), a clear and colorless fluid that can be found in the brain and spinal cord.

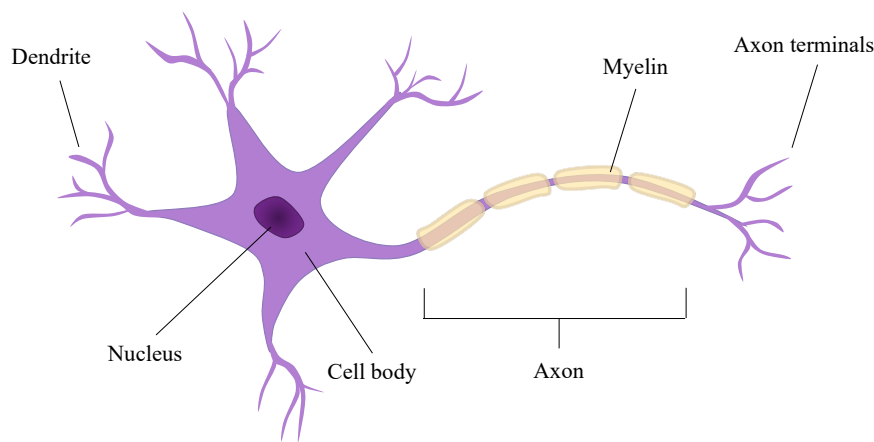


Figure 4.3: Simplified anatomy of a neuron. The human brain has two kinds of tissue, gray matter and white matter. Gray matter contains neuron cell bodies, dendrites and axon terminals, and white matter is made of myelinated axons. The dendrites receive signals, while the axon terminals transmit signals. The axons in white matter form nerve tracts, connecting the various parts of gray matter to each other.

Diffusion is free and isotropic in regions with few or no barriers, which is typical in the CSF. Diffusion is more anisotropic in white matter, because the water molecules tend to diffuse along the nerve tracts. Gray matter has a complex structure and contains lots of biological barriers, but is not organized in the same way as white matter. The typical resolution of a diffusion-weighted MR image is of the order of 2-3 mm in all three directions. With an observational window of this size, the diffusion will still appear isotropic, even though the diffusion is hindered. This happens because the gray matter structures change direction on a smaller scale than the length scale of a voxel. On the other hand, many regions of white matter have nerve tracts that are 2 mm or longer, which is why the diffusion anisotropy of white matter stands out in DTI. The diffusion processes are illustrated in figure 4.4.

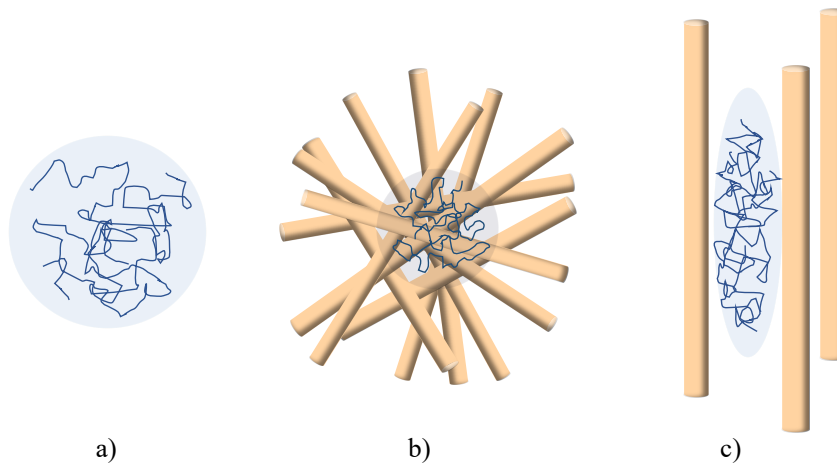


Figure 4.4: The diffusion process in different regions of the brain. a) The diffusion process is unrestricted and isotropic in the CSF. b) Diffusion in gray matter is hindered, but also isotropic on the length scale of a voxel because of the structural complexity. c) Diffusion in white matter is anisotropic, with primary diffusion direction along the nerve fibers. Adapted from[21].

4.6 DTI parameters

Anisotropic diffusion can not be fully described by a single scalar, because of its direction specific properties. In order to generate an MR image that illustrates the diffusion anisotropy, a scalar value has to be derived from the diffusion tensor \mathbf{D} . The scalar value reflects the degree of diffusion anisotropy in that specific voxel. Each voxel can then be assigned a scalar value, and the intensity of each voxel in the image is dependent on the value of the scalar. These scalars are referred to as *DTI parameters*. There are several different DTI parameters that describe anisotropic diffusion, the two most frequently used being the mean diffusivity (MD) and fractional anisotropy (FA).

The direction of the eigenvectors depend on the orientation of the diffusion tensor. If the patient moves, or the slice orientation is changed, the direction of the eigenvectors also change. The tensor eigenvalues however, remain the same, because they are *rotationally invariant*. This means that eigenvalues do not change under rotation. The eigenvalues are the same even when the patient moves, or the orientation of the slice is changed. The tensor eigenvalues are ideal for characterizing anisotropic diffusion, because of their rotationally invariant properties. DTI parameters can therefore be compared between patients.

4.6.1 Mean Diffusivity

The mean diffusivity describes the average diffusion properties of a voxel. It represents the degree of diffusivity in a voxel, where unrestricted, free diffusion results in a high MD, and hindered diffusion results in a low MD. The value of MD gives an indication of the space available for the diffusion process to occur. The mean diffusivity can be found by taking the mean of the diagonal elements of \mathbf{D} , which is equivalent to the mean of the tensor eigenvalues,

$$\text{MD} = \frac{D_{xx} + D_{yy} + D_{zz}}{3} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \quad (4.39)$$

The mean diffusivity can also be calculated by taking the average of the ADC values,

$$\text{MD} = \frac{\sum_{i=1}^N \text{ADC}_i}{N} \quad (4.40)$$

where the ADC values are given in equation (3.18). The mean diffusivity can be represented by an MD map where the intensity of the pixel is proportional to the MD value (see figure 4.5). Regions with free and unrestricted diffusion result in high intensity pixels. The CSF has a high degree of diffusion and appears bright in an MD map. In regions where the diffusion is limited, like in the white matter, the pixels appear darker.

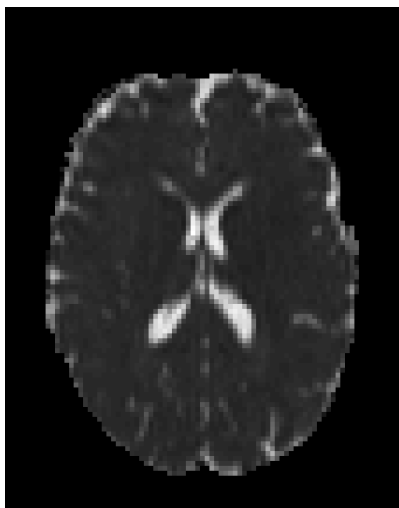


Figure 4.5: Example of an MD map. The CSF typically has a high degree of diffusion, hence a high MD, and appears bright in an MD map. Regions with high degree of hindrance have lower MD and appears darker.

4.6.2 Axial Diffusivity

The axial diffusivity (AD, or λ_{\parallel}) is simply the eigenvalue of the eigenvector oriented along the primary axis of the diffusion ellipsoid,

$$\text{AD} = \lambda_1 \quad (4.41)$$

and is a measure of the diffusivity in the axial direction. An AD map can be generated in the same manner as an MD map, where the intensity of the pixel is proportional to the AD value, as illustrated in figure 4.6a.

4.6.3 Radial Diffusivity

The radial diffusivity (RD, or λ_{\perp}) is the mean diffusivity in the transverse plane of the diffusion ellipsoid, and given by

$$\text{RD} = \frac{\lambda_2 + \lambda_3}{2}. \quad (4.42)$$

The axial diffusivity is usually larger than the radial diffusivity, because it is oriented along the primary axis of diffusion. When AD is larger than RD, the diffusion is anisotropic. If AD is close to RD, the diffusion is more isotropic. An RD map is illustrated in figure 4.6b.

4.6.4 Fractional Anisotropy

Fractional anisotropy (FA) quantifies the pointedness of the diffusion ellipsoid, without providing any information about the direction it is pointing to. Several different formulas exist for calculating the FA index. The formula used in this project is given by

$$\text{FA} = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (4.43)$$

where

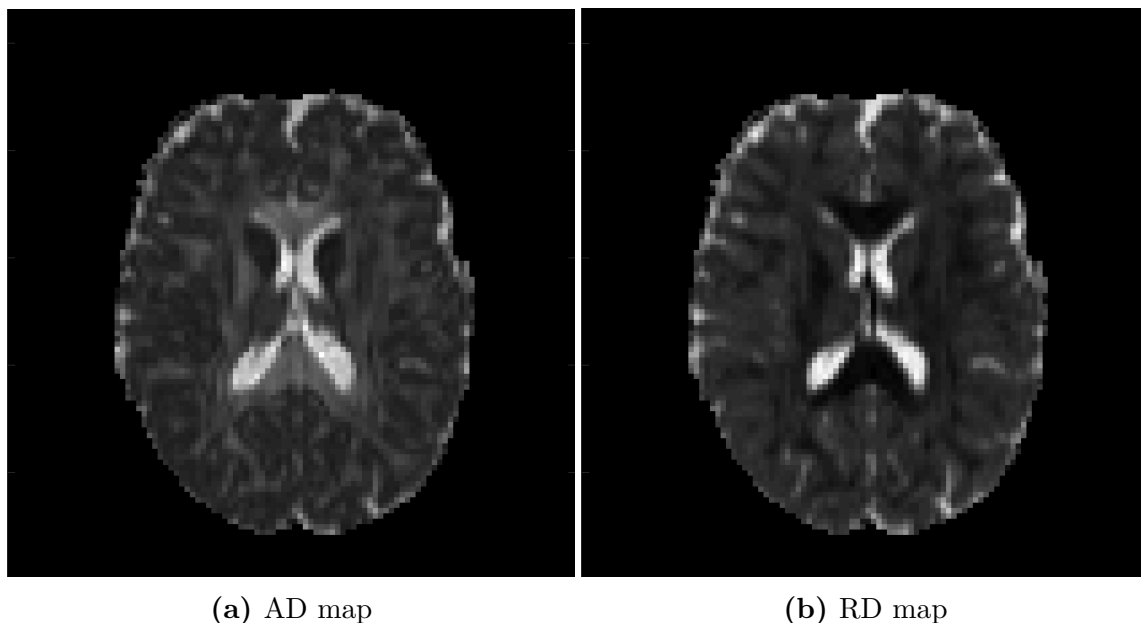


Figure 4.6: AD and RD parametric maps illustrates the degree of axial and radial diffusion, respectively. High intensity pixels represent a high degree of diffusion in the axial (a) or radial (b) direction.

$$\bar{\lambda} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \quad (4.44)$$

is the mean diffusivity. The values of FA range from 0 to 1. If $\lambda_1 = \lambda_2 = \lambda_3$, the diffusion is completely isotropic, and the FA index is reduced to 0. The FA index gets closer to one with progressive diffusion anisotropy (see figure 4.7). An FA index of 1 means that the diffusion is completely anisotropic, it is then fully restricted, and diffusion only occurs along one axis. In this case the diffusion tensor only has one non-zero eigenvalue, one eigenvector, and the ellipsoid is reduced to a line pointing in the direction of the eigenvector. This is however rarely the case in real measurements.

The FA values can be used to generate an FA image. An FA image, also called an FA map, is an MR image where the intensity of the pixel is related to the FA value in that pixel (see figure 4.8). The intensity of a pixel is represented by a grayscale. A pixel with $FA = 1$ appears white, and a pixel with $FA = 0$ appears black.

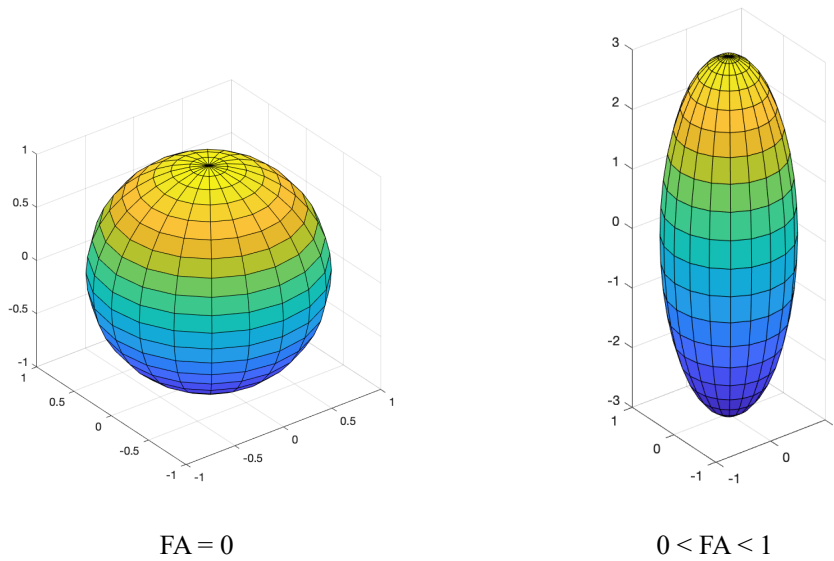


Figure 4.7: The FA index quantifies the pointedness of the diffusion ellipsoid, without providing any information about the direction of the principal axes. Isotropic diffusion results in an FA index of 0, and the diffusion ellipsoid becomes a sphere. Anisotropic diffusion has an FA index between 0 and 1, and the more anisotropic the diffusion is, the closer the FA index is to 1.



Figure 4.8: An FA map illustrates the degree of anisotropy in the tissue. The nerve tracts in the white matter of the brain are highly anisotropic and appears bright in an FA map, while the gray matter is less anisotropic and appears darker.

Chapter 5

Diffusion Kurtosis Imaging

The theory in this chapter is collected from the articles *Diffusion kurtosis imaging: an emerging technique for evaluating the microstructural environment of the brain* (A. J. Steven, J. Zhuo and E. R. Melhem)[22] and *MR diffusion kurtosis imaging for neural tissue characterization* (Ed X. Wu and Matthew M. Cheung)[23].

Diffusion kurtosis imaging (DKI) is a diffusion-weighted imaging model that takes into account the non-Gaussian diffusion of water molecules in biological systems. It is a simple extension of the diffusion tensor imaging (DTI) model. The DTI model assumes a monoexponential signal decay as a function of b , which is typical for Gaussian mediums, where the diffusion is free and unrestricted, or so-called isotropic. It is important to note that the DTI model does not assume that the diffusion occurs in a Gaussian medium, but rather that biological tissue is indistinguishable from a Gaussian medium when lower b -values are used[24]. However, when higher b -values are used, the assumption about the monoexponential signal decay is no longer valid. The microstructural environment of the tissue is filled with biological barriers like macromolecules, cell membranes, and nerve axons. These barriers restrict the diffusion of water molecules, leading to a diffusion pattern that deviates from the Gaussian diffusion distribution. DKI provides a more accurate diffusion model, especially at higher b -values, by quantifying the deviation from a Gaussian diffusion distribution. The deviation is quantified by the kurtosis¹ K , and the value of K may be used as an indicator of the complexity of the microstructural environment of the tissue. In the DKI model, the signal attenuation is related to both the apparent diffusivity, D_{app} , and the apparent kurtosis K_{app} .

DKI provides a new imaging metric useful in many clinical applications. Several different pathologies cause changes in the morphology of the brain, which may alter the diffusion anisotropy of the tissue. The mean kurtosis, which is the average kurtosis measured along with all diffusion gradient directions, has been shown to

¹In this document, the term kurtosis refers to the *excess kurtosis*. The excess kurtosis is simply the kurtosis minus 3, to provide the comparison to the normal distribution.

offer an improved sensitivity in detecting pathological changes, in comparison to conventional DTI.

Some neurodegenerative diseases, like Alzheimer's, cause neuronal loss, which can result in a decreased kurtosis value. This is because the loss of neuron cell bodies, synapses and dendrites lead to an increase in extracellular space and less restricted diffusion[25]. Neuronal loss is also associated with the aging of the brain, and there is a significant relationship between the age of the brain and the mean kurtosis value[26]. Diffusion kurtosis imaging can therefore be used in the evaluation of the aging brain or as a biological marker for neurodegenerative diseases.

Demyelination of axons increases the water exchange between the intra-axonal and extra-axonal compartments. The diffusion in the radial direction is therefore increased due to demyelination, resulting in a decreased radial kurtosis[27]. The radial kurtosis may therefore act as a marker for demyelinating diseases, like multiple sclerosis.

5.1 Kurtosis

In probability theory and statistics, the kurtosis K is a term that reflects the deviation of a probability distribution from the Gaussian distribution. The kurtosis is a unitless measure, and the higher the K , the bigger the deviation. When water is diffusing freely, the diffusion distribution is Gaussian, with $K = 0$. A probability distribution with a positive kurtosis has a higher peak and heavier tails, and a distribution with a negative kurtosis has a lower peak and lighter tails, compared to the Gaussian distribution, as illustrated in figure 5.1.

A positive, non-zero kurtosis value, $K > 0$, indicates tissue heterogeneity and that the diffusion is directional, while $K = 0$ signifies Gaussian diffusion (see figure 5.2). The (mean) kurtosis typically becomes higher with increasing tissue complexity and is related to the degree of non-Gaussian diffusion. White matter usually has a higher kurtosis value, because the diffusion is directional along the nerve fibers. Gray matter has a complex structure, but is not as organized as white matter, and has a lower kurtosis value. In the CSF, the kurtosis is zero, because the diffusion follows a Gaussian distribution probability. In biological systems, the kurtosis term is always non-negative, $K \geq 0$. A negative diffusion kurtosis would not be possible, because the diffusion process can not become more homogeneous than it already is in a Gaussian distribution, where $K = 0$.

5.2 The DKI signal

The signal attenuation as predicted by the DTI model can be expressed as

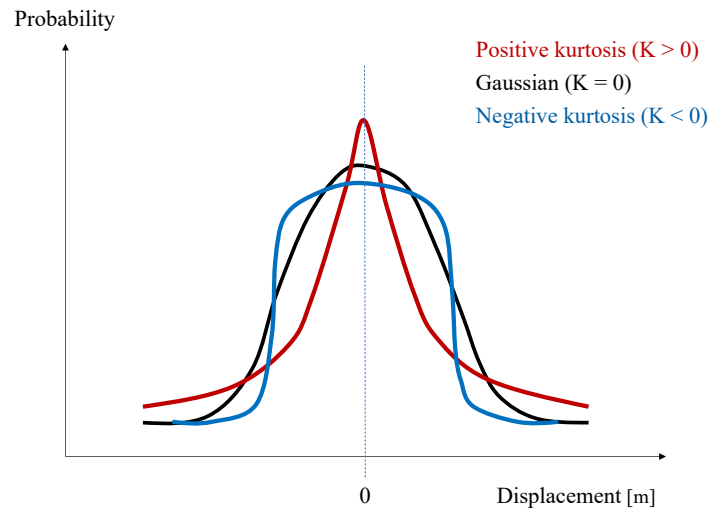


Figure 5.1: Different degrees of kurtosis. A probability distribution with positive kurtosis has a higher peak and heavier tails. A negative kurtosis results in a lower peak and lighter tails. A Gaussian distribution has a kurtosis of zero. Biological systems always have a non-negative kurtosis, $K \geq 0$. Adapted from [22].

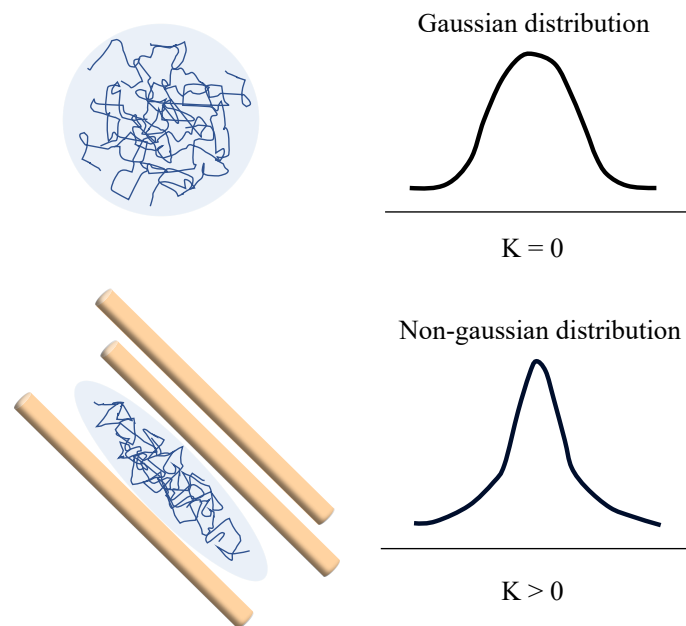


Figure 5.2: Free and unrestricted diffusion (top figure) can be described by a Gaussian diffusion distribution, where the diffusion is equal in all directions. A Gaussian distribution has a kurtosis of zero, $K = 0$. Diffusion in biological tissue (bottom figure) can be both restricted and hindered, which results in a non-Gaussian diffusion distribution. The diffusion is directional, and has a positive, non-zero kurtosis, $K > 0$. Adapted from [22].

$$\ln\left(\frac{S(b)}{S_0}\right) = -bD_{app} \quad (5.1)$$

where $S(b)$ is the diffusion-weighted signal in a certain direction with diffusion weighting b , S_0 is the signal without diffusion weighting ($b = 0$), and D_{app} is the apparent diffusion coefficient (ADC). Equation (5.1) predicts a monoexponential signal decay, and there is a linear relationship between the logarithmic signal decay $\ln(S(b)/S_0)$ and the b -value, as illustrated in figure 5.3. However, when diffusion is measured at higher b -values, the diffusion pattern deviates more and more from a Gaussian distribution. This is often the case in biological systems, where diffusion is restricted. An excess kurtosis term might be added to the previous expression (5.1), resulting in the DKI signal attenuation

$$\ln\left(\frac{S(b)}{S_0}\right) = -bD_{app} + \frac{1}{6}b^2D_{app}^2K_{app} \quad (5.2)$$

where K_{app} is the apparent diffusion kurtosis in a certain direction. When $K_{app} = 0$, the kurtosis term disappears, and the DKI model is reduced to the DTI model (5.1). The DKI model describes the diffusion-weighted signal attenuation without imposing any biophysical modeling. In other words, it is simply a mathematical model that describes the signal behavior without making any assumptions about the underlying tissue being imaged. Thus, it can be used on any type of tissue[24].

Measurements of the diffusion-weighted signal indicate that the signal attenuation deviates from the monoexponential decay predicted by the DTI model. This becomes evident at higher b -values, where $b > 1500$ s/mm². Higher b -values give the signal a stronger diffusion-weighting, making the diffusion gradient more sensitive to shorter diffusion distances. A high b -value therefore highlights the tissue complexity even more, and the diffusion gradient picks up more signals. The signal attenuation is therefore not completely monoexponential at higher b -values. The DKI model predicts a signal attenuation that deviates from a monoexponential decay. Diffusion signal measurements have shown that the real data fits the curve shape of the DKI model better than the monoexponential DTI model at higher b -values.

5.3 The diffusion kurtosis tensor

The diffusivities along the different directions are characterized by the *diffusion tensor* (DT), a 2nd order 3×3 tensor whose eigenvectors point in the direction of the principal axes of the diffusion ellipsoid and eigenvalues represent the diffusivities

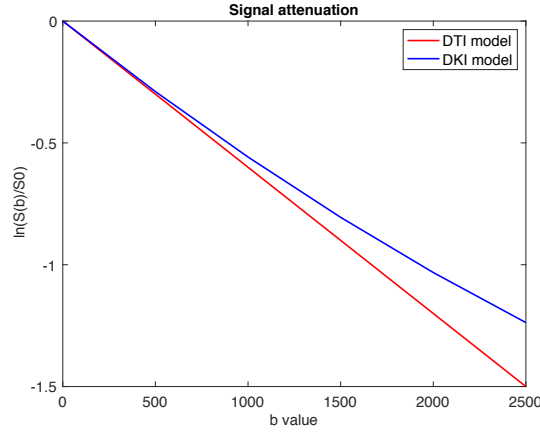


Figure 5.3: Graph showing the logarithmic signal decay as predicted by the DTI model and the DKI model in equation (5.1) and (5.2), with apparent diffusion coefficient $D_{app} = 0.6 \times 10^{-3} \text{ mm}^2/\text{s}$ and $K_{app} = 0.7$. The DTI model predicts a monoexponential signal decay, which results in a linear decay in a logarithmic plot. However, at higher b -values, measurements has shown that the real signal deviates from the linear diffusion function, and that the DKI model fits the real data better.

along the principal axes. Since the diffusion is directional, the diffusion kurtosis also varies with direction. The directional kurtosis is represented by a 4th order $3 \times 3 \times 3 \times 3$ tensor called the *kurtosis tensor* (KT), often denoted as W_{ijkl} . The kurtosis tensor describes the non-Gaussian behavior of the diffusing water molecules. The spatial distribution of the kurtosis tensor is complex and can not be illustrated by an ellipsoid, like the diffusion tensor. Instead, it is often illustrated by an oblate ellipsoid (see figure 5.4), with the directional kurtosis values K_1 , K_2 and K_3 along the eigenvectors ϵ_1 , ϵ_2 and ϵ_3 of DT.

The apparent kurtosis coefficient (AKC, or K_{app}) along an arbitrary direction is estimated through the kurtosis tensor W_{ijkl} :

$$K_{app} = \frac{MD^2}{D_{app}^2} \cdot \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 n_i n_j n_k n_l W_{ijkl} \quad (5.3)$$

where MD is the mean diffusivity and n_i is the i th element of the diffusion direction. The diffusion process may be asymmetric in biological systems. However, in the scale of a voxel, the diffusion can be seen as symmetric, because the diffusion distance of a water molecule is much smaller than the voxel dimension, during a measurement. The symmetry of the diffusion process reduces equation (5.3) to

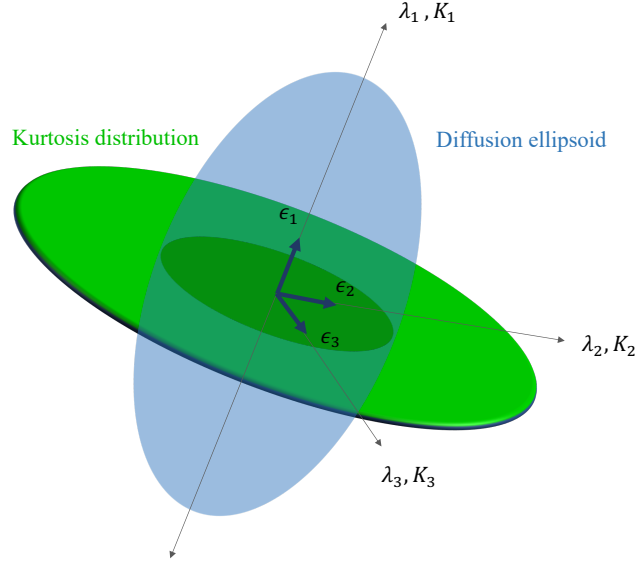


Figure 5.4: Graphical illustration of the 3D diffusion distribution and kurtosis distribution. The diffusion distribution is illustrated by an ellipsoid, where the eigenvectors ϵ_1 , ϵ_2 and ϵ_3 of the diffusion tensor define the principal axes of the ellipsoid. The 3D kurtosis distribution is complex and can not be illustrated by an ellipsoid. For simplicity, the kurtosis distribution is represented by an *oblate* ellipsoid, with directional kurtoses K_1 , K_2 and K_3 along the eigenvectors ϵ_1 , ϵ_2 and ϵ_3 . Adapted from [23].

$$\begin{aligned}
K_{app} &= \frac{MD^2}{D_{app}^2} \left(\frac{4!}{4!} \sum_{i=1}^3 n_i^4 W_{iiii} + \frac{4!}{3!} \sum_{i=1}^3 \sum_{\substack{j=1 \\ j \neq i}}^3 n_i^3 n_j W_{iiij} + \frac{4!}{2!2!} \sum_{i=1}^3 \sum_{\substack{j=1 \\ j \neq i}}^3 n_i^2 n_j^2 W_{iijj} \right. \\
&\quad \left. + \frac{4!}{2!} \sum_{i=1}^3 \sum_{\substack{j=1 \\ j \neq i}}^3 \sum_{\substack{k=1 \\ k \neq i \\ k \neq j}}^3 n_i^2 n_j n_k W_{iijk} \right) \\
&= \frac{MD^2}{D_{app}^2} \left[n_1^4 W_{1111} + n_2^4 W_{2222} + n_3^4 W_{3333} \right. \\
&\quad + 4(n_1^3 n_2 W_{1112} + n_1^3 n_3 W_{1113} + n_1 n_2^3 W_{2221} + n_2^3 n_3 W_{2223} + n_1 n_3^3 W_{3331} + n_2 n_3^3 W_{3332}) \\
&\quad + 6(n_1^2 n_2^2 W_{1122} + n_1^2 n_3^2 W_{1133} + n_2^2 n_3^2 W_{2233}) \\
&\quad \left. + 12(n_1^2 n_2 n_3 W_{1123} + n_1 n_2^2 n_3 W_{1223} + n_1 n_2 n_3^2 W_{1233}) \right]. \tag{5.4}
\end{aligned}$$

The symmetric properties of KT (W_{ijkl}) results in 15 unique tensor elements. In order to construct KT and estimate K_{app} , diffusion has to be measured along at least 15 different gradient directions. The diffusion tensor DT has 6 unique tensor

elements. The number of diffusion measurements must therefore be greater than or equal to 21 ($6 + 15$) in order to uniquely define the diffusion and kurtosis tensor in DKI.

5.4 DKI parameters

Both diffusivity and kurtosis parameters are estimated in DKI. The DKI parameters are extracted from the kurtosis tensor W_{ijkl} . Similar to the DTI parameters, the DKI parameters are rotation invariant and can be presented in DKI parametric maps where the pixel intensity signifies the parameter value.

5.4.1 Mean kurtosis

The mean kurtosis (MK) is the average diffusion kurtosis along all diffusion directions, given by

$$\text{MK} = \frac{1}{n} \sum_{i=1}^n (K_{app})_i \quad (5.5)$$

where n is the number of gradient directions, and K_{app} is the apparent kurtosis along a certain direction, given by equation (5.4). The MK is a measure of the overall kurtosis and is believed to be proportional to the heterogeneity and complexity of the tissue microstructure. Higher MK values are typically found in white matter, due to the higher degree of restriction and diffusion complexity. An increased MK may indicate an environment with more densely packed cells, which is common in cancer tumors. A decreased MK however, may indicate a loss of cellular structure. An example of an MK map is given in figure 5.5a.

5.4.2 Axial kurtosis

Axial kurtosis (AK, or K_{\parallel}) is the kurtosis along the principle diffusion tensor eigenvector (ϵ_1),

$$\text{AK} = K_1 \quad (5.6)$$

To calculate the value of K_1 , the kurtosis tensor first have to be transformed from

the Cartesian coordinate system to the coordinate system defined by the three eigenvectors of the diffusion tensor,

$$\hat{W}_{ijkl} = \sum_{i'=1}^3 \sum_{j'=1}^3 \sum_{k'=1}^3 \sum_{l'=1}^3 e_{i'i} e_{j'j} e_{k'k} e_{l'l} W_{i'j'k'l'} \quad (5.7)$$

with a rotation $x = Pn$, where P is the 3D rotation matrix, and e_{ij} are the elements of P . The value of K_1 can then be calculated by the following general equation, which expresses the diffusion kurtosis along the i th diffusion tensor eigenvector,

$$K_i = \frac{MD^2}{\lambda_i^2} \cdot \hat{W}_{iiii}, \quad i = \{1, 2, 3\} \quad (5.8)$$

where λ_i is the i th DT eigenvalue. The axial kurtosis is usually low in white matter because the main diffusion direction is along the axial direction of the diffusion ellipsoid. Water molecules diffuse more freely and unrestricted in this direction. This leads to a diffusion probability with a small deviation from the Gaussian distribution, hence a small AK. An AK map is shown in figure 5.5b.

5.4.3 Radial kurtosis

The radial kurtosis (RK, or K_{\perp}) is the kurtosis perpendicular to the principal diffusion eigenvector of the diffusion ellipsoid and is defined as

$$RK = \frac{K_2 + K_3}{2}. \quad (5.9)$$

The radial kurtosis is typically high in white matter, because diffusion in the radial direction is often restricted, leading to a non-Gaussian diffusion distribution of water molecules. It is therefore the radial kurtosis that contributes to the overall high MK values in white matter. Evaluation of RK may reveal the integrity of the nerve cell membranes and myelin sheaths. An RK map is shown in figure 5.5c.

5.4.4 Fractional anisotropy of kurtosis

The fractional anisotropy of kurtosis (FA_K) is a measure of the anisotropy of the kurtosis distribution, and can be calculated in the same manner as the fractional anisotropy of diffusivity (FA) in DTI, defined as

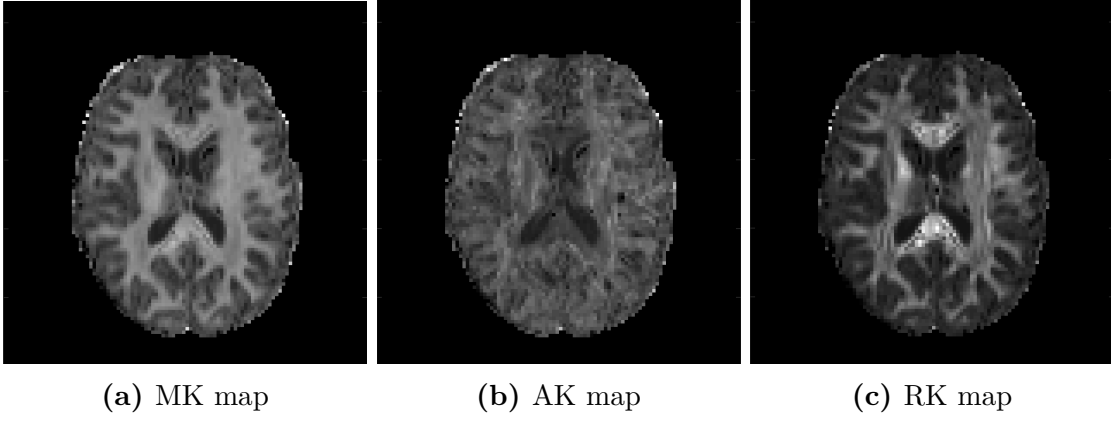


Figure 5.5: DKI-derived parametric maps. Note that AK is generally lower than RK in white matter. Water diffuses more freely in the axial direction than in the radial direction, leading to a low diffusion kurtosis in the axial direction and hence a low AK.

$$FA_K = \sqrt{\frac{3}{2}} \sqrt{\frac{(K_1 - \bar{K})^2 + (K_2 - \bar{K})^2 + (K_3 - \bar{K})^2}{K_1^2 + K_2^2 + K_3^2}} \quad (5.10)$$

where

$$\bar{K} = \frac{1}{3} \sum_{i=1}^3 K_i \quad (5.11)$$

is the average of the three directional kurtoses K_1 , K_2 and K_3 . \bar{K} is not the same as the mean kurtosis (MK), as it only takes the average of the three principal kurtoses. The fractional anisotropy of kurtosis will not be considered in this project, however.

Part III

Materials and methods

Chapter 6

Experimental data

Real diffusion-weighted signals were extracted from DW-MRI acquisitions in three healthy volunteers and used as a baseline for the signal generation. The signals were extracted from voxels located in three different regions of the brain; white matter, gray matter and the cerebrospinal fluid. One voxel was selected from each region for each of the subjects. Hence a total of nine voxels provided the basis for the signal generation. The following sections describe the details about the image acquisition, preprocessing of data and the selection of the voxels used in the analysis, and covers:

1. Extraction of baseline MRI diffusion signals
2. Diffusion MRI data
3. Data preprocessing
4. Slice selection
5. Tissue segmentation
6. Voxel selection

The images were preprocessed in FSL (FMRIB Software Library v6.0), a comprehensive library of analysis tools for FMRI, MRI and DTI brain imaging data¹. Matlab was used for the tissue segmentation and voxel selection.

6.1 Extraction of baseline MRI diffusion signals

In theory, synthetic MRI signals could be generated based on the mathematical expression of the diffusion model, given predefined parameters.

¹

<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL>

For the DTI model, a signal generation based on the mathematical expression is straightforward. The 3D geometry of the diffusion tensor (DT) is intuitive and the eigenvalues and eigenvectors of the tensor have a direct physical relevance to the diffusion processes. The tensor elements could therefore be user-defined to simulate a certain diffusion distribution.

For the DKI model, the signal generation proved to be more challenging. The derivation of the kurtosis term K_{app} requires a 4th order kurtosis tensor (KT), containing 15 independent elements. The KT is also needed when parameters like MK, RK and AK are calculated. The tensor elements could in theory be user-defined to resemble a certain diffusion kurtosis distribution. Nevertheless, the determination of the tensor elements is not straightforward, due to the mathematical complexity of KT. As of today, a direct interpretation of the individual elements of KT has yet to be explored[23].

Because of the challenges of generating a realistic KT from scratch, a high quality diffusion-weighted MRI data set was imported into Matlab and used as a baseline for the signal generation. This was done with the intention of minimizing the difference between the computer simulations and real-life diffusion signals. A few voxels that represented white and gray matter as well as CSF were selected from the real DWI data set, and the signal from these voxels were extracted. Diffusion and kurtosis tensors could then be estimated and used in the generation of the synthetic DKI signal.

To ensure consistency, the same approach was used for the DT in the DTI model. The simulated DTI signals were therefore generated based on DTs found in real DWI signals.

6.2 Diffusion MRI data

Diffusion-weighted images covering the whole brain (76 axial slices) from three healthy, young adults were used in this project. The scans were performed on a 3.0 Tesla Siemens Prisma scanner in connection with an ongoing sleep-deprivation study at Oslo University Hospital, Rikshospitalet. Images for each subject were obtained using a full-brain multishell Stejskal-Tanner sequence with single shot SE-EPI. Diffusion was measured along N_{grad} equally spaced directions for 4 b -values (500, 1000, 2000 and 3000 s/mm²). Six b_0 images were also obtained. The gradient directions were noncollinear and isotropically distributed in space to obtain an optimal sampling of diffusion. The 3D distribution of the gradient directions is plotted in figure 6.1. Details about the scan protocol can be found in table 6.1.

Table 6.1: Sequence parameters, b -values and number of gradient directions N_{grad} used per b -value in the image acquisition.

Number of slices	76
TR [ms]	2450
TE [ms]	85
FoV [mm ²]	212×212
Slice thickness [mm]	2
Matrix size	106×106
Voxel size [mm ³]	2×2×2
Flip angle (α)	78°
Multiband (MB) acceleration factor	4
Phase encode direction	anterior-posterior
Total scan time	8 min 31 sec
b-value [s/mm²]	N_{grad}
0	6
500	12
1000	30
2000	40
3000	50

6.3 Data preprocessing

The image data was preprocessed prior to the analysis². The technique involves several methods to remove noise and artefacts, according to the following steps:

1. **Noise correction:** A LPCA (local principal component analysis) noise filter was applied to reduce the random noise in the images. The denoising filter improves data quality by taking into account the multicomponent nature of multi-directional DWI datasets, thereby increasing the SNR. The denoise algorithm is described in Manjón *et al.*, 2013[28].
2. **Gibbs-ringing correction:** A truncation in k-space gives rise to Gibbs-ringing artefacts. The tool for removal of the Gibbs-ringing artefact is based on the algorithm described in Kellner *et al.*, 2016[29].
3. **Correction of susceptibility induced distortions:** The diffusion weighted SE-EPI sequence is very sensitive to non-zero off-resonance fields caused by susceptibility variations and eddy currents. Diffusion weighted images are normally acquired using the same phase encode direction. A separate b_0 image is therefore acquired using the opposite phase encode direction. This b_0 image, along with the b_0 image from the DWI set, creates an image pair with distortions going in opposite directions. The FSL tool `topup`[30] estimates and

²The images were preprocessed by supervisor Tuva Hope.

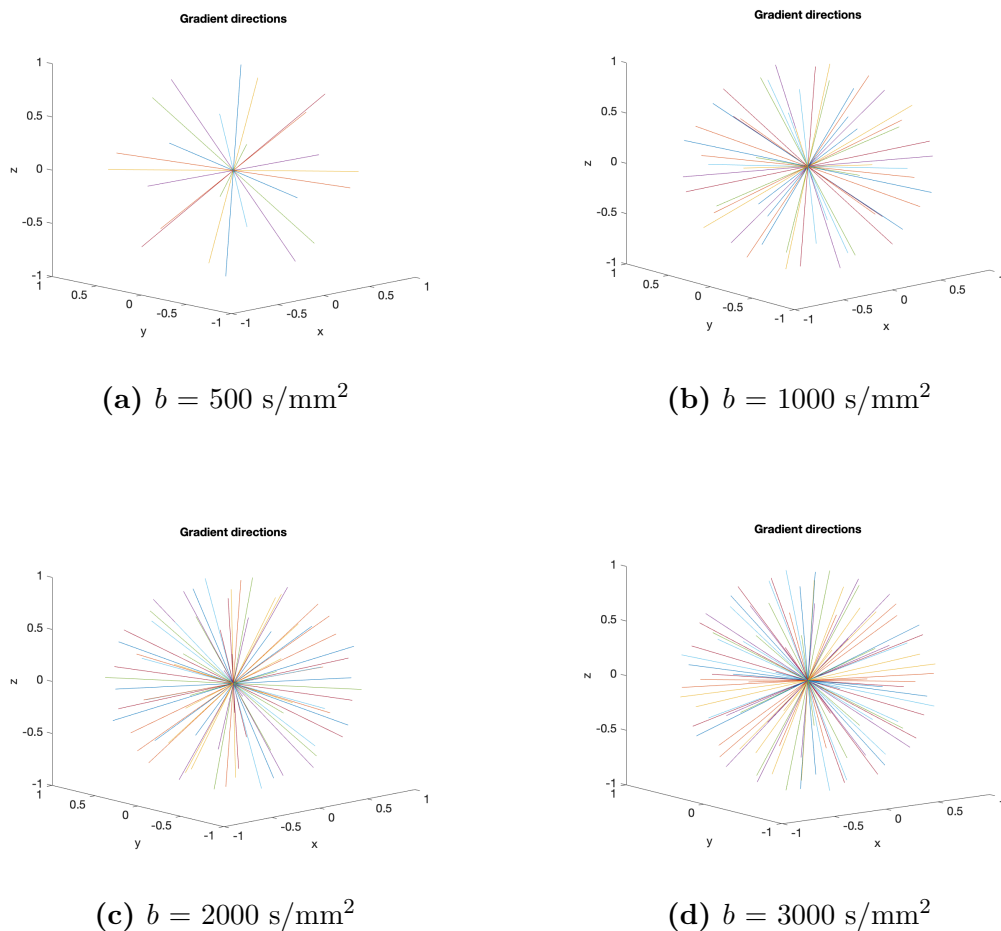


Figure 6.1: The coordinates of the diffusion gradients plotted as vectors in space, illustrating the isotropic distribution of gradient directions for the DWI data sets used as basis for the simulations. The number of gradient directions typically increases with the b -value.

corrects for the susceptibility induced field using the b_0 image pair. Based on the correction of the image pair, `toptup` correction is applied to the rest of the DWIs, creating a corrected multishell DWI set. The susceptibility induced fields are estimated using a method similar to that described in Andersson *et al.*, 2003[31].

4. **Correction of motion induced artefacts and eddy currents:** The FSL tool `eddy` corrects for eddy current-induced distortions and subject movements in the image data. `eddy` distinguishes between signal variations caused by diffusion and by eddy currents or subject movements. The method is described in detail in Andersson *et al.*, 2006[32].

6.4 Slice selection

Mainly two types of tissue were of interest in this project; white matter (WM) and gray matter (GM). In addition to those, the properties of CSF were also investigated. One slice from each volume was selected based on the amount of WM, GM and CSF present in the slice. The most interesting slices were located in the middle of the brain, where large white matter fiber tracts, as well as areas GM and CSF were visible. The slices used in this project are illustrated as DWIs³ along with FA and MD maps in figure 6.2.

6.5 Segmentation

The diffusion patterns of WM, GM and CSF are all different, and the degree of anisotropy can be described by the FA value. The FA values measured in WM are typically much higher than in GM, while in the CSF, where diffusion is free, the value should be close to zero.

The brain tissue in each of the selected slices was therefore segmented into two groups based on the FA value of the voxels (threshold value defined below). Voxels with small FA values were sorted into segment 1, while voxels with large FA values were sorted into segment 2. Segment 1 could then represent GM and CSF, and segment 2 represent WM.

A threshold value had to be established in order to define the upper and lower limits of FA in the two segments. The limits were determined by calculating the 90th and 10th percentile of the FA values of the whole brain. When sorting all the FA values in ascending order, the 10th percentile marks the upper limit of the 10% smallest FA values, and the 90th percentile marks the lower limit of the 10% largest FA values⁴. Voxels with FA values lower than the 10th percentile would then belong to segment 1, while voxels with FA values higher than the 90th percentile would belong to segment 2.

The voxels in the two segments are visualized in figure 6.3. This visualization confirms that voxels with higher FA values are mainly located in the WM, while the voxels with lower FA values are found in the GM and CSF regions of the brain. Histograms of the FA and MD values for the two segments are illustrated in figure 6.4.

³Visualization provided by Horos, an open source medical image viewer.
<https://horosproject.org/>

⁴Details can be found in the Matlab script `percentile.m`. A code snippet is provided in Appendix B.5

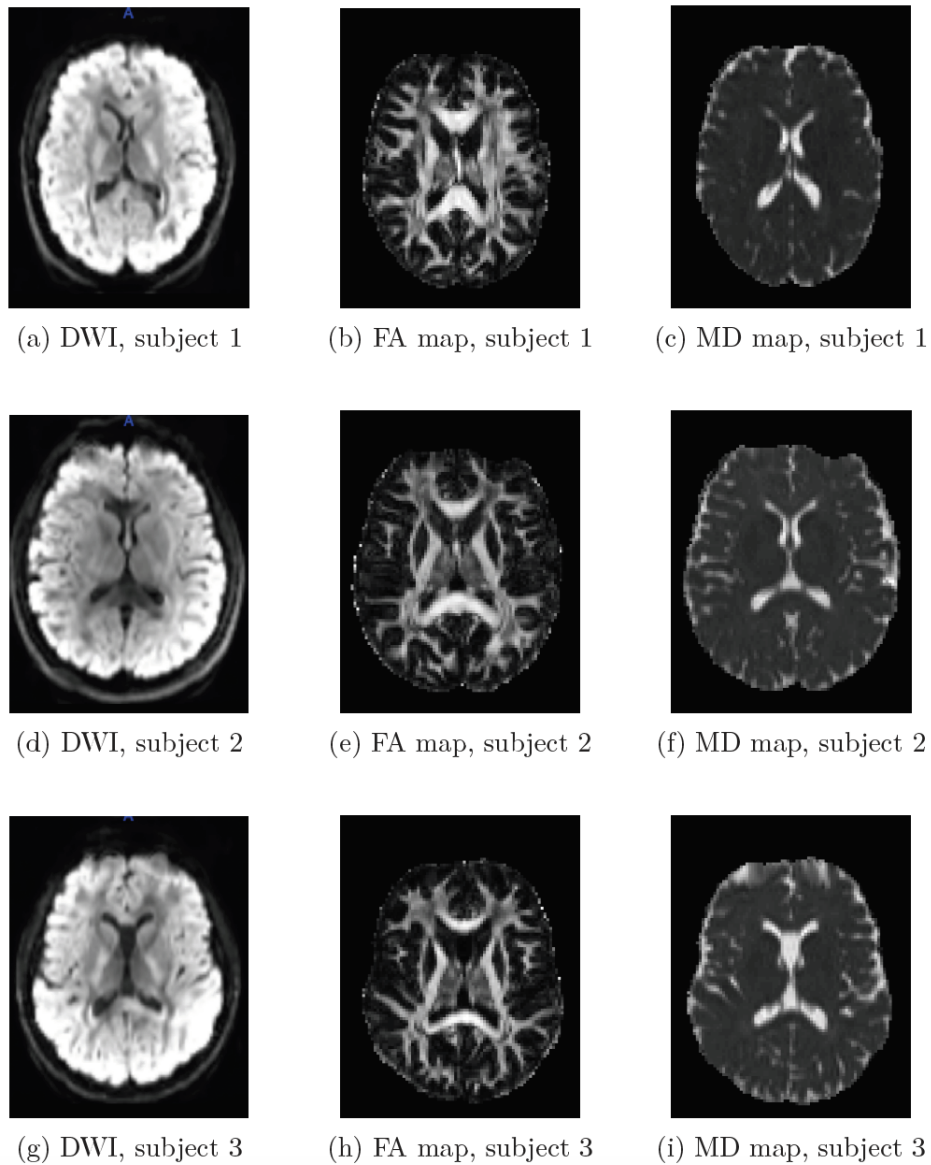
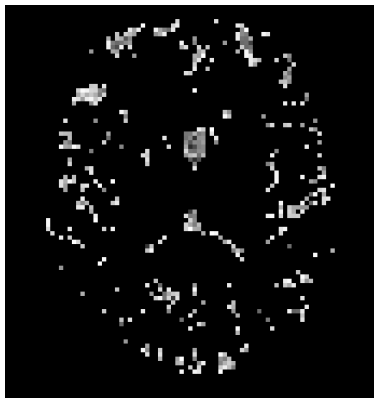


Figure 6.2: The slices were selected based on the content of WM, GM and CSF. Here the DWIs are presented with diffusion-weighting $b = 1000 \text{ s/mm}^2$ in one direction. The FA map clearly visualizes the WM fiber tracts, while the MD map provides visualization of the CSF in the ventricles and around the brain.

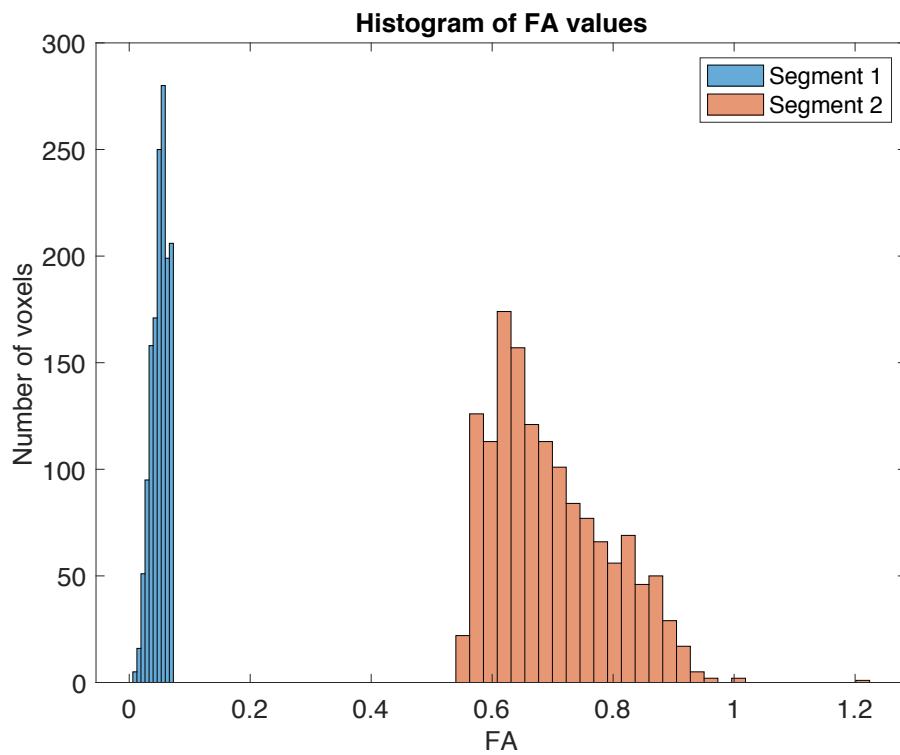


(a) Segment 1

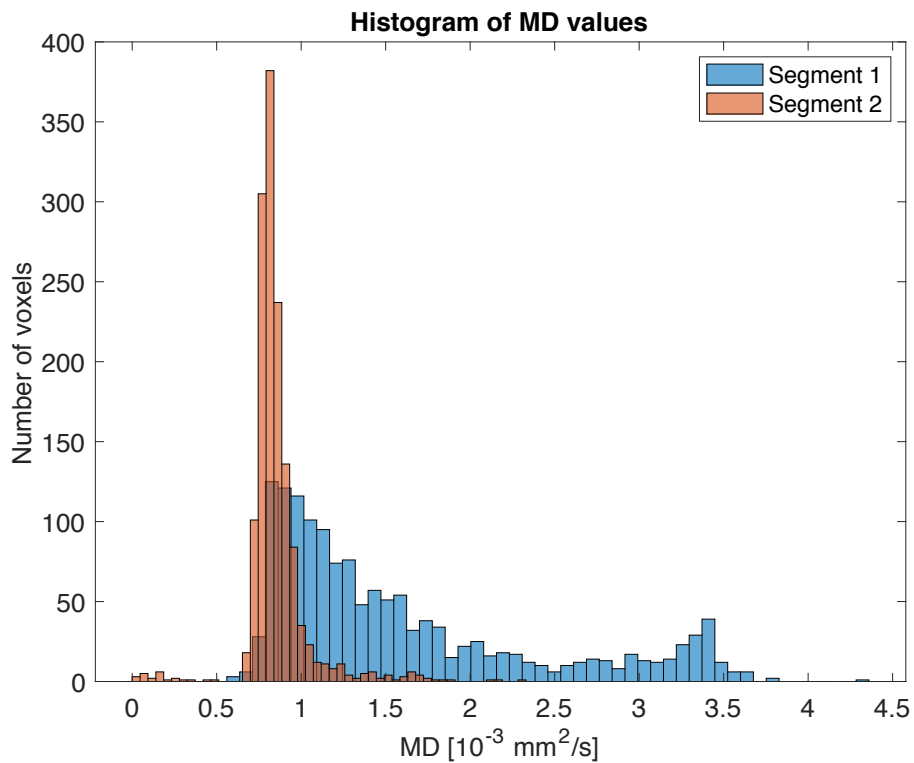


(b) Segment 2

Figure 6.3: The tissue was segmented into two groups based on the voxel's FA value. Segment 1 contains pixels with FA values smaller than the 10th percentile and could be classified as GM or CSF, while segment 2 contains pixels with FA values higher than the 90th percentile, classified as WM.



(a) FA



(b) MD

Figure 6.4: Histograms illustrating the frequency distribution of FA and MD values in the two segments. Segment 1 represents gray matter and CSF and segment 2 represents white matter. Per definition, there is a large difference in FA values in the two segments. The MD values however show a slight overlap between the two segments.

6.6 Voxel selection

The purpose of dividing the brain into two segments was to make the process of selecting ideal voxels from each tissue type more convenient. Two voxels were selected from segment 1, and one voxel was selected from segment 2. This resulted in three voxels per subject; one in white matter (voxel 1), one in gray matter (voxel 2) and one in the CSF (voxel 3). The reason for selecting single voxels instead of a larger group of voxels in a region of interest (ROI), was to ensure consistency in the orientation of the diffusion tensor. It was discovered that even in smaller ROIs, there was a noticeable variation between neighboring voxels. The selected voxels are presented in figure 6.5, and the extracted signals are shown in figure 6.6 for one diffusion direction.

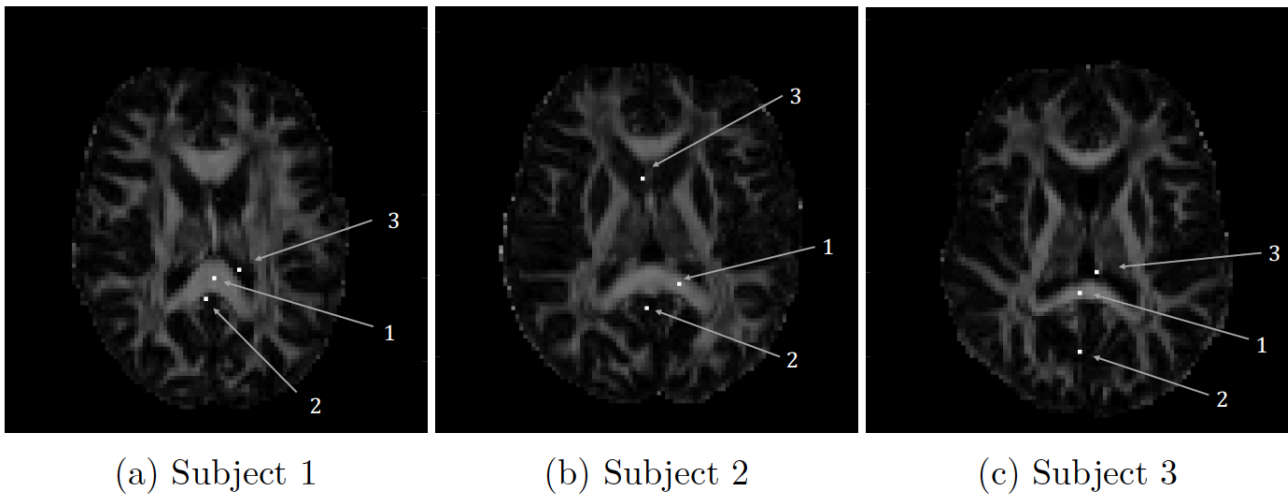


Figure 6.5: Three voxels were selected from each subject, each representing a different type of tissue. Voxel 1 is white matter, voxel 2 is gray matter, and voxel 3 is CSF.

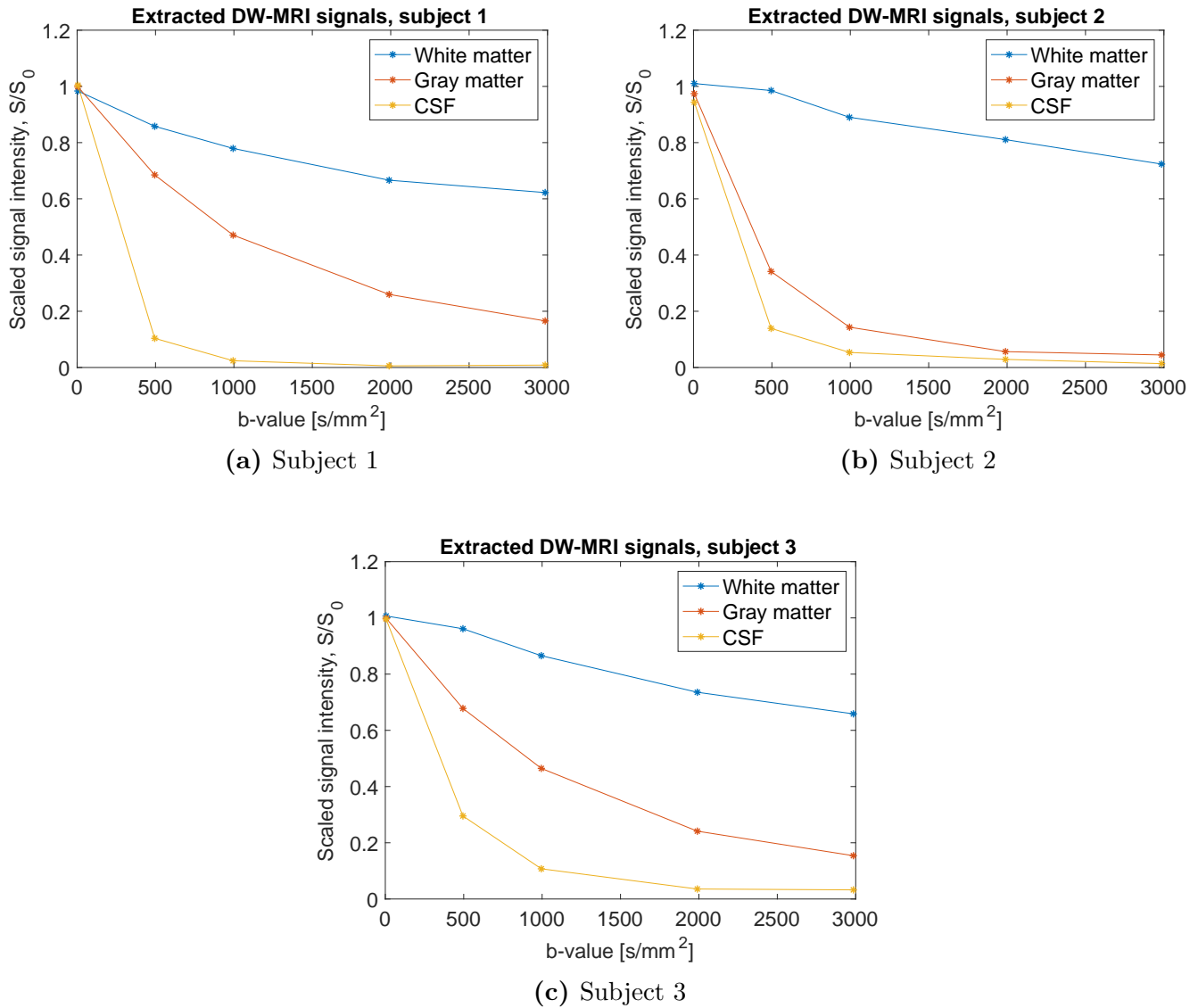


Figure 6.6: DW-MRI signals from the selected voxels in one diffusion direction. For each subject, signals were extracted from white matter, gray matter and CSF.

Chapter 7

Simulations

DTI and DKI signals were generated based on real DWI signals extracted from DW-MRI acquisitions in healthy volunteers. The generated signals were used as baseline signals in the simulations. The following sections describe the pipeline of the simulation process, and covers:

1. Tensor estimation
2. Signal generation
3. Establishing ground truth
4. Adding synthetic noise
5. Calculation of diffusion parameters
6. Statistical analysis

In addition, completely synthetic DTI signals were generated based on diffusion tensors extracted from the real DWI signals by using the mathematical expression of the DTI model. All computer simulations were executed in Matlab.

7.1 Tensor estimation

For the DKI analysis, DT and KT were calculated for each voxel using the Matlab code `dki_fit`[33], developed by Veraart and colleagues at the MRI Biophysics group at the New York University School of Medicine¹. `dki_fit` estimates the tensor elements of DT and KT using constrained weighted linear least squares estimation (WLS) (weights given by equation (4.34)). WLS is the method of choice because

¹More information can be found at:
<https://github.com/NYU-DiffusionMRI/DESIGNER>
Code snippet is provided in Appendix B.1

it is better at modeling signals containing noise than the OLS method, where all measurements are assumed to be equally precise. For the DTI analysis, a modified version of `dki_fit` was developed, namely `dti_fit`², which do not take the kurtosis effects into consideration. The input data included:

- **Diffusion-weighted images:** The DWIs were preprocessed prior to the tensor fitting as outlined in chapter 6.
- **Diffusion encoding information:** All gradient directions and b -values (for each specific analysis) which were used for the tensor fitting.
- **Brain mask:** A mask is a binary matrix with elements that are either zero or one. The matrix must have the same dimension as the diffusion image, where each matrix element corresponds to a specific voxel in the diffusion image. If the matrix element has the value zero, the corresponding voxel is ignored. By providing a mask, the calculations are limited to a user-defined region of interest (ROI). A considerable amount of time can be saved if the voxels outside of the brain are neglected, as these only contain background noise. A brain mask was therefore generated in FSL.
- **Constraints:** An array which imposes a user-defined constraint to the WLS estimation of the tensors. The constraints are given as a boolean, $[c1, c2, c3]$, where

```
c1: Dapp > 0
c2: Kapp > 0
c3: Kapp > b/(3*Dapp)
```

Although a negative K_{app} is mathematically possible, empirical evidence indicates that the kurtosis (and diffusivity) is always non-negative in biological tissue[34]. The constraints were therefore set to $[1 \ 1 \ 1]$ for the DKI analysis and $[1 \ 0 \ 0]$ for the DTI analysis, as it does not consider kurtosis effects.

- **Maximum b -value:** A scalar which puts an upper bound on the b -values being used in the analysis. The value was set to 3000 s/mm² for the DKI analysis, and 1000 s/mm² for the DTI analysis, as this is considered optimal for the DTI model[35].

7.2 Signal generation

The elements of the diffusion and kurtosis tensor were estimated using `dki_fit` and `dti_fit` for the DKI and DTI model, respectively. The scripts use a WLS approximation given by equation (4.32), where the weights are given by equation

²Code snippet is provided in Appendix B.2

(4.34). Baseline signals were then generated based on the estimated diffusion and kurtosis tensors through the expression

$$\hat{S} = \exp(-b \cdot dt) \quad (7.1)$$

where b is a matrix containing all the b -values and gradient directions and dt is a vector containing all the unique tensor elements. For the DTI model, dt contains 6 unique elements, and (7.1) is equivalent to calculating the DTI signal given by equation (4.7). For the DKI model, dt contains 21 unique elements, and the logarithm of (7.1) is equivalent to the DKI signal given by equation (5.2).

7.3 Establishing ground truth

Diffusion and kurtosis tensors were estimated from the generated signals using `dki_fit` and `dti_fit`. 6 b_0 -images, 12 $b = 500$ and 30 $b = 1000$ s/mm² measurements were used in estimating of ground truth for the DTI model, while 6 b_0 -images, 12 $b = 500$, 30 $b = 1000$, 40 $b = 2000$ and 50 $b = 3000$ s/mm² were used for DKI, resulting in a total of 48 images for DTI and 138 images for DKI.

The DKI- and DTI-derived diffusion parameters were then calculated from the tensor elements using the Matlab scripts `dki_paramters`[36] and `dti_parameters`, respectively³. An overview of the ground truth parameter values for the baseline signals are given in table 7.1.

7.4 Adding synthetic noise

Synthetic noise was added to the baseline signals after the ground truth values had been established. Conventional MR images have a Rician noise distribution[37]. Synthetic noise was therefore incorporated by adding Gaussian noise to the signal S , through the expressions

$$S_r = S + N_r \quad (7.2)$$

and

$$S_i = N_i \quad (7.3)$$

³Code snippets can be found in Appendix B.3 and B.4

Table 7.1: Table showing the ground truth values for the baseline signals. MD, RD and AD are given in 10^{-3} mm²/s.

Subject 1						
Diffusion model	DKI			DTI		
Voxel	1	2	3	1	2	3
Tissue type	WM	GM	CSF	WM	GM	CSF
FA	0.8974	0.1497	0.1284	0.8730	0.1435	0.0883
MD	0.7629	0.7900	3.3811	0.7384	0.7254	3.2055
RD	0.1823	0.7327	3.1306	0.2078	0.6771	3.0466
AD	1.9241	0.9047	3.8820	1.7996	0.8221	3.5233
MK	1.1866	0.6263	0.2975			
RK	4.3780	0.6408	0.3200			
AK	0.5266	0.6001	0.2580			
Subject 2						
Diffusion model	DKI			DTI		
Voxel	1	2	3	1	2	3
Tissue type	WM	GM	CSF	WM	GM	CSF
FA	0.8922	0.1423	0.1127	0.8897	0.1251	0.1176
MD	0.7644	1.9270	2.6081	0.6981	1.7760	2.8186
RD	0.1937	1.7736	2.4531	0.1796	1.6534	2.6370
AD	1.9059	2.2338	2.9181	1.7352	2.0212	3.1818
MK	0.9876	0.5156	0.3854			
RK	2.3172	0.5550	0.4088			
AK	0.4541	0.4460	0.3433			
Subject 3						
Diffusion model	DKI			DTI		
Voxel	1	2	3	1	2	3
Tissue type	WM	GM	CSF	WM	GM	CSF
FA	0.9084	0.1596	0.1371	0.9237	0.1557	0.0973
MD	0.7019	0.7782	2.5100	0.6223	0.7184	2.1814
RD	0.1531	0.7128	2.3390	0.1170	0.6590	2.0821
AD	1.7995	0.9089	2.8519	1.6331	0.8372	2.3799
MK	0.8148	0.4883	0.4009			
RK	0.0784	0.5052	0.4295			
AK	0.4415	0.4119	0.3513			

where N_r and N_i are random numbers generated from a Gaussian distribution with zero mean and standard deviation $\sigma = S_0/\text{SNR}$. Rician noise could then be added to the baseline signals though the following expression

$$\tilde{S} = \sqrt{S_r^2 + S_i^2} \quad (7.4)$$

where \tilde{S} is the noise-incorporated signal. The SNR level was varied from 10 to 100 in 10 steps; 10, 15, 20, 25, 30, 40, 50, 60, 80 and 100.

The SNR is defined as the ratio of the average signal value in a selected region-of-interest (ROI) to the standard deviation of the background noise,

$$\text{SNR} = \frac{\text{signal}}{\text{noise}} = \frac{\mu}{\sigma} \quad (7.5)$$

where μ is the signal mean and σ is the standard deviation of the background noise. However, it is important to stress that the baseline signals already contained a certain level of SNR, as they were based on tensors estimated from real DW-MRI signals that were not completely noise-free (in practice, real signals can never be completely noise-free). At higher signal intensities, the Rician noise distribution approaches a Gaussian distribution[38], and the process of adding noise to a signal that already contains some level of noise simply results in an extra noise contribution to the signal S . The signal with added noise can then be expressed as

$$S_{noise} = S + \sigma_{measured} + \sigma_{added} \quad (7.6)$$

where S is the DW-MRI signal, $\sigma_{measured}$ is the noise measured in the DWI, and σ_{added} is the added noise. On average, the standard deviation of the background noise in the b_0 -images were $\sigma_{measured} \approx 3,7$, with an SNR of approximately 532, 836 and 1253 in white matter, gray matter and CSF, respectively. However, since the noise has a Rician distribution at lower SNR, equation (7.6) should be seen as an approximation.

7.5 Calculation of diffusion parameters

For the DKI analysis, DT and KT were again estimated using `dki_fit`, only this time from the noise-corrupted signals \tilde{S} . The DKI-derived parameters were then calculated using `dki_parameters`.

The DTI analysis used `dti_fit` in the estimation of DT, and the DTI-derived parameters were calculated using `dti_parameters`.

Each parameter was estimated $n = 1000$ times per voxel.

7.6 Statistical analysis

7.6.1 Outlier removal

Outliers were removed from the parameter estimation sample using Tukey's method[39]. A parameter is considered as an outlier if the value is less than

$$Q_1(\hat{\theta}_i) - 1.5Q_I(\hat{\theta}_i) \quad (7.7)$$

or greater than

$$Q_3(\hat{\theta}_i) + 1.5Q_I(\hat{\theta}_i) \quad (7.8)$$

where $Q_1(\hat{\theta}_i)$ is the first quartile, $Q_3(\hat{\theta}_i)$ is the third quartile, and $Q_I(\hat{\theta}_i)$ is the interquartile range of the estimated parameter, $\hat{\theta}_i$. A quartile divides a sorted data set into four equal parts where each part represents 1/4 of the data set. The first quartile is the same as the 25th percentile, and the third quartile is the same as the 75th percentile. The interquartile range is the difference between the first and the third quartile.

7.6.2 Relative error

The performance of the estimations was evaluated by calculating the relative error for each fitted parameter. The relative error is a measure of accuracy[37] and gives an indication of how good an estimation is relative to the ground truth. The relative error is measured in percent, and is defined as

$$\text{Relative error} = \frac{\hat{\theta}_i - \theta}{\theta} \times 100 \quad (7.9)$$

where $\hat{\theta}_i$ is the estimated parameter and θ is the ground truth.

7.6.3 Box plots

The results from the three subjects were pooled together and presented in box plots (see figure 7.1). The box plot combines the minimum and maximum values, or the

range, of the data points with the quartiles into one useful graph. The central mark in each box indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, or the interquartile range, respectively. The whiskers extend to the most extreme data points, and the outliers are plotted individually using the '+' symbol. The default whisker length extends to approximately $\pm 2.7\sigma$, which corresponds to 99.3 % coverage if the data are normally distributed.

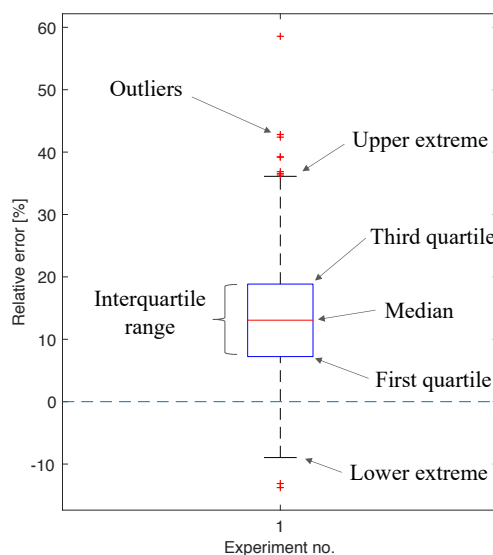


Figure 7.1: The results are presented in box plots.

7.7 Noise images

To demonstrate the effect of adding noise to the signal, parametric maps were generated for various levels of noise. Signals were estimated for each voxel in the image using `dki_fit`. Rician noise was then added to the signals, and the diffusion parameters were estimated for each voxel.

7.8 Analytic DTI signal generation

For comparison purposes, completely synthetic DTI signals were generated based on the mathematical expression of the DTI signal attenuation, using the diffusion tensors estimated from the real DW-MRI signals. By using the expression $\mathbf{Y} = \mathbf{H}\mathbf{d}$ given by (4.19), where \mathbf{H} is the matrix containing all the gradient directions, and \mathbf{d} is the vector with all the tensor elements from the DWI signal, S was calculated through the expression

$$S = S_0 \exp(-\mathbf{b} \cdot \mathbf{Y}) \quad (7.10)$$

where S_0 signal intensity at b_0 , and \mathbf{b} is the vector containing all the b -values. The signal attenuation vector, given by (4.12), could then be expressed as $Y = \frac{1}{b} \ln(\frac{S_0}{S})$.

The tensor elements were calculated using OLS (4.27), where $\mathbf{d} = \mathbf{H}^\Psi \mathbf{Y}$, and \mathbf{H}^Ψ is the pseudoinverse of \mathbf{H} , given by $\mathbf{H}^\Psi = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ in equation (4.26). A snippet of the code is provided in Appendix B.7.

The diffusion parameters were then estimated using `dti_parameters.m`, with the elements of \mathbf{d} as input.

The diffusion tensors were visualized as ellipsoids to demonstrate the degree of diffusion anisotropy. This was done by calculating the eigenvalues of the diffusion tensor using the Matlab command `eig(D)`. The eigenvalues λ_1 , λ_2 and λ_3 could then define the semi-axes length of the ellipsoid.

Chapter 8

Simulation experiments

A number of simulation experiments using different levels of SNR, number of gradient directions and b -values were executed and the results analyzed by investigating the accuracy and precision of the parameter estimation.

8.1 Parameter estimation as a function of SNR

The parameter estimations were evaluated as a function of SNR. In this analysis, 6 b_0 , 12 $b = 500$, 30 $b = 1000$, 40 $b = 2000$ and 50 $b = 3000$ s/mm² were used for the DKI parameter estimation, and 6 b_0 , 12 $b = 500$ and 30 $b = 1000$ s/mm² were used for the DTI parameter estimation. Ground truth values were calculated before increasing levels of noise was added to the signal¹.

8.2 Parameter estimation as a function of NSA

There are several parameters that may influence the SNR. The number of signal averages (NSA) is one such example. The NSA (also called the number of excitations, NEX) represents the number of times that each line in k-space is sampled, and is normally used to increase the SNR. Increasing the NSA is commonly used in clinical practice if the SNR is marginal. The cost of increasing the NSA is a longer acquisition time. The SNR scales with \sqrt{N} , where N is the number of acquisitions. Doubling the NSA only improves the SNR by $\sqrt{2}$ because the background noise is also sampled during the acquisitions.

NSA was increased from NSA = 1 to NSA = 2 to investigate the effect of multiple acquisitions. The effect of multiple signal averages was included in the simulations

¹A code snippet is provided in Appendix B.6

by repeating the acquisition loop N times. In this analysis, 6 b_0 , 12 $b = 500$, 30 $b = 1000$, 40 $b = 2000$ and 50 $b = 3000$ s/mm² were used for the DKI parameter estimation, while 6 b_0 , 12 $b = 500$ and 30 $b = 1000$ s/mm² were used for the DTI parameter estimation.

8.3 Parameter estimation as a function of number of gradient directions and b -values

The influence of the number of gradient directions, N_{grad} , per b -value shell was evaluated by the accuracy of the parameter estimations. The SNR was kept at a constant level of 50. For the DTI model, the minimum requirement for estimating the diffusion tensor is one b -value, where diffusion is measured in 6 directions (in addition to one b_0 image). For the DKI model, the minimum requirement is two b -values, measured in 6 and 15 directions, respectively (in addition to one b_0 image)[24]. Based on these requirements, a number of gradient sets were generated.

Some of the gradient sets were reduced to half its size by manually removing every second gradient direction in Matlab. The coordinates of the remaining gradients were then plotted as vectors in space (using the same method as in figure 6.1) to ensure that the distribution of gradients remained fairly isotropic and noncollinear². The gradient sets are listed in table 8.1 below.

Table 8.1: Gradient sets used in the DTI and DKI analysis. b -values are given in s/mm².

DTI	
Set name	N_{grad} per b -value
1	1 $b = 0$; 15 $b = 1000$
2	1 $b = 0$; 30 $b = 1000$
3	6 $b = 0$; 30 $b = 1000$
4	1 $b = 0$; 6 $b = 500$; 15 $b = 1000$
5	1 $b = 0$; 12 $b = 500$; 30 $b = 1000$
6	6 $b = 0$; 12 $b = 500$; 30 $b = 1000$
7	1 $b = 0$; 12 $b = 500$; 30 $b = 1000$; 50 $b = 3000$
8	6 $b = 0$; 12 $b = 500$; 30 $b = 1000$; 50 $b = 3000$
DKI	
Set name	N_{grad} per b -value
1	6 $b = 0$; 12 $b = 500$; 30 $b = 1000$
2	1 $b = 0$; 6 $b = 500$; 25 $b = 3000$
3	1 $b = 0$; 12 $b = 500$; 50 $b = 3000$
4	6 $b = 0$; 12 $b = 500$; 50 $b = 3000$
5	1 $b = 0$; 6 $b = 500$; 15 $b = 1000$; 25 $b = 3000$
6	1 $b = 0$; 12 $b = 500$; 30 $b = 1000$; 50 $b = 3000$
7	6 $b = 0$; 12 $b = 500$; 30 $b = 1000$; 50 $b = 3000$
8	6 $b = 0$; 12 $b = 500$; 30 $b = 1000$; 40 $b = 2000$; 50 $b = 3000$

²A code snippet is provided in Appendix B.8

Part IV

Results

Chapter 9

Results

A number of simulation experiments were conducted in order to study the effect of SNR, NSA, number of gradient directions and number of b -values on the diffusion parameter estimations of the DTI and the DKI model. The results were evaluated in terms of accuracy by calculating the relative error of the parameter estimation. Baseline DTI and DKI signals simulating white matter, gray matter and CSF were used in the experiments. A total of nine signals were investigated; three signals simulating white matter, three simulating gray matter and three simulating CSF.

The most relevant results are presented in the following sections. Additional results can be found in the Appendix. The results are presented as follows:

1. Results from the signal generation
2. Results from the parameter estimation as a function of SNR
3. Results from the parameter estimation as a function of NSA
4. Results from the parameter estimation as a function of number of gradient directions and b -values

9.1 Signal generation

Baseline signals were generated based on the estimated diffusion and kurtosis tensors of the extracted DW-MRI signals in the white matter, gray matter and CSF. Figure 9.1 illustrates the fitted (baseline) signals together with the actual DW-MRI signals in one gradient direction for one of the subjects.

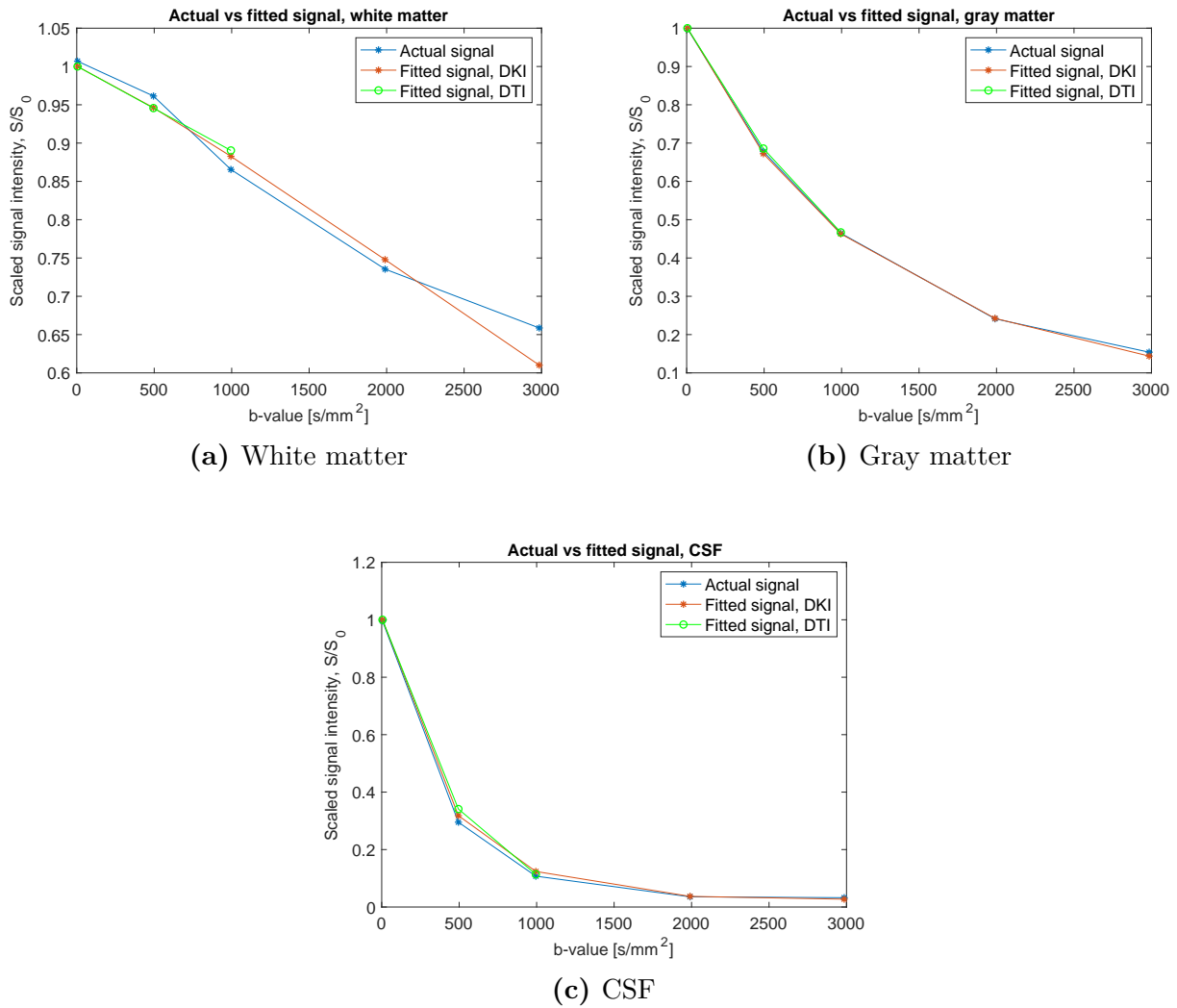


Figure 9.1: Actual DW-MRI signals vs the fitted DKI and DTI signals in the white matter, gray matter and CSF in one diffusion direction for one of the subjects. The upper bound on the DTI analysis was set to $b = 1000$ s/mm².

9.1.1 Noise images

To demonstrate the effect of adding noise to the signal, parametric maps were generated for various levels of SNR. The FA and MK maps for one of the subjects are shown in figure 9.2 and 9.3, respectively. Histograms showing the distribution of FA and MK values in the parametric maps for various noise levels are presented in figure 9.4.

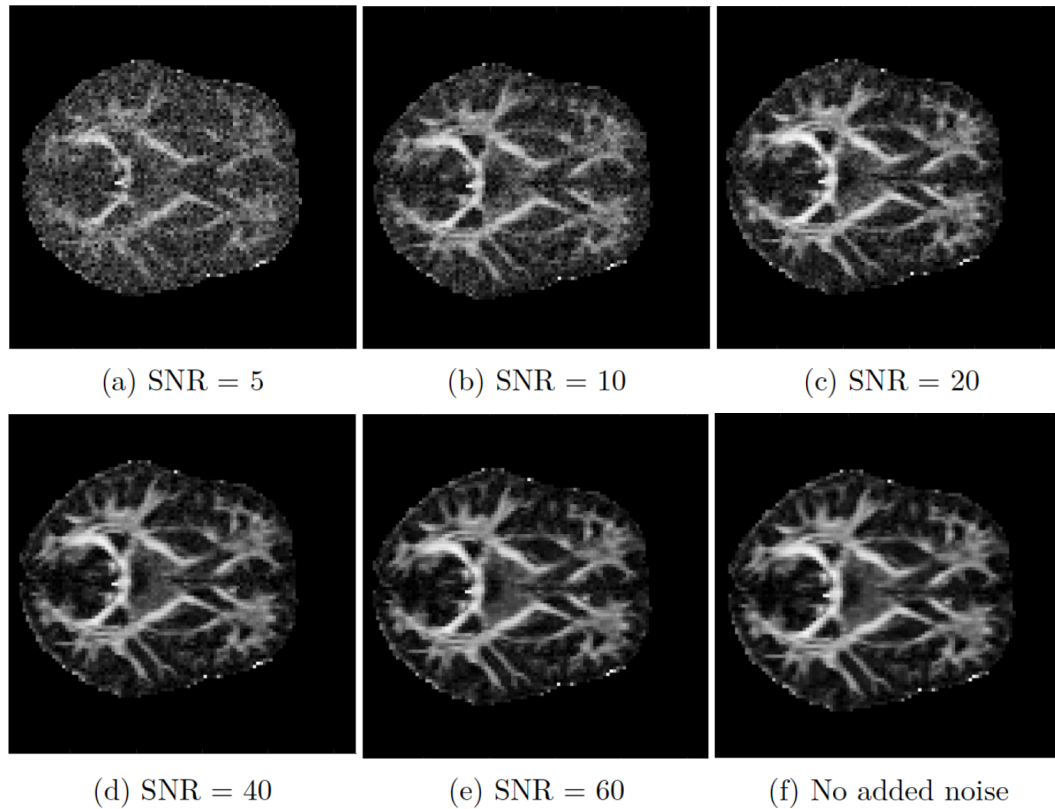


Figure 9.2: FA maps generated with various levels of noise added to the baseline signals. The FA map without added noise is presented in (f).

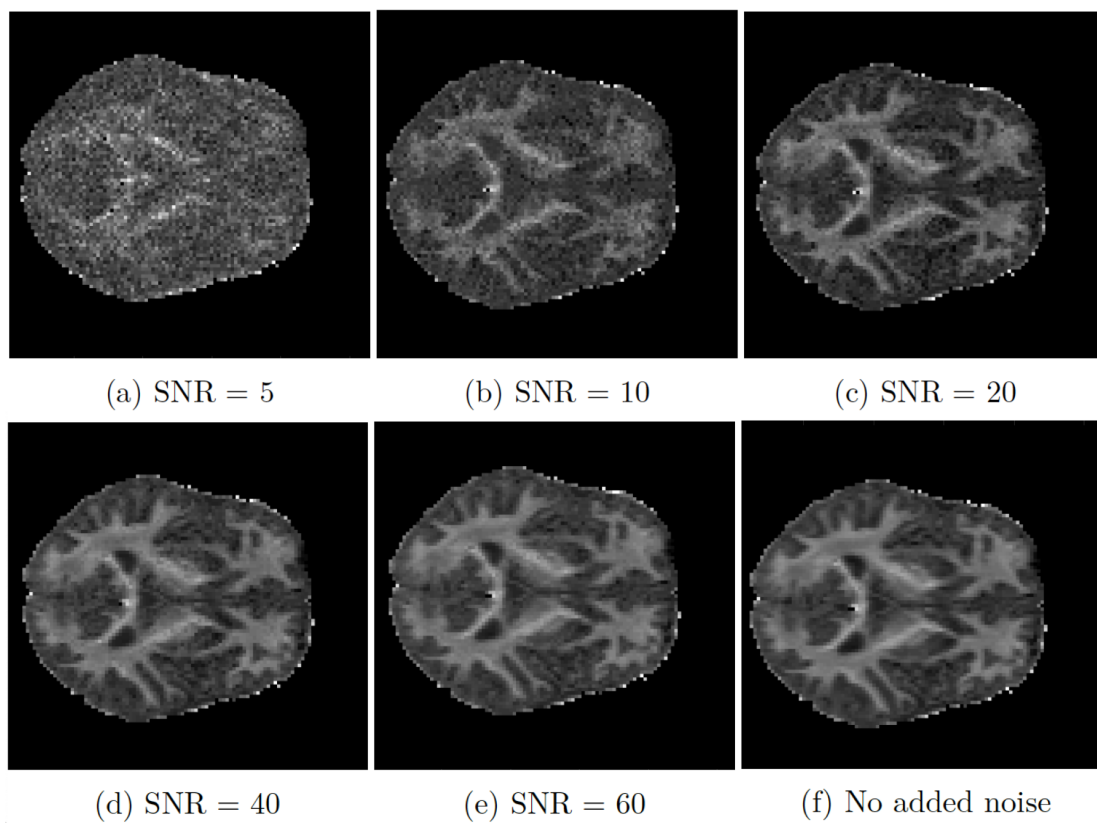


Figure 9.3: MK maps generated with various levels of added noise. (f) illustrates the MK map without any noise added to the baseline signal.

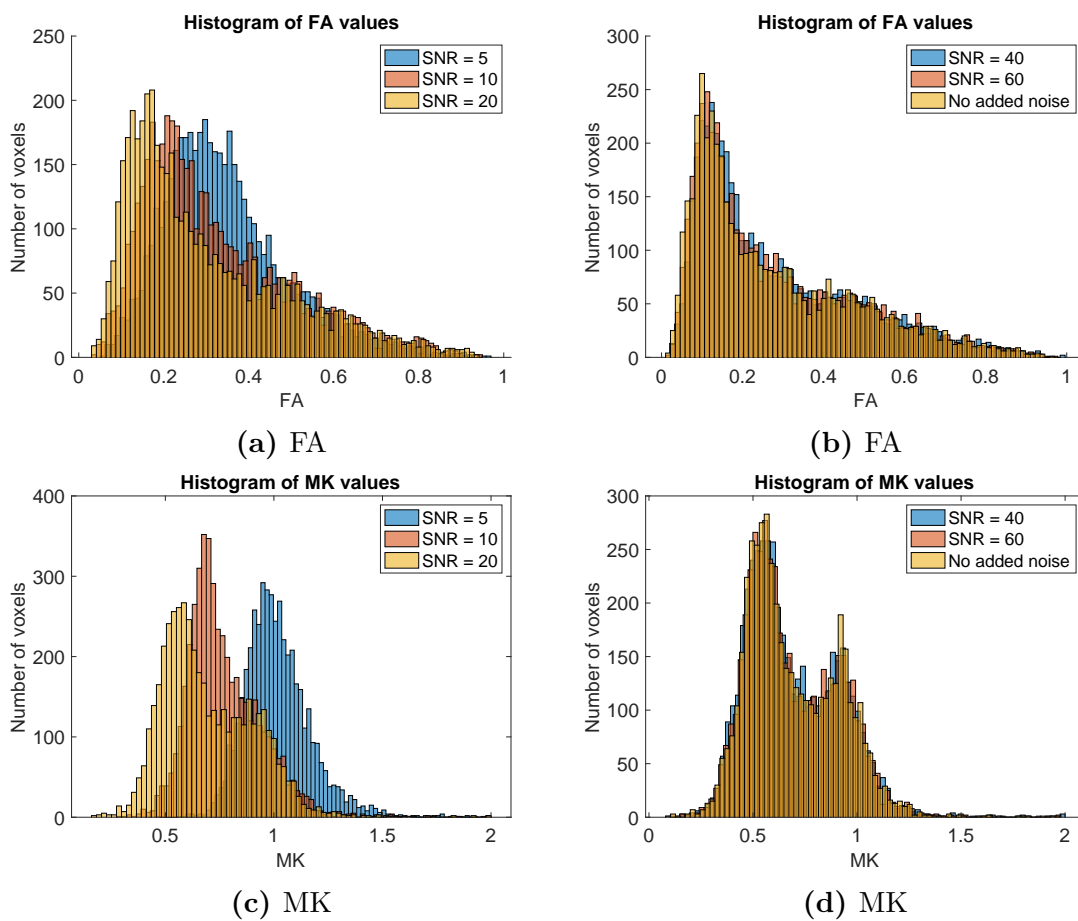


Figure 9.4: Histograms illustrating the FA and MK values for increasing levels of SNR in the whole brain.

9.1.2 Analytic DTI signal generation

Analytic DTI signals were generated based on the mathematical expression of the DTI signal attenuation, using the tensor elements of the extracted DW-MRI signals. The generated signals are plotted in figure 9.5 together with the actual DW-MRI signals and the DTI signals approximated by the WLS method. The diffusion tensors used in the signal generation are visualized as ellipsoids in figure 9.6 for one of the subjects.

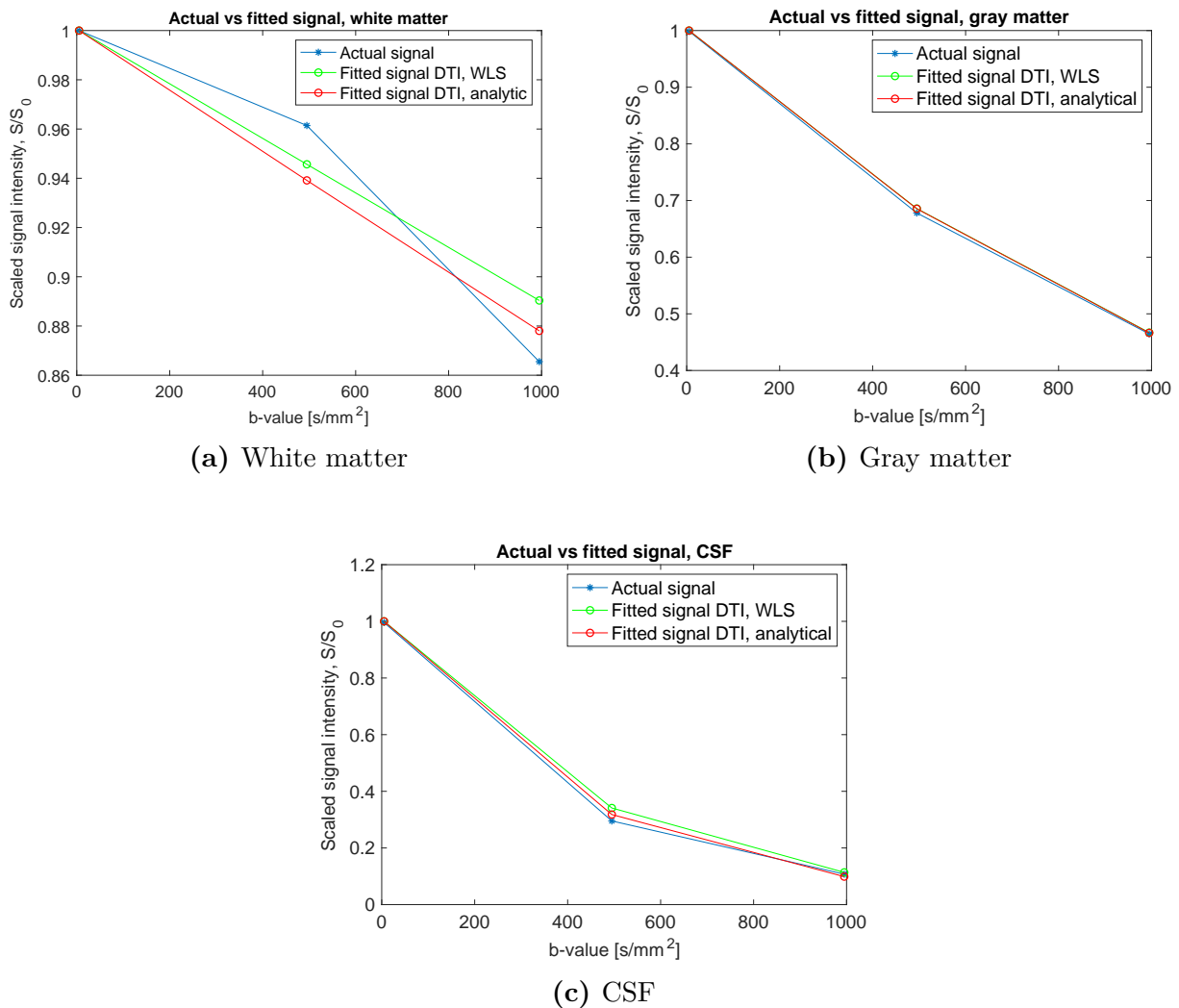


Figure 9.5: Actual DW-MRI signals vs the generated DTI signals in the white matter, gray matter and CSF in one gradient direction for one of the subjects. The signals were generated by WLS approximation and by using the mathematical expression for the DTI signal attenuation.

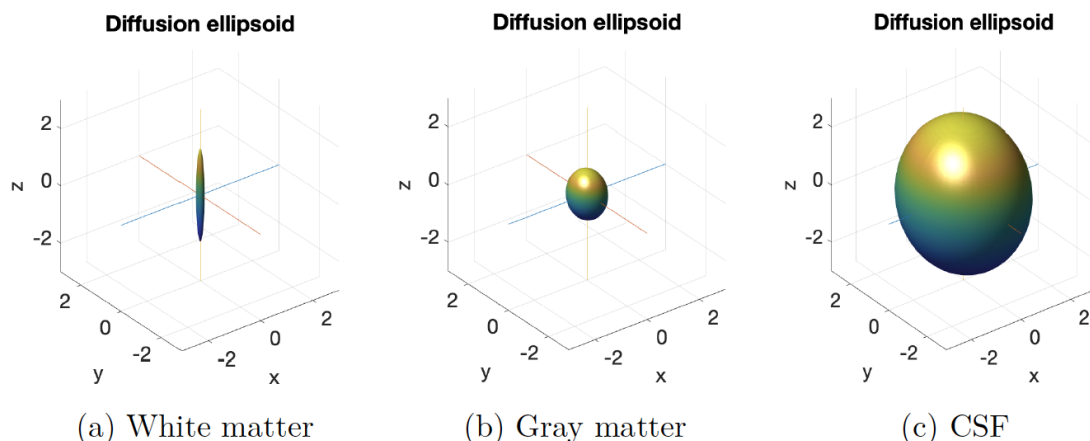


Figure 9.6: Diffusion tensors from the white matter, gray matter and CSF for one of the subjects, visualized as ellipsoids. The tensor elements were used in the generation of the analytic DTI signal.

9.2 Parameter estimation as a function of SNR

The results of the parameter estimations as a function of SNR are presented in box plots where the x -axis represents SNR and the y -axis represents the relative error of the parameter estimation. Note however that the x -axis represents the level of *added* SNR, as the baseline signals already contained a certain level of noise. A blue, horizontal line marks where the relative error is equal to zero; this is where the estimations are equal to the ground truth. The results of the DTI and DKI analyses are presented in the same box plot for FA, MD, RD and AD, as these parameters are common for both models. The DKI-derived parameters MK, RK and AK are presented in separate figures. Next follows the results of the simulations for signals representing white matter, gray matter and CSF. The results of the analytic DTI signal are included for comparison purposes.

9.2.1 White matter

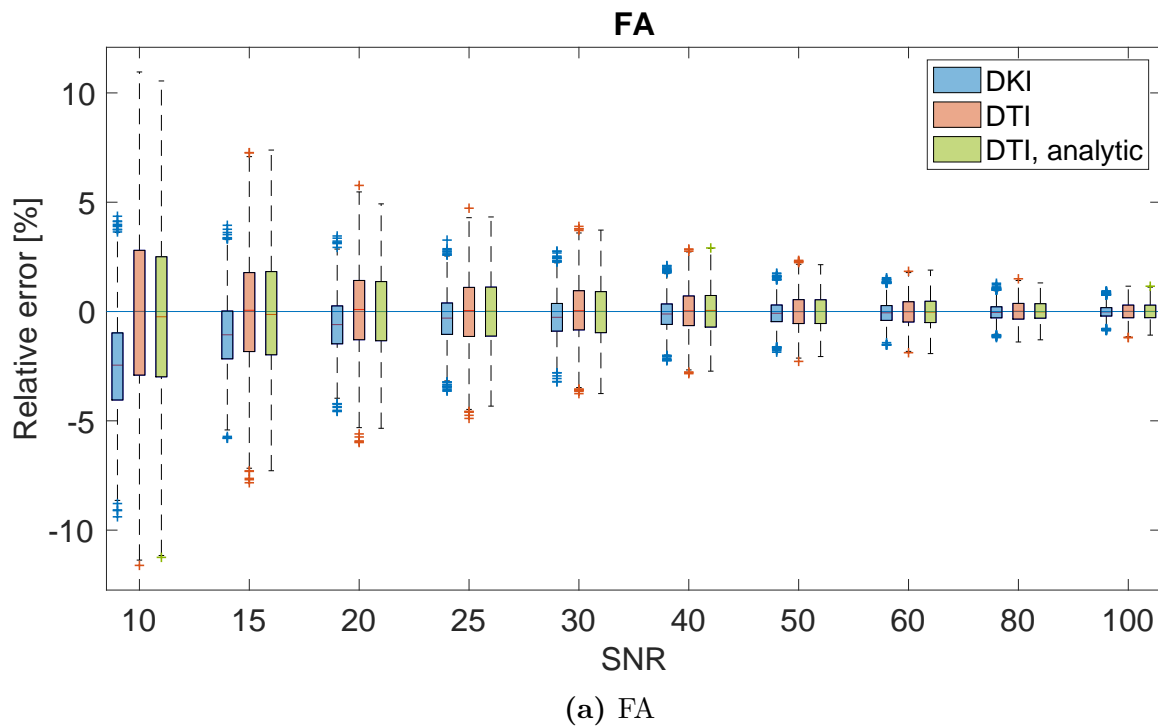
The relative error of the parameter estimations of FA, MD, RD and AD in the white matter is illustrated in figure 9.7. For convenience, the DTI and DKI estimations are presented side by side in the same box plot. The results of the analytic DTI signal are included for comparison. The differences between the analytic DTI and the WLS approximated DTI analyses are generally small, although the errors appear to be a bit smaller for the analytic DTI signal.

While the parameter estimations become progressively better for increasing SNR in both models, figure 9.7 demonstrates a trend where the error-range is larger for the DTI model than for the DKI model. This goes for both the interquartile range and

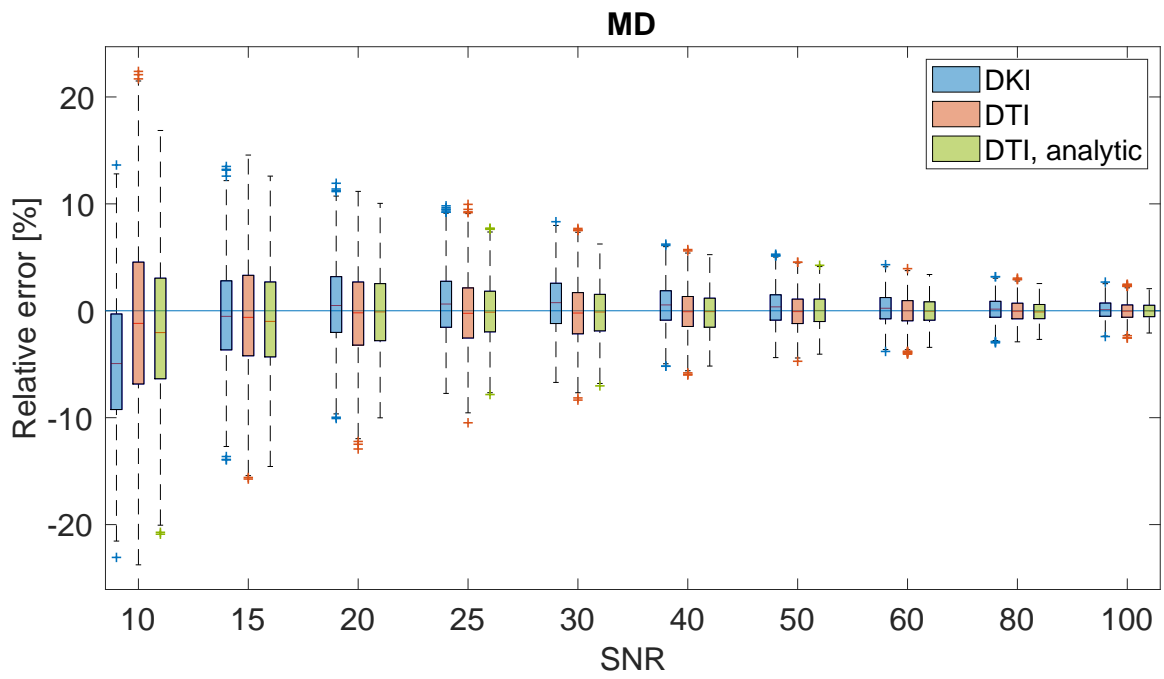
the extent of the whiskers, and becomes evident especially at lower SNR. Despite the low precision in the DTI estimations, the median seems to remain accurate and close to ground truth, even at lower SNR. The DKI estimations however have a higher precision, but lower accuracy at lower SNR¹.

Furthermore, the estimation of RD should be mentioned. While the other parameters have relatively small errors ranging from $\pm 20\%$ or less, RD has a relative error of $\pm 100\%$ at worst.

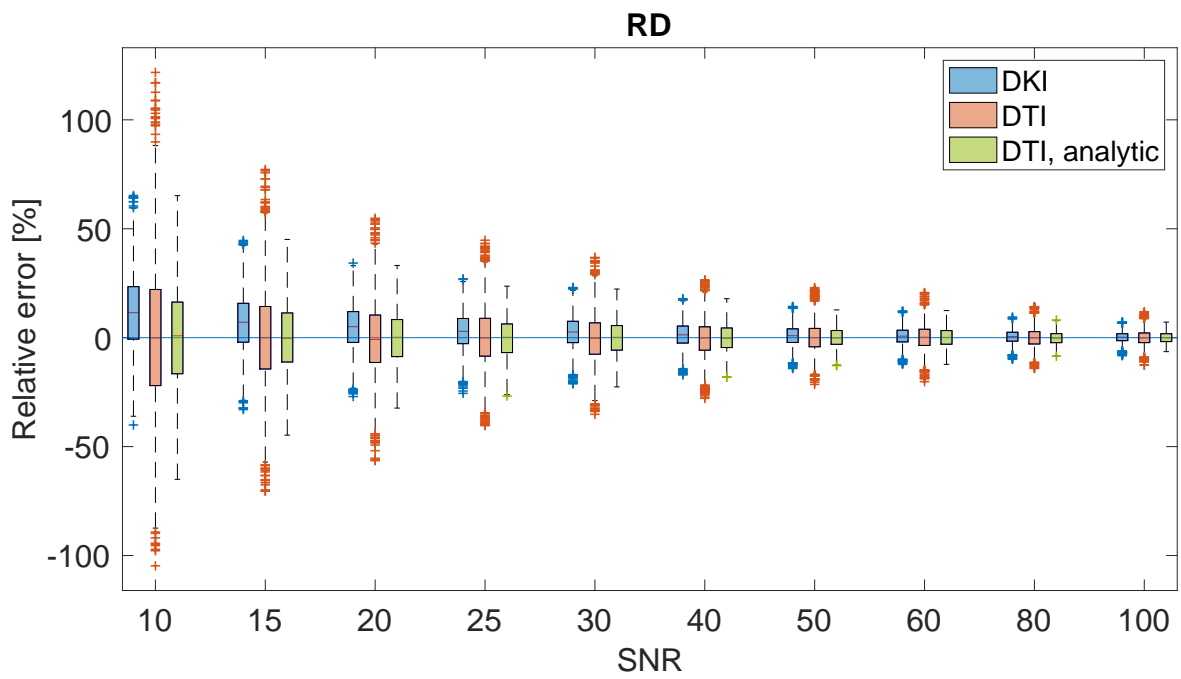
In general, it seems like the estimations are stabilizing around $\text{SNR} = 50$, as a further increase in SNR does not improve the estimations noteworthy.



¹The accuracy is measured by the median, and the precision is measured by the range of the data set.



(b) MD



(c) RD

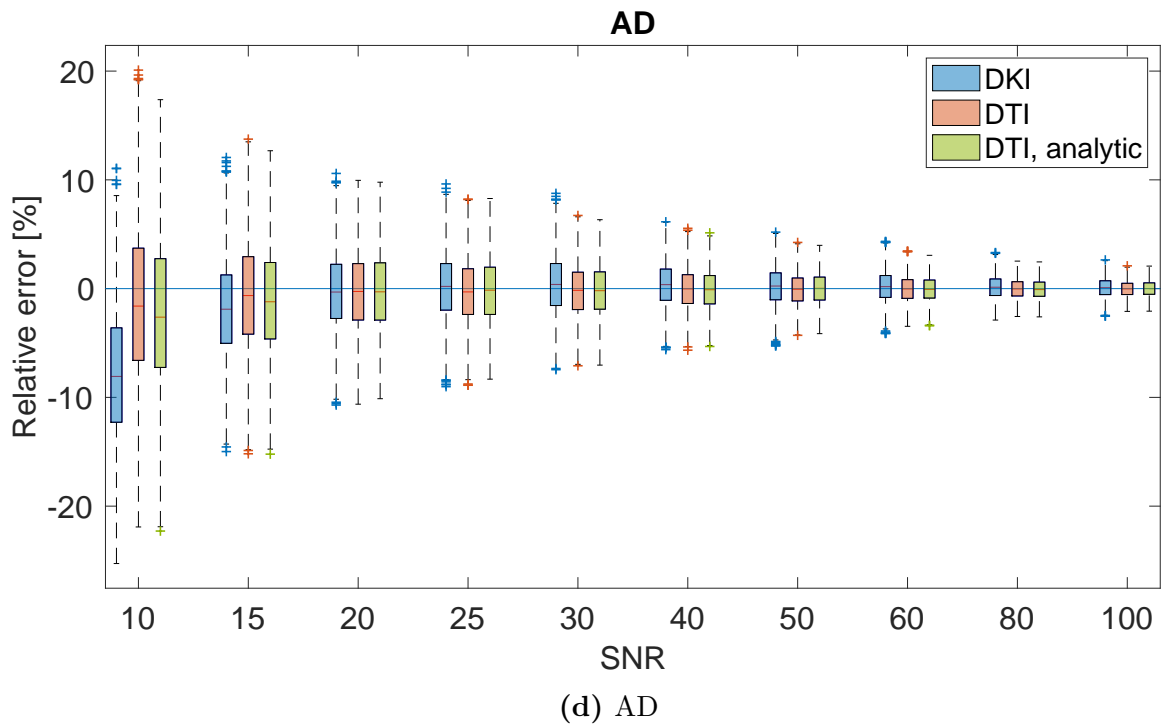
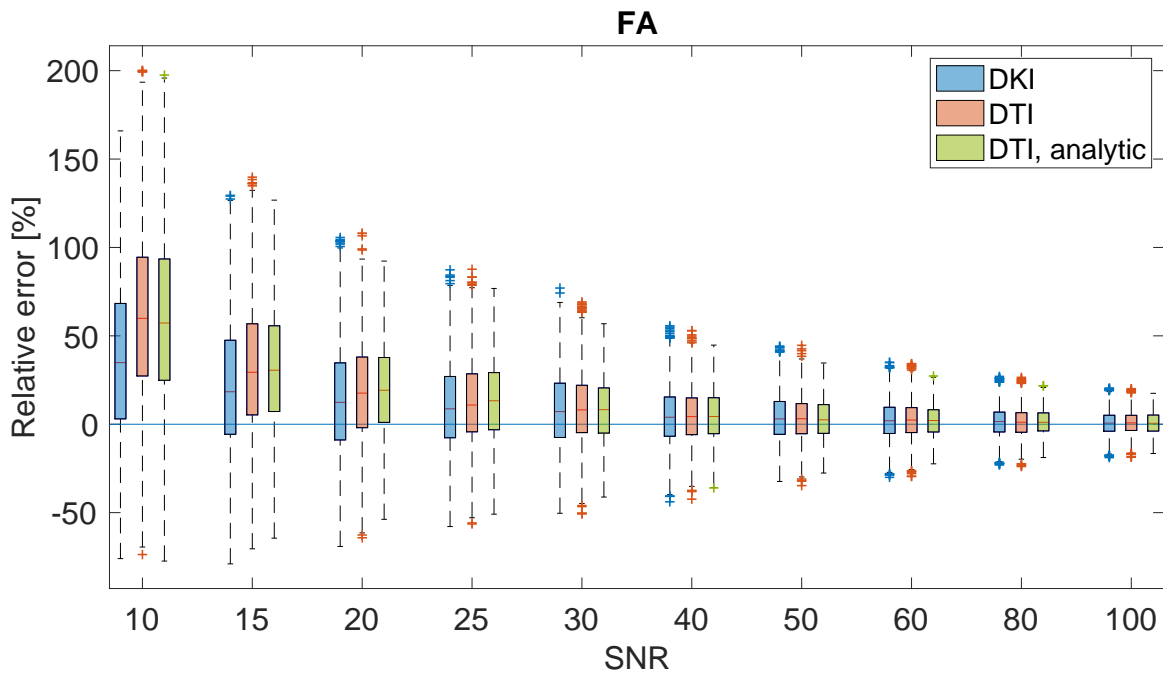


Figure 9.7: Relative error of the parameter estimations in the white matter as a function of SNR for the DTI and DKI model. The results of the analytic DTI signal are included for comparison purposes. The precision of the estimation increases with decreasing levels of noise. The figure demonstrates a general trend where the DKI estimations appear to have a higher precision, but a lower accuracy than the DTI estimations at lower SNR. Of all diffusion parameters, the estimations of RD are the least precise.

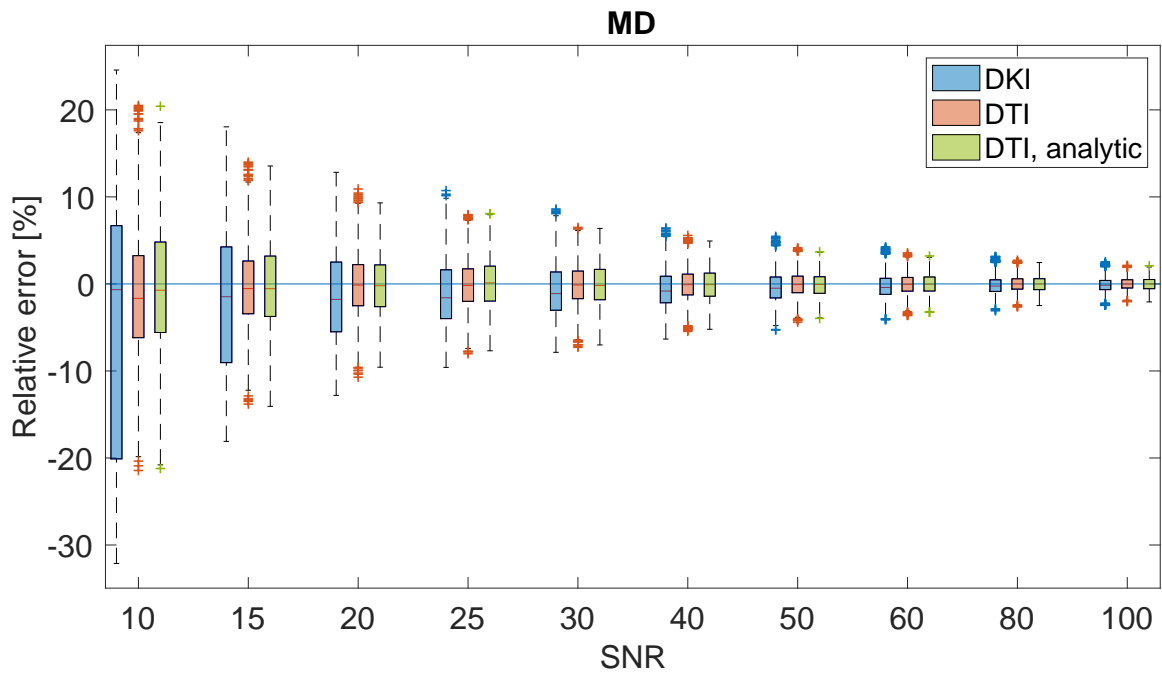
9.2.2 Gray matter

The results of the estimations of FA, MD, RD and AD in the gray matter are presented in figure 9.8. The precision of the estimations decreases with increasing noise. Interestingly, the results from the gray matter demonstrate the opposite of what was observed in white matter; the DTI model has an overall higher precision than the DKI model in estimating MD, RD and AD, this is especially evident at lower SNR. The differences between the models are however not as apparent as in the white matter.

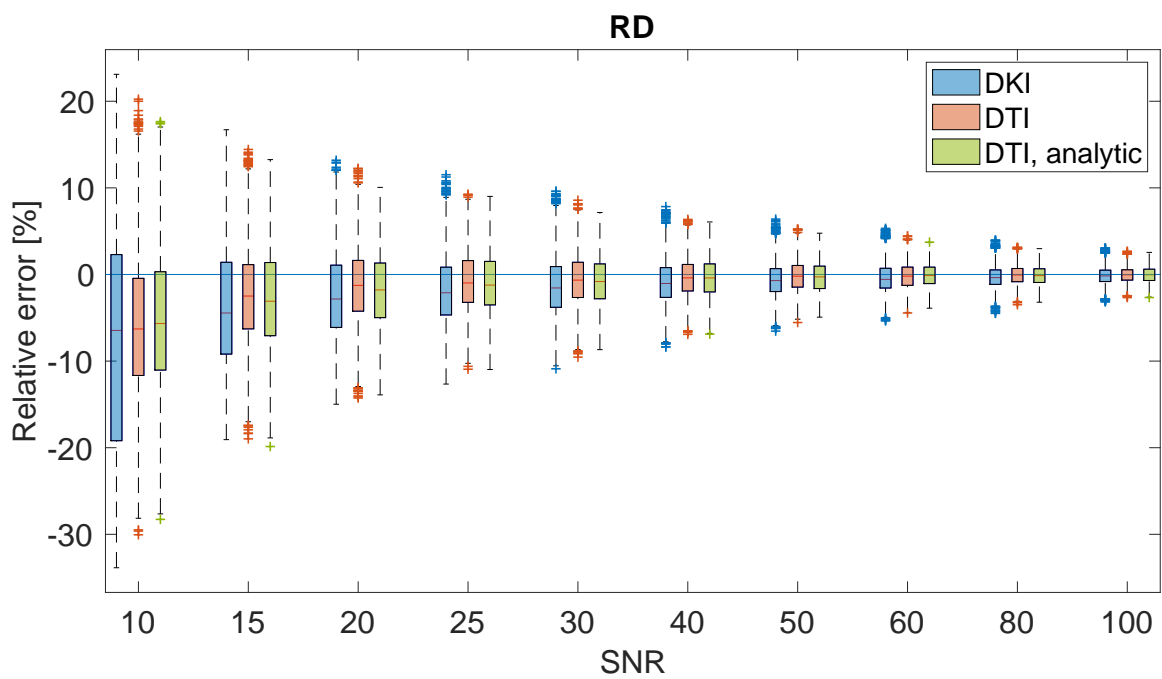
The largest errors are observed in the estimations of FA, with relative errors stretching from - 50 % to 200 % at worst. The estimations seem to stabilize around SNR = 50.



(a) FA



(b) MD



(c) RD

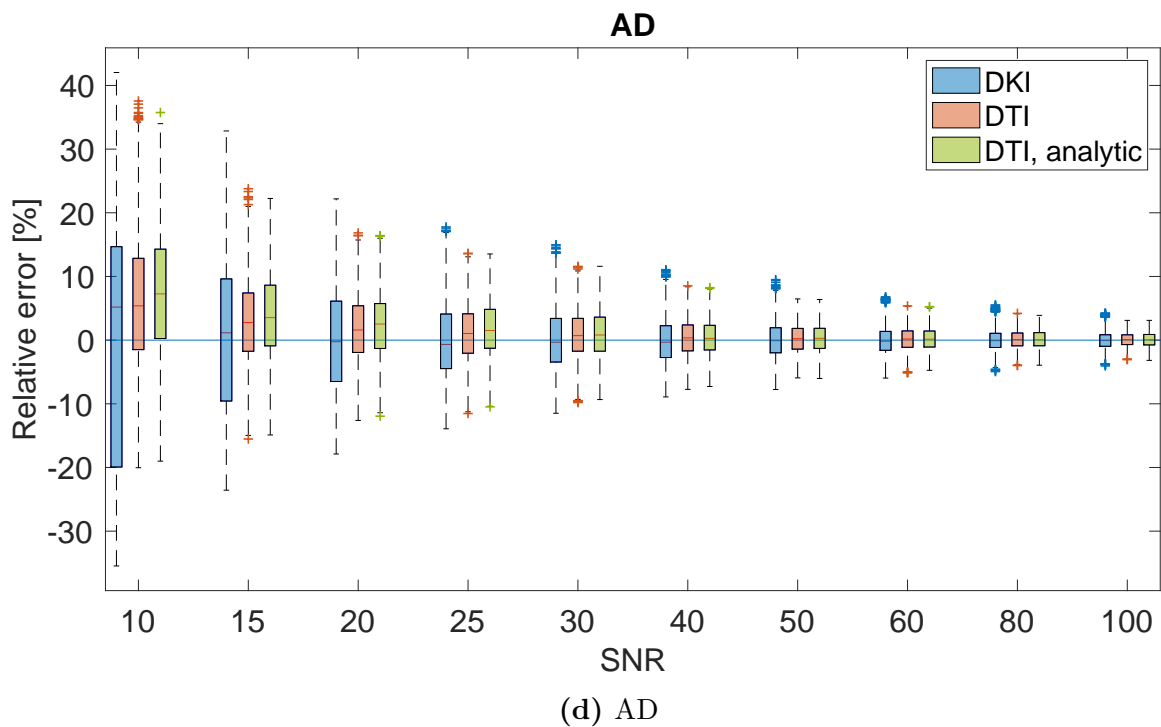


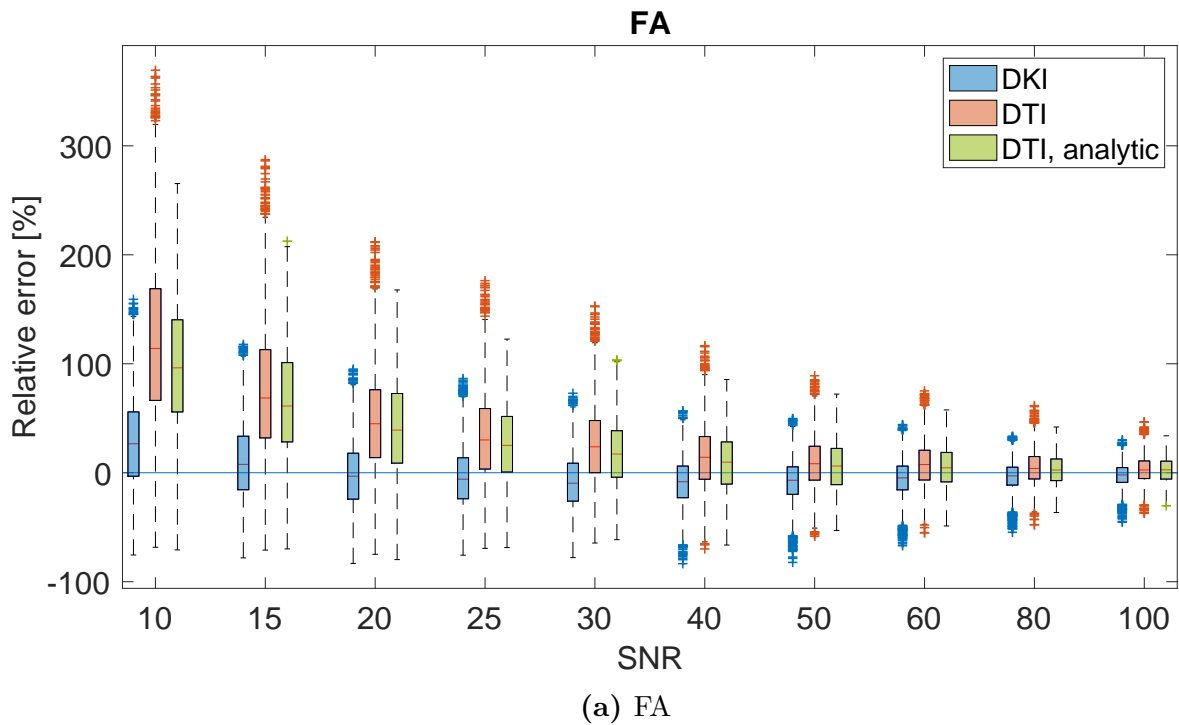
Figure 9.8: Relative error of the parameter estimations in the gray matter as a function of SNR. The precision increases with the SNR. As opposed to what was observed in white matter, the precision of the parameter estimations in DTI are higher than DKI in gray matter. The estimations of FA were the least precise of all diffusion parameters in gray matter.

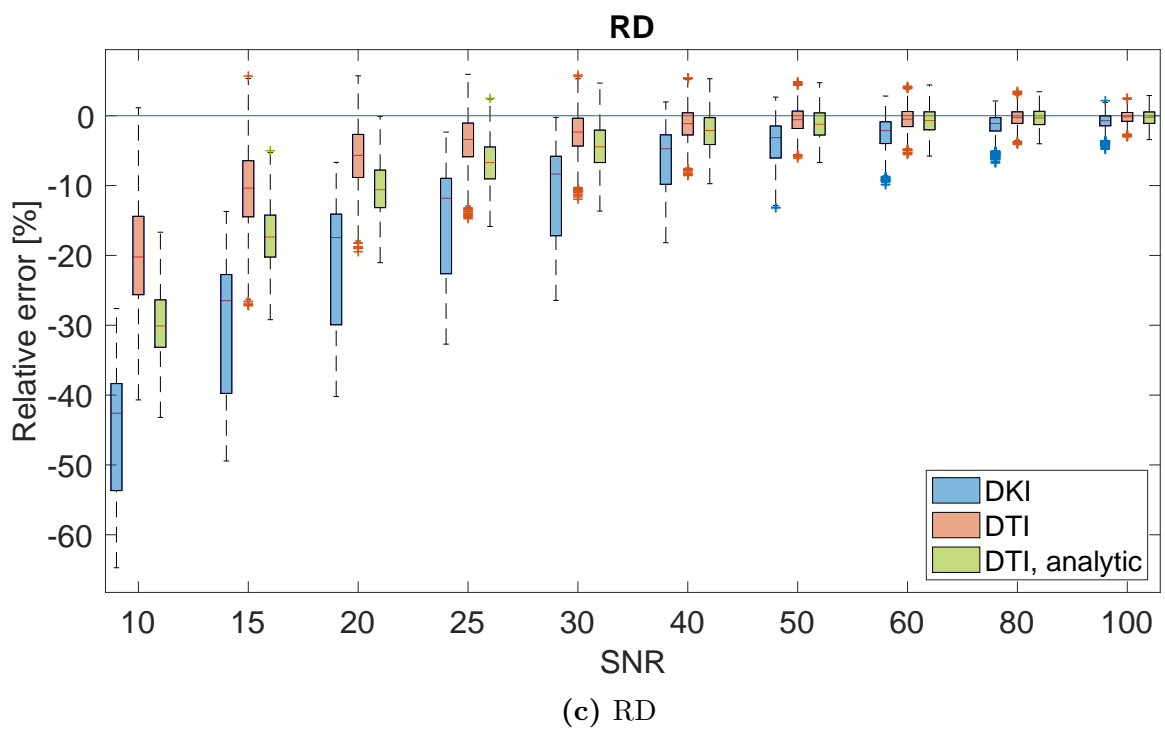
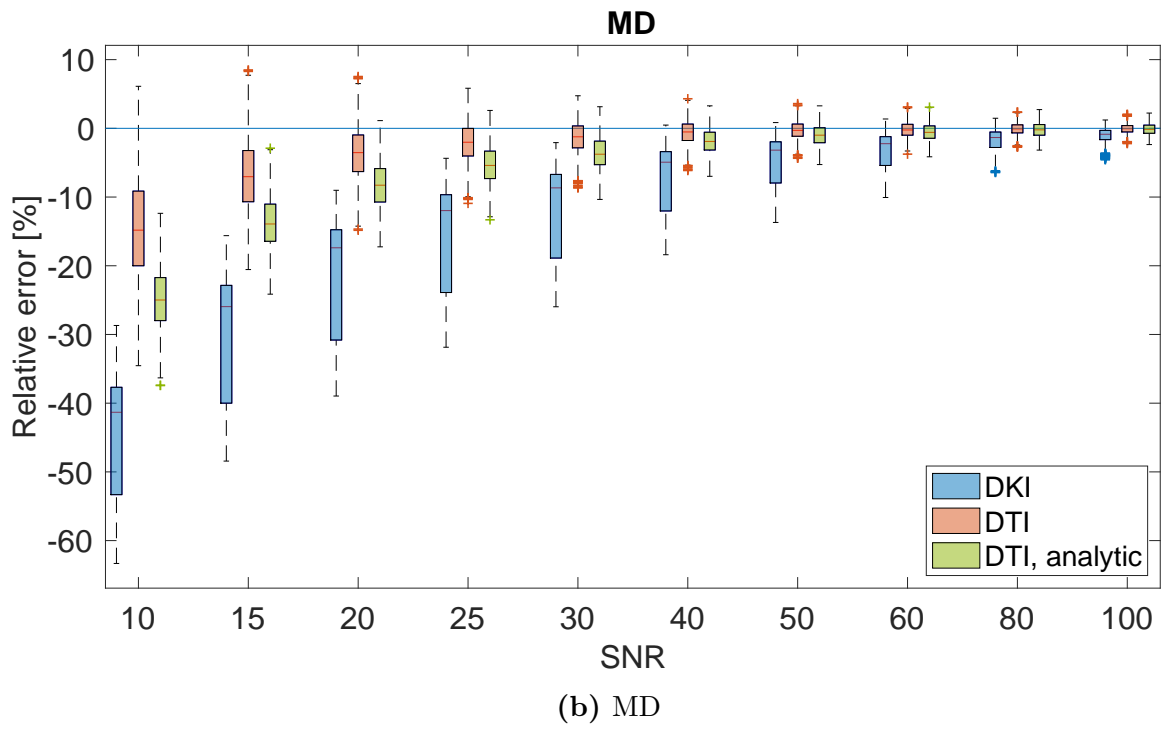
9.2.3 CSF

The results of the parameter estimations in the CSF are presented in figure 9.9. There is a clear distinction in both accuracy and precision between the models that can be observed among all diffusion parameters. For MD, RD and AD, the DTI analysis is most accurate, while the DKI analysis was the least accurate.

In FA however, the situation appears to be the opposite; the DTI analysis is less accurate than the DKI analysis. Furthermore, FA seems to have the largest range of relative errors, regardless of model. In general, the relative error in the parameter estimations is larger in the CSF than in the white and gray matter.

When the SNR reaches a level of around 50-60, the precision in the estimations seem to stabilize and a further increase in SNR does not have any remarkable effect on the estimations.





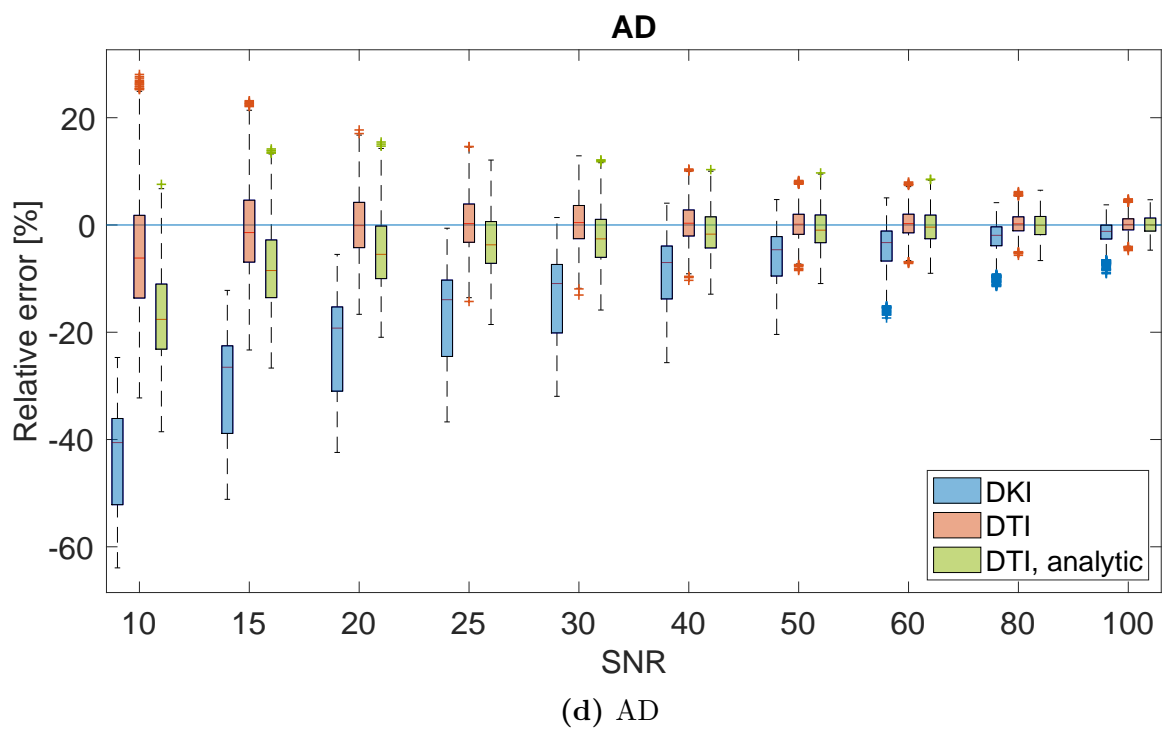
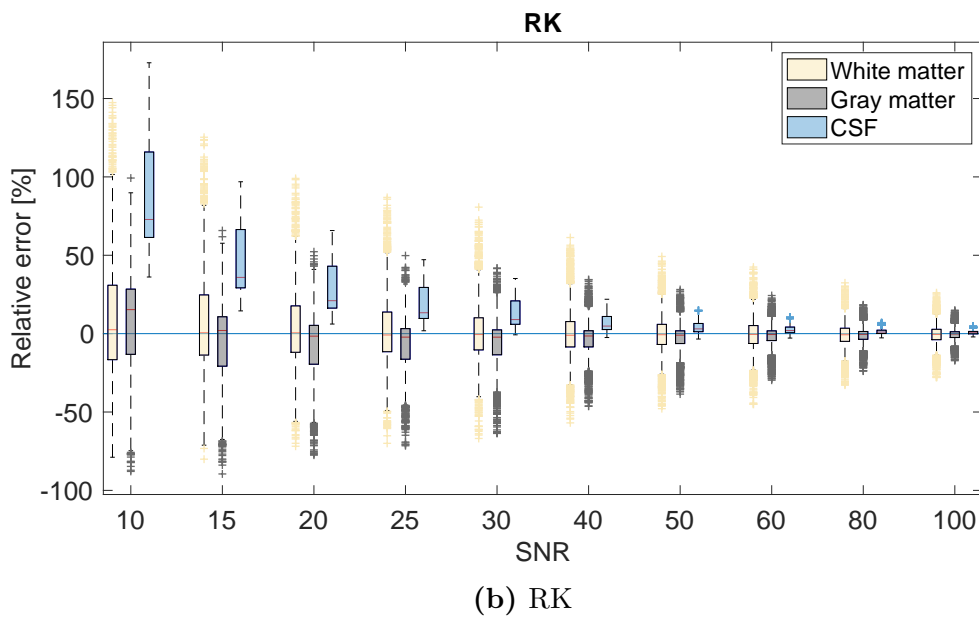
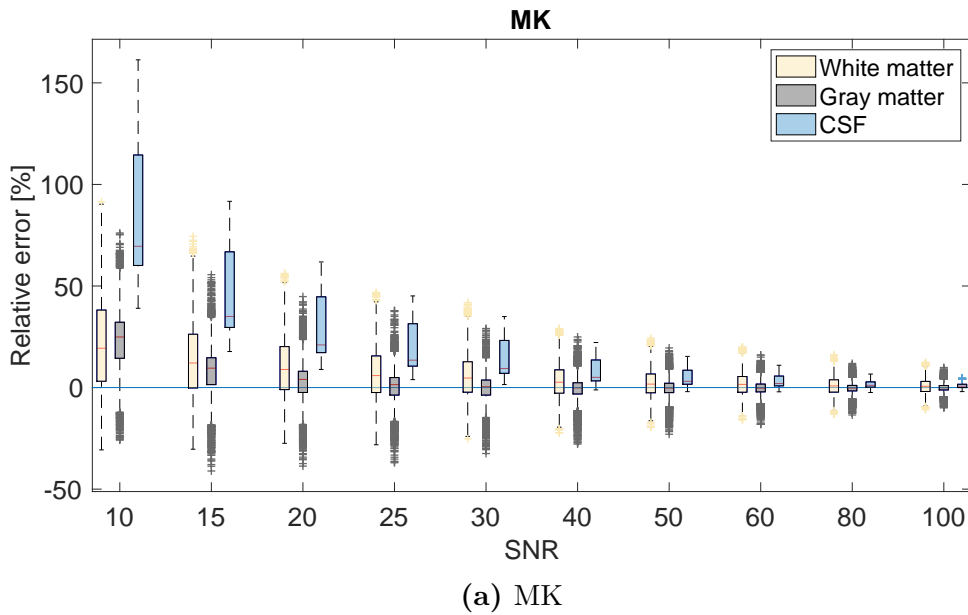


Figure 9.9: Relative error of the parameter estimations in the CSF as a function of SNR. A clear distinction between the models is observed in all parameters. In the estimations of MD, RD and AD, DTI is the most accurate, and the DKI model is the least accurate. In the estimations of FA, the DKI model is the most accurate. Additionally, FA seems to have the largest range of relative errors.

9.2.4 MK, RK and AK

Estimations of the DKI parameters MK, RK and AK are presented in figure 9.10 as a function of SNR for signals simulating white matter, gray matter and CSF. A general trend is observed where the accuracy of the parameter estimations is remarkably lower in the CSF than in the white and gray matter. The estimations seem to stabilize at an SNR of 50-60.



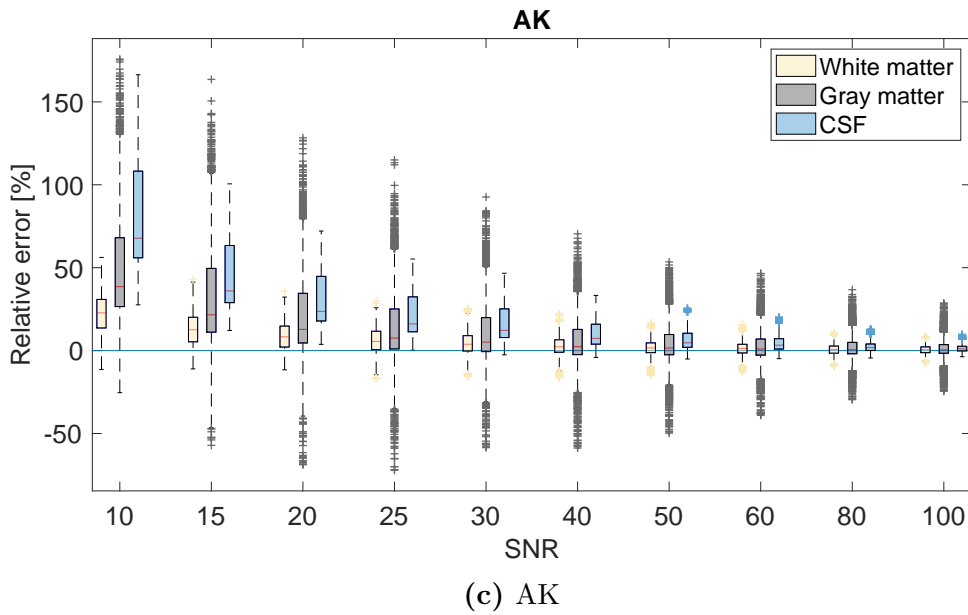


Figure 9.10: Relative error of the parameter estimations of MK, RK and AK from signals simulating white matter, gray matter and CSF as a function of SNR. The estimations in the CSF are particularly inaccurate compared to the estimations in white and gray matter.

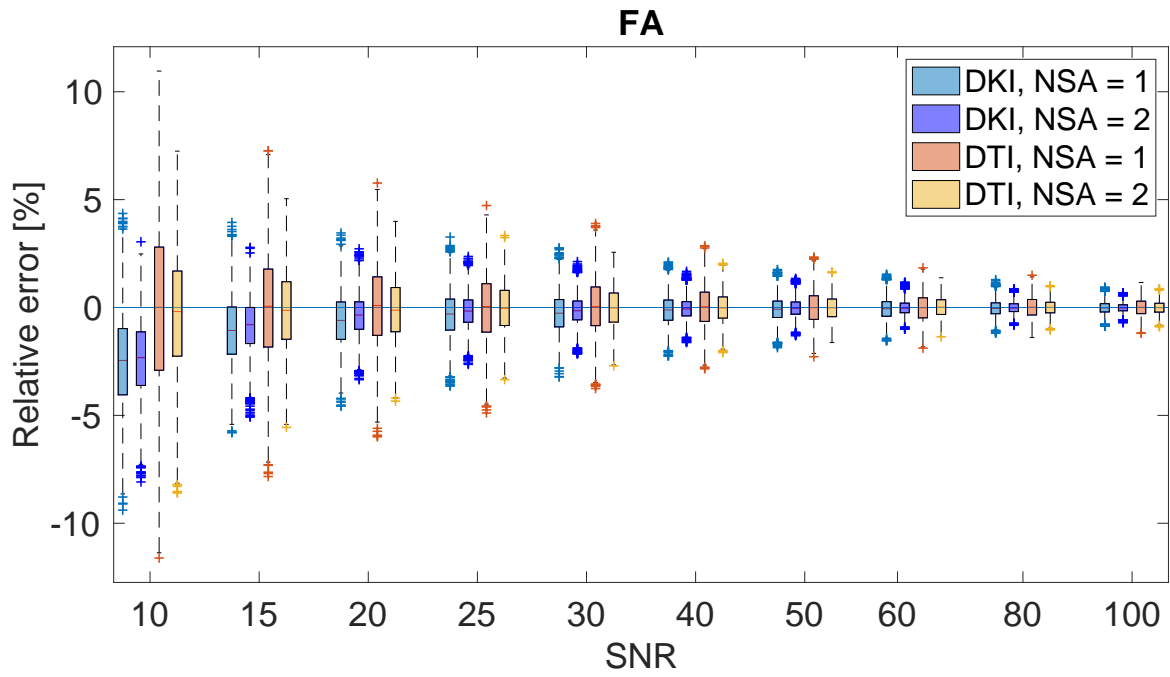
9.3 Parameter estimation as a function of NSA

The effect of doubling the NSA from 1 to 2 was investigated for signals simulating white matter, gray matter and CSF for various levels of SNR. The results are presented in boxplots illustrating the relative error as a function of SNR.

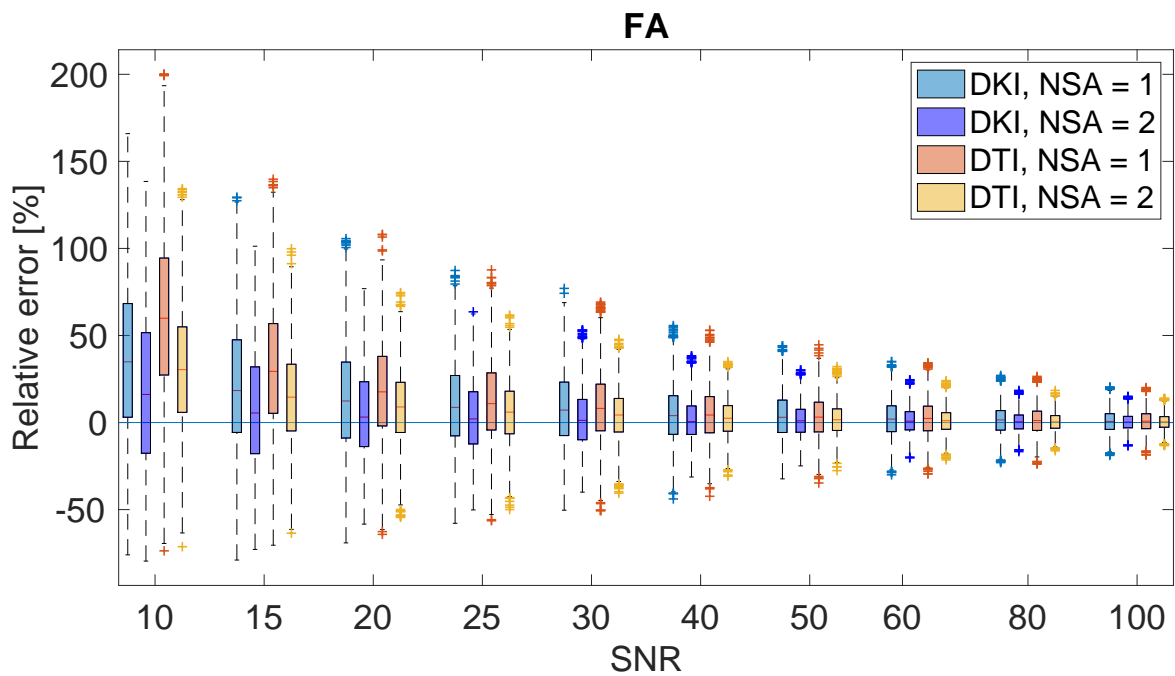
9.3.1 FA

The effect of doubling the NSA on the estimations of FA for signals simulating white matter, gray matter and CSF is illustrated in figure 9.11. Both models are included in the plot for $NSA = 1$ and $NSA = 2$. An increased NSA seems to have a noticeable effect on the parameter estimations in both models. The overall trend is an improved precision, where the improvement appears to be equivalent to a + 10 increase in SNR, as well as an increase in accuracy. The effect of an increased NSA is most prominent in the gray matter and CSF, where the errors are of a considerable size. The same result was observed in MD, RD and AD^2 .

²These results can be found in Appendix A, figure A.1, A.2 and A.3



(a) FA in white matter



(b) FA in gray matter

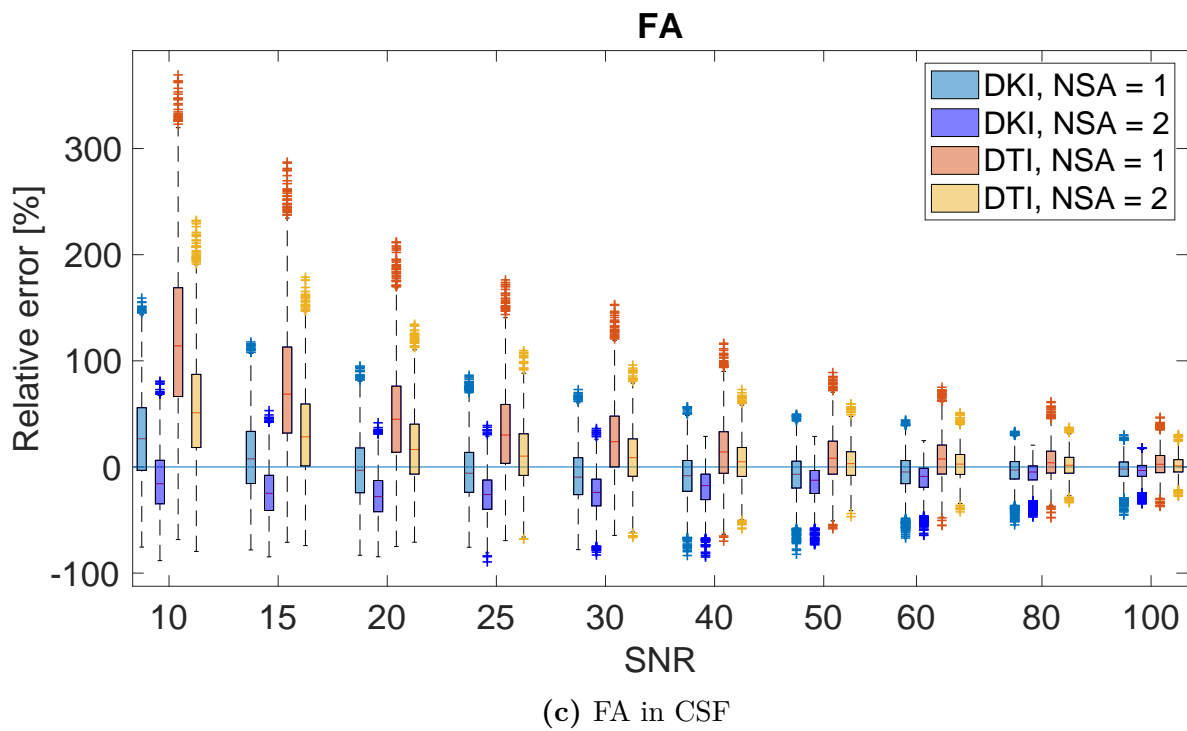


Figure 9.11: Relative error of the estimations of FA from signals simulating white matter, gray matter and CSF, for $\text{NSA} = 1$ and $\text{NSA} = 2$. Doubling the NSA resulted in an increase in both the precision and accuracy of the estimations. The same effect was observed for the diffusion parameters MD, RD and AD.

9.3.2 MK

The effect of doubling the NSA on the estimations of MK, RK and AK was investigated for increasing levels of SNR. The results for MK are shown in figure 9.12 for signals simulating white matter, gray matter and CSF. The apparent effect of increasing the NSA from 1 to 2 was an increased precision. The parameter estimations of RK and AK yielded the same results³.

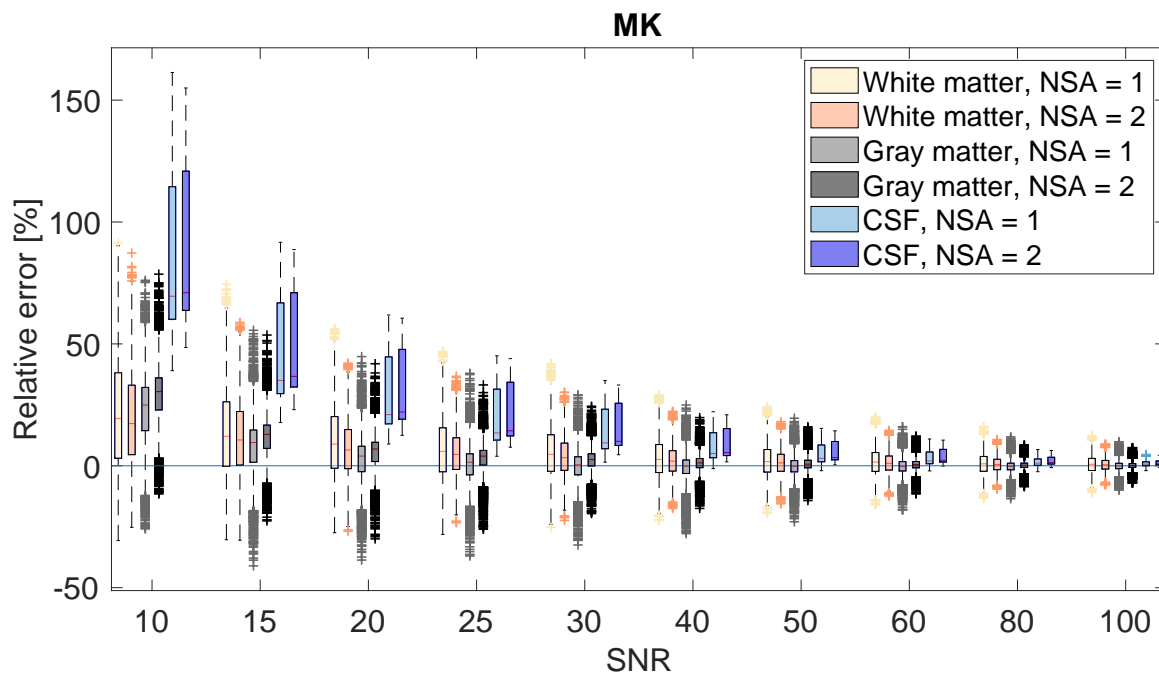


Figure 9.12: Estimations of MK as a function of SNR for signals simulating white matter, gray matter and CSF. Doubling the NSA resulted in an increase in the precision of the estimations. The same effect was observed in RK and AK.

9.4 Parameter estimation as a function of number of gradient directions and b -values

Diffusion parameters were estimated using gradient sets with various numbers of gradient directions and b -values. The complete list of gradient sets is given in table 8.1. The results of the simulations are presented in box plots showing the relative error of the parameter estimations for the various gradient sets. Next follows the

³The results of RK and AK are presented in Appendix A, figure A.4 and A.5

results of the DTI and DKI analysis, presented in separate sections, for signals representing white matter, gray matter and CSF.

9.4.1 DTI analysis

The parameters FA, MD, RD and AD were estimated using the DTI model for various numbers of gradient directions and b -values. The results of the FA and MD estimations are given in the following sections. The results of RD and AD can be found in Appendix A, figure A.6 and A.7.

FA

The estimations of FA in white matter, gray matter and CSF are presented in figure 9.13. The most prominent result is the error-range in white matter, being remarkably smaller than in the gray matter and the CSF. In the white matter, the largest errors are to be found in set 1 and set 2, containing only 1 b_0 -image and 15 and 30 gradient directions for $b = 1000$ s/mm², respectively⁴. Furthermore, there is only a minor improvement in the estimations between set 1 and 2, despite the doubling of gradient directions in set 2. Also, note the difference between set 2 and 3; when going from only 1 b_0 -image to 6 b_0 -images, the precision increases drastically. Surprisingly, the error in the estimations are larger in set 4 and 5 than in set 3, despite set 4 and 5 having an additional b -value of 500 s/mm². However, set 4 and 5 contain only 1 b_0 -image, while set 3 contains 6 b_0 -images.

A step-wise improvement in the estimations is observed when going from set 4 to set 5 and 6, as the number of gradient directions increases for each set⁵. Interestingly, the results of set 3 and set 6 seem to be equally precise, indicating that adding $b = 500$ s/mm² to the measurements has no major effect on the estimations. Also, adding $b = 3000$ s/mm² did not improve estimation accuracy or precision.

As can be seen in 9.13c, the errors of set 7 and 8 in the CSF were of such magnitude that the results of the remaining sets got overshadowed. Both sets contain b -values up to 3000 s/mm², however, while set 7 contains only 1 b_0 -image, set 8 contains 6 b_0 -images⁶. The differences in the estimations are marginal. The results are presented

⁴A small reminder of the gradient sets (b -values are given in s/mm²):

Set 1: 1 b_0 , 15 $b = 1000$

Set 2: 1 b_0 , 30 $b = 1000$

Set 3: 6 b_0 , 30 $b = 1000$

⁵

Set 4: 1 b_0 , 6 $b = 500$, 15 $b = 1000$

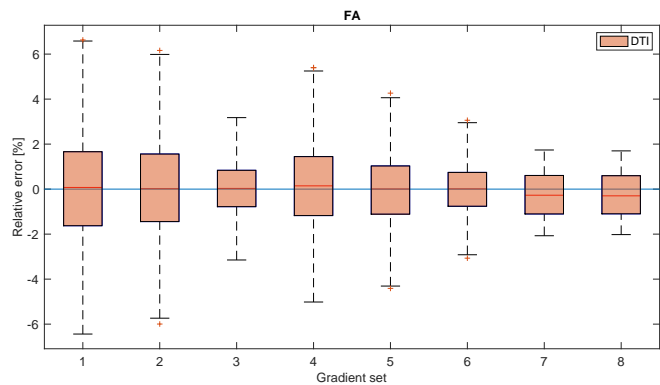
Set 5: 1 b_0 , 12 $b = 500$, 30 $b = 1000$

Set 6: 6 b_0 , 12 $b = 500$, 30 $b = 1000$

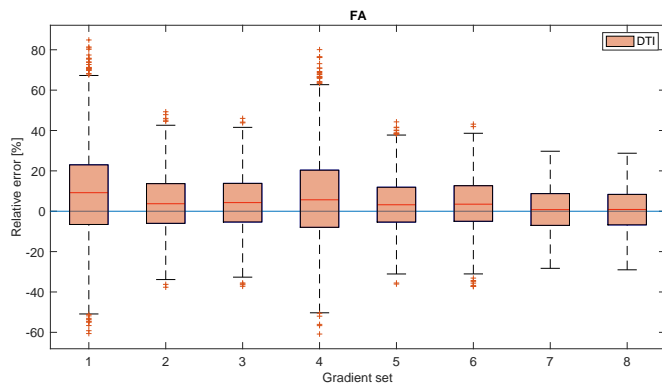
⁶

Set 7: 1 b_0 , 12 $b = 500$, 30 $b = 1000$, 50 $b = 3000$

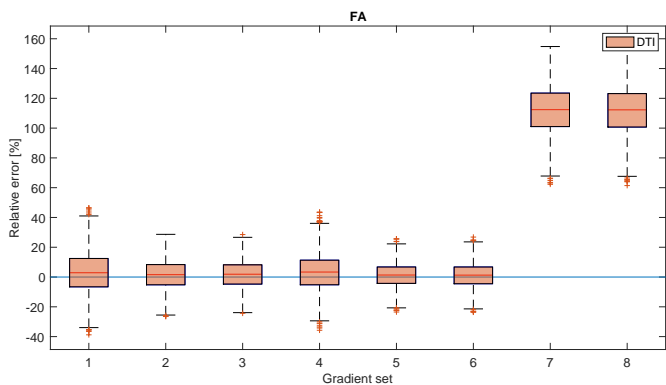
without set 7 and 8 in figure 9.13d. With set 7 and 8 removed from the plot, it appears that the estimations in set 1 and 4 are the least precise in the gray matter and CSF. Both sets contain only 1 b_0 -image and have a reduced number of gradient directions. Otherwise, there seems to be only minor differences across the gradient sets in the gray matter and CSF.



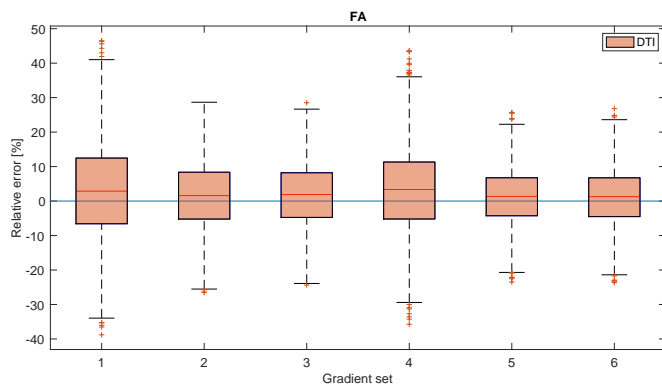
(a) FA in white matter



(b) FA in gray matter



(c) FA in the CSF



(d) FA in the CSF, without set 7 and 8

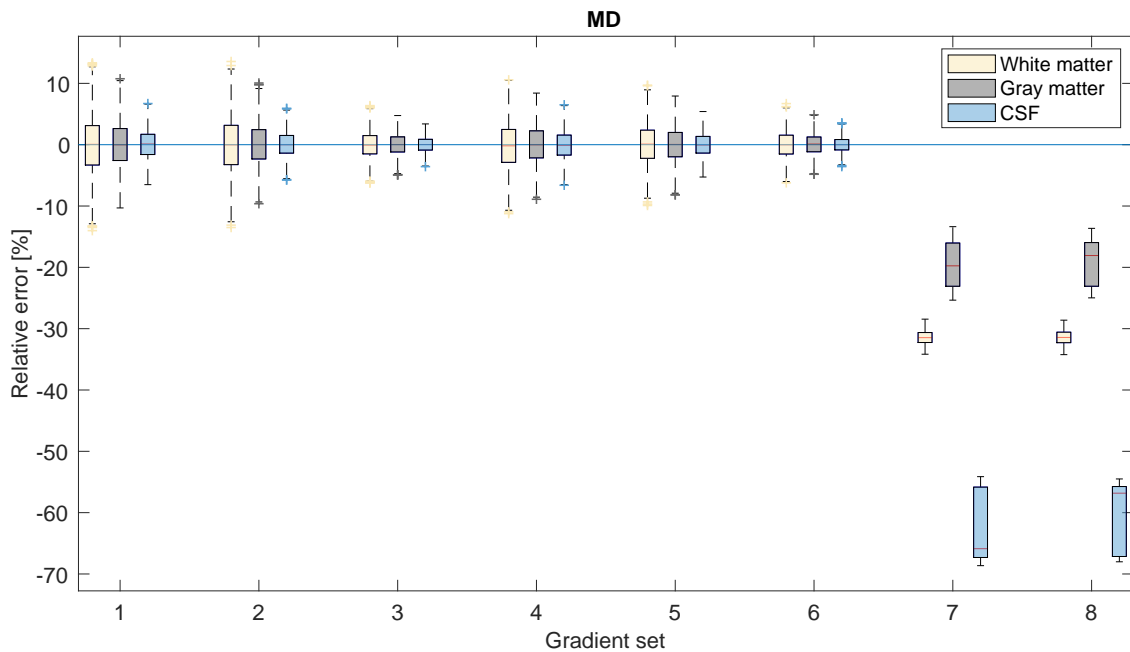
Figure 9.13: Estimations of FA in the white matter, gray matter and CSF for various gradient sets using the DTI model with $SNR = 50$. In general, the errors are remarkably smaller in the white matter. Also, increasing the number of b_0 -images from 1 to 6 results in a noteworthy improvement in the estimations, as can be seen in (a) when going from set 2 to set 3. Furthermore, adding an additional $b = 500$ s/mm² to the gradient set had no considerable effect on the estimations, as can be seen in set 3 and set 6. In the gray matter and CSF, set 1 and set 4 yield the least precise results, as they contain only 1 b_0 -image and have a reduced number of gradient directions.

Set 8: 6 b_0 , 12 $b = 500$, 30 $b = 1000$, 50 $b = 3000$

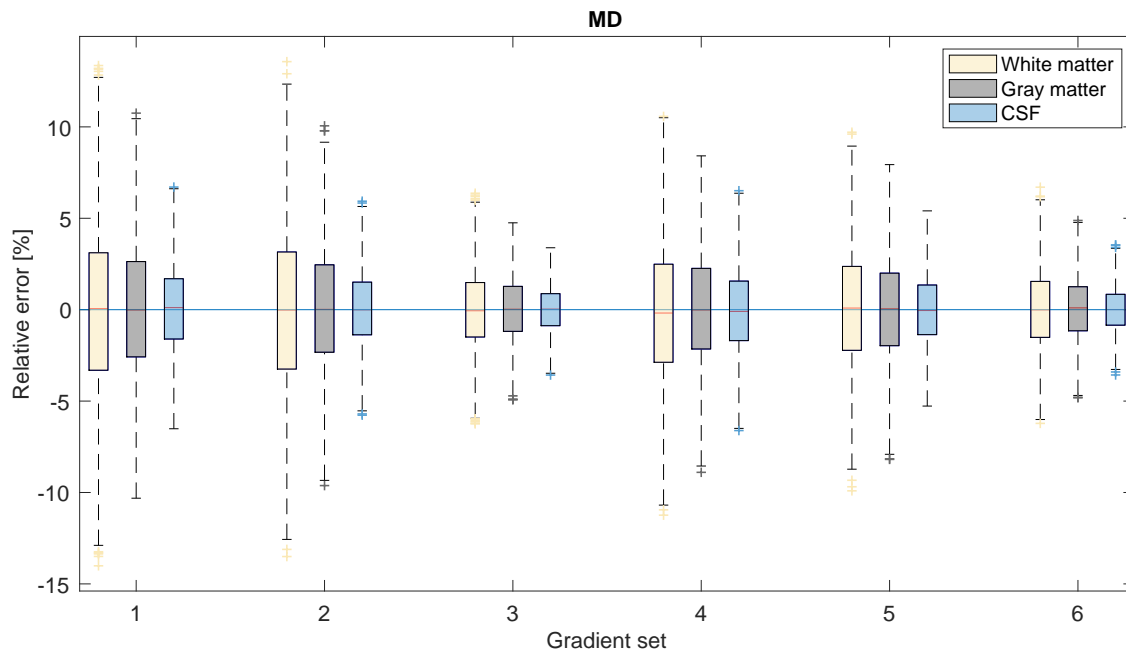
MD

The estimations of MD in the white matter, gray matter and CSF is presented in figure 9.14a. The errors in set 7 and 8 were the largest, regardless of tissue type. The results are presented without set 7 and 8 in figure 9.14b.

First and foremost, the estimations seem to remain reasonably accurate across all tissue types, with relative errors within $\pm 10\%$. In general, the results found in the estimations of FA also holds for MD. Set 1 and 2 are the least precise, as they contain the least amount of gradient directions and b -values. A noteworthy improvement is observed from set 2 to set 3, where 6 b_0 -images are used instead of just 1 b_0 -image. Set 4 and set 5 have larger errors than set 3, despite the fact that they both contain an additional b -value of 500 s/mm^2 . Set 3 and set 6 appears to be equally precise, even though set 6 has the additional $b = 500\text{ s/mm}^2$ measurements.



(a) Estimations of MD in the white matter, gray matter and CSF for various gradient sets. The errors in set 7 and 8 are considerable compared to the errors in the remaining sets.



(b) Estimations of MD, with set 7 and 8 removed from the plot.

Figure 9.14: Estimations of MD with in white matter, gray matter and CSF using the DTI model. For all tissue types, the errors remain within a 10 % range. The improvement seen when going from set 2 to set 3 indicates that adding additional b_0 -images to the gradient set has a positive effect on the estimations of MD. Also, the similarity between set 3 and set 6 implies that adding an additional b -value of 500 s/mm² only has a minor effect on the estimations.

9.4.2 DKI analysis

The DKI parameter estimations of FA, MD and MK for signals simulating white matter, gray matter and CSF for various gradient sets are presented in the next pages. The results of RD, AD, RK and AK can be found in Appendix A, figure A.8, A.9, A.10 and A.11.

FA

The effect of number of gradient directions and b -values on the estimation of FA in the white matter, gray matter and CSF is illustrated in figure 9.15. As shown by the figure, the accuracy of the estimations are largely dependent on the anatomy of the tissue. The estimations are particularly accurate in the white matter, with relatively small errors across the gradient sets. In the gray matter and CSF however, the errors range over several hundred percents.

In white matter, the least accurate results are found in set 1, which contains b -values up to only 1000 s/mm^2 . Set 2 and 3 are almost equally precise, despite the fact that set 3 contains twice as many gradient directions than set 2⁷. Furthermore, there seems to be no clear distinction between set 4 and set 5. While set 4 has 6 b_0 -images, set 5 has only 1 b_0 -image, but with an additional b -value of 1000 s/mm^2 and half the number of gradient directions⁸. A minor improvement is observed from set 6 to set 7, when going from 1 b_0 image to 6 b_0 images⁹.

Interestingly, the errors are largest in set 2 in the gray matter and CSF. In general, set 7 and 8 yield the most accurate results for all tissue types, where set 8 is the full gradient set with all b -values and gradient directions, containing the additional $b = 2000 \text{ s/mm}^2$ ¹⁰.

7

Set 1: 6 b_0 , 12 b 500, 30 b 1000

Set 2: 1 b_0 , 6 b 500, 25 b 3000

Set 3: 1 b_0 , 12 b 500, 50 b 3000

8

Set 4: 6 b_0 , 12 b 500, 50 b 3000

Set 5: 1 b_0 , 6 b 500, 15 b 1000, 25 b = 3000

9

Set 6: 1 b_0 , 12 b 500, 30 b 1000, 50 b = 3000

Set 7: 6 b_0 , 12 b 500, 30 b 1000, 50 b = 3000

10

Set 8: 6 b_0 , 12 b = 500, 30 b = 1000, 40 b = 2000 and 50 b = 3000

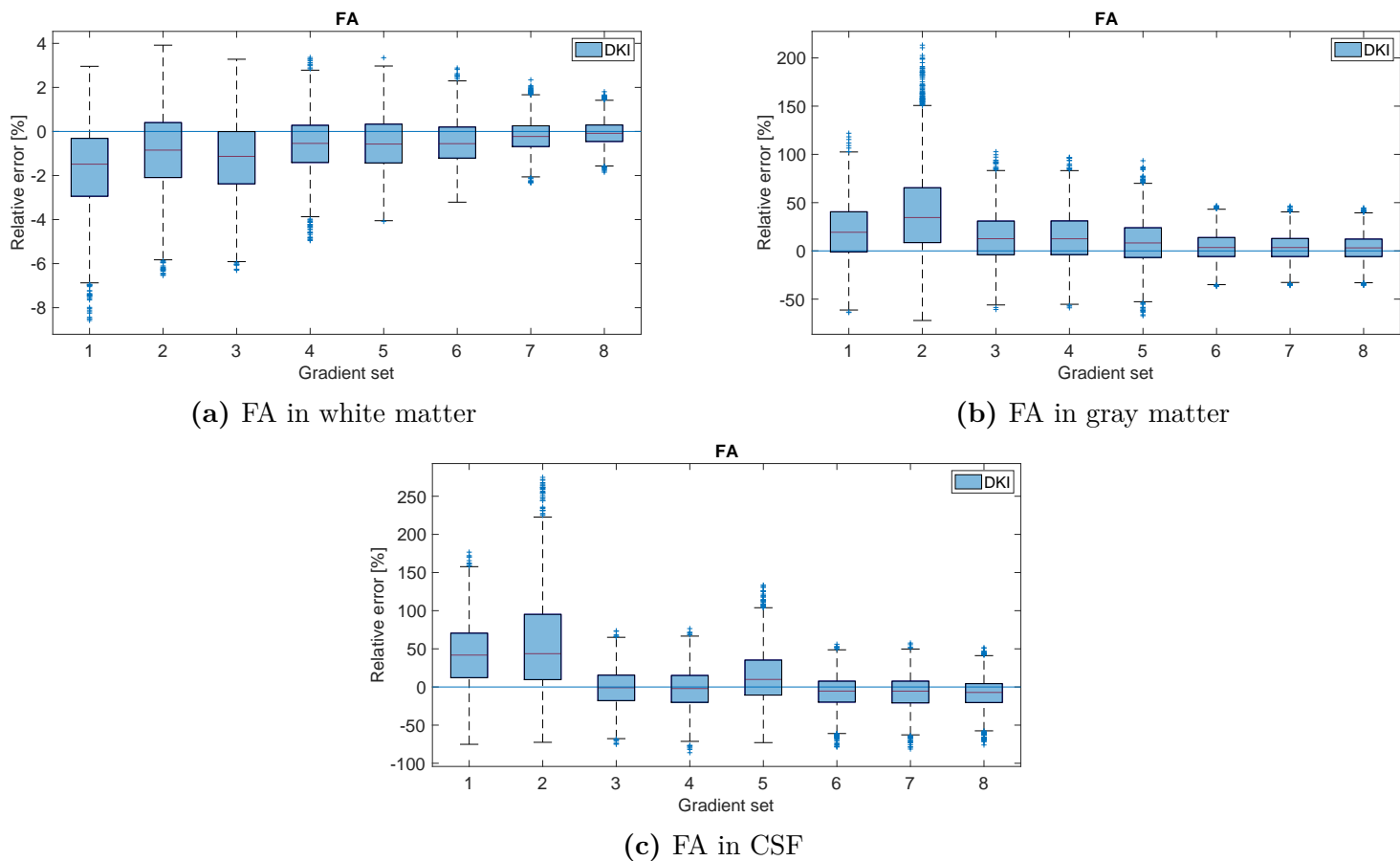


Figure 9.15: Estimations of FA in white matter, gray matter and CSF for various gradient sets using the DKI model. The errors are substantially larger in the gray matter and CSF than in the white matter, ranging over several hundred percents. In white matter, set 1, with b -values up to only 1000 s/mm^2 , gave the least accurate results, while set 7 and 8, with the highest number of gradient directions and b -values, gave the most accurate results.

MD

The estimations of MD in white matter, gray matter and CSF are presented in figure 9.15 for the various gradient sets. The accuracy of the estimations are particularly inaccurate in the CSF, even in set 8, where all gradient directions and b -values are included.

There also seems to be no apparent change in the estimations between set 2 and set 3, despite the fact that set 3 contains twice as many gradient directions than set 2. They both contain only 1 b_0 -image, however.

Furthermore, there is no clear distinction between the results of set 4, 5 and 6, which indicates that having 6 b_0 , 12 $b = 1000 \text{ s/mm}^2$ and 50 $b = 3000 \text{ s/mm}^2$ is equivalent

to having 1 b_0 with an additional 30 directions for $b = 1000$ s/mm² in the gradient set.

As expected, the most accurate estimations of MD were found in set 7 and set 8.

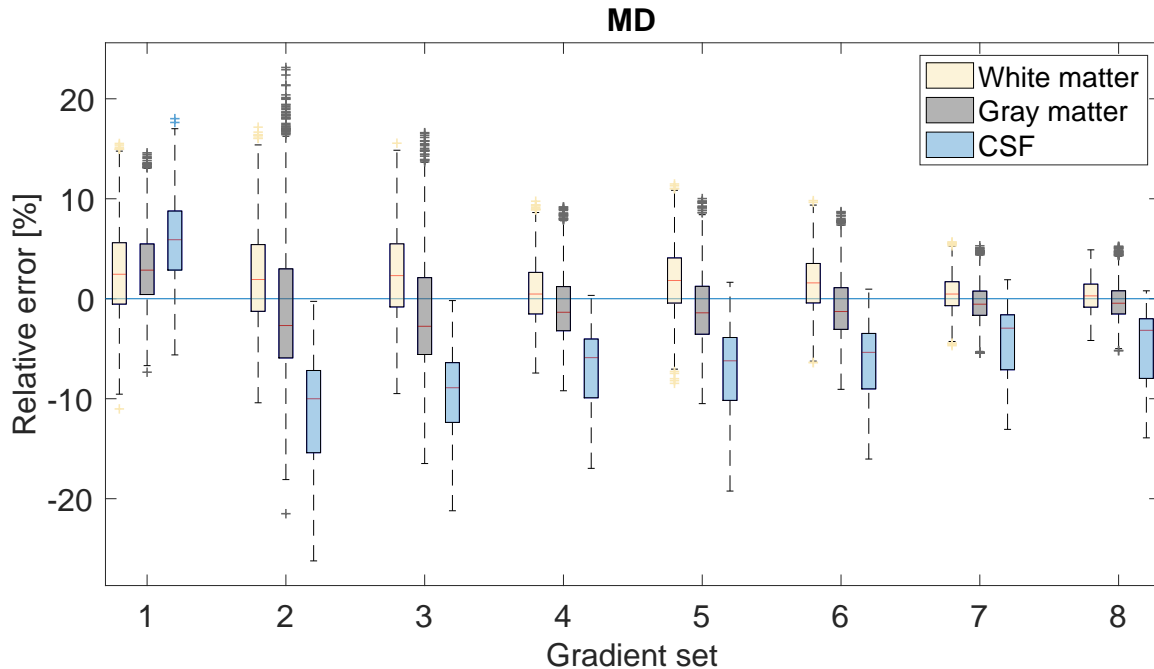


Figure 9.16: Relative error in the estimation of MD in white matter, gray matter and CSF for various gradient sets using the DKI model. The figure clearly illustrates that the estimations of MD are particularly inaccurate in the CSF. The largest errors are observed in set 2 and set 3, where both sets contain only 1 b_0 measurement. The estimations in set 4, 5 and 6 are fairly equal, while set 7 and 8 are the most precise.

MK

The estimations of MK in the white matter is given in figure 9.17. As the figure illustrates, the errors in set 1, containing b -values up to only 1000 s/mm², were of such magnitude that the results of the remaining sets got overshadowed. The same result was observed in the gray matter and CSF. The estimations of MK is therefore illustrated in figure 9.18 without set 1. There were only minor variations in the estimations across the gradient sets and tissue types, with the exception of set 8, which interestingly showed a decreased accuracy despite being the full gradient set with all diffusion MK directions and b -values. Set 7 seems to provide the most accurate estimations of MK.

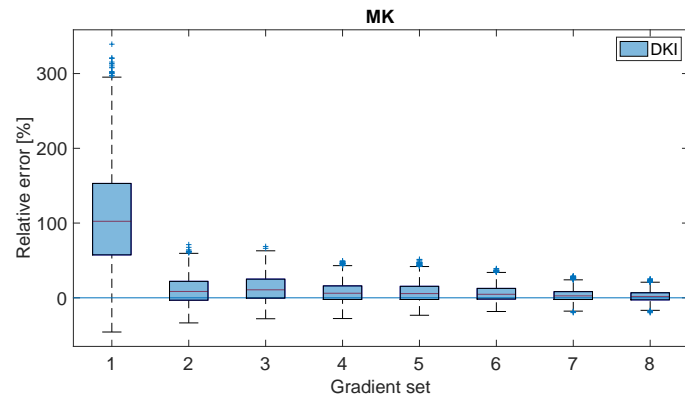


Figure 9.17: Relative error in the estimation of MK in the white matter for various gradient sets. The errors in set 1, with b -values up to only 1000 s/mm^2 , were of such magnitude that the results of the remaining sets got overshadowed. The same result was observed in the gray matter and CSF.

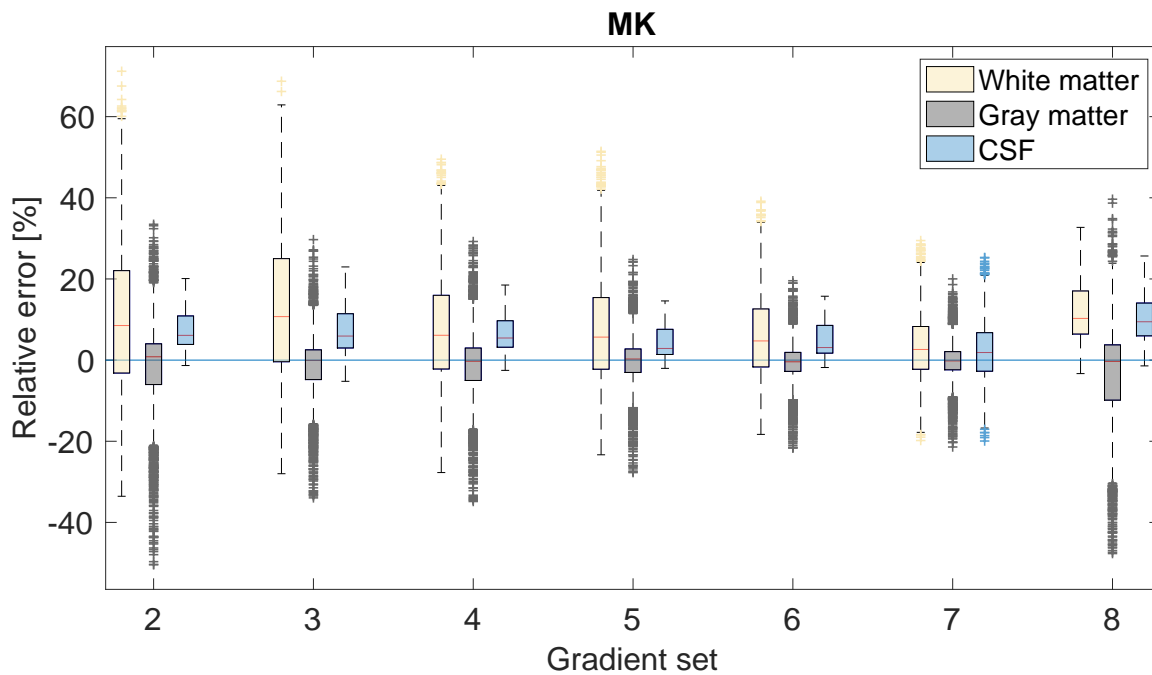


Figure 9.18: Relative error in the estimation of MK in the white matter, gray matter and CSF for various gradient sets. Set 1 was removed from the plot due to the large errors in the estimations. Only minor variations are observed across the gradient sets. Interestingly, the errors were smaller in set 7 than set 8, even though set 8 contains an additional b -value of 2000 s/mm^2 .

9.5 Summary

A summary of the most important observations is given in the list below:

1. Parameter estimation as a function of SNR
 - In general, the parameter estimations seem to stabilize at an SNR level of about 50 or 60, as a further increase in SNR did not result in any further improvement in the estimations.
 - The estimations of FA in white matter were least affected by limited SNR, with errors ranging from $\pm 10\%$.
 - The estimations on FA in gray matter and CSF seem to be most affected by limited SNR, with errors ranging over several hundred percents.
2. Parameter estimation as a function of NSA
 - In the estimations of FA, doubling of the NSA resulted in an increase in both precision and accuracy. In the estimations of MK, doubling the NSA resulted in an increase in precision. The result holds for both models.
3. Parameter estimation as a function of number of gradient directions and b -values
 - In general, gradient sets containing 6 b_0 -images yielded more precise results than gradient sets containing only 1 b_0 -image.
 - In the DTI analysis, using a gradient set with only 1 b_0 -image and either 15 or 30 gradient directions for $b = 1000$ s/mm² gave the least accurate results in white matter.
 - In the DTI analysis, using a gradient set with 6 b_0 -images and 30 $b = 1000$ s/mm² directions was equally precise as using a gradient set with 6 b_0 -images, 12 $b = 500$ and 30 $b = 1000$ s/mm² directions.
 - In the DTI analysis, gradient sets containing b -values up to 3000 s/mm² did not improve the estimations noteworthy. In some cases, the errors when using these sets were substantial.
 - In the DKI analysis, using a gradient set with b -values up to only 1000 s/mm² gave the least accurate results in white matter and should not be preferred.
 - In the DKI analysis, going from 1 b_0 -image to 6- b_0 images did only have a minor impact on the estimations when the gradient set already contained the maximum number of gradient directions for the b -values 500, 1000 and 3000 s/mm².

- In the DKI analysis, using a gradient set with 6 b_0 -images, 12 $b = 500$, 30 $b = 1000$ and 50 $b = 3000$ s/mm² was equally precise as using the full gradient set, containing the additional 40 $b = 2000$ s/mm² measurements.

Part V

Discussion and conclusion

Chapter 10

Discussion and conclusion

The main objective of this thesis was to compare the two DWI models DTI and DKI by studying the effect of SNR, NSA, number of gradient directions and b -values on the diffusion parameter estimation. The motivation for conducting this study was based on the axiom that the simplest model able to properly describe a given phenomenon should be preferred. With more complex diffusion models, more model parameters are needed, and larger data sets must be acquired per scan, resulting in a longer acquisition time. On the other hand, the increase in model complexity may be justified by an increased sensitivity and the detection of subtle changes in the tissue microstructure.

The results from each simulation experiment are further discussed in the following sections.

10.1 Parameter estimation as a function of SNR

There are some differences in the estimation of the DTI-parameters FA, MD, RD and AD between the DTI model and the DKI model in the white matter. As figure 9.7 suggests, the errors are generally larger in DTI than in DKI, especially in settings of limited SNR. Despite the fact that the DKI model is more precise in its estimations, it demonstrates a lower accuracy at the lowest levels of SNR, being particularly inaccurate at $\text{SNR} = 10$. Nevertheless, the accuracy of the DKI model improves gradually as the SNR level increases.

It could be speculated that the reduced accuracy of the DKI model in settings of limited SNR is a case of overfitting, which may occur when a statistical model contains more model parameters than can be justified by the data. When the SNR is sufficiently low, the model may extract some of the residual variation in the signal, i.e. noise, as if that variation represented the underlying structure of the model[40].

The errors in the estimations of RD are particularly large, as seen in figure 9.7c, the errors range from $\pm 100\%$ at the lowest levels of SNR. A possible explanation for this may be the anisotropic structure of the white matter. When diffusion is highly anisotropic, the size of RD is especially small. As can be seen in table 7.1, the ground truth values of RD in the selected white matter voxels are all in the range $0.12\text{-}0.21 \times 10^{-3} \text{ mm}^2/\text{s}$. As stated in formula (4.42), RD is calculated as the mean of the two eigenvalues λ_2 and λ_3 . The estimation of RD therefore depends on diffusion measurements in two orthogonal axes. When λ_2 and λ_3 are small, a small error in these eigenvalues may lead to substantial errors in the estimation of RD.

In gray matter, the values of RD are generally larger. Figure 9.8c illustrates that the errors in RD are much smaller in gray matter, in contrast to what was observed in the white matter. This may substantiate the fact that large errors may arise in the estimations when diffusion is more restricted in the radial plane and RD is small.

In contrast, there is a substantial increase in the errors in the estimations of FA in gray matter (figure 9.8a) compared to white matter. As stated in equation (4.43), FA is calculated through the three eigenvalues of the diffusion tensor. Gray matter is almost isotropic to water diffusion, as it has a randomly ordered microstructure, at least on the scale of a voxel. Thus, the differences between RD and AD, and therefore also the eigenvalues λ_1 , λ_2 and λ_3 , are small in gray matter. This can be seen in the ground truth values of RD and AD in table 7.1. Small deviations in one (or more) of the eigenvalues may therefore have an impact on the calculation of FA, as the diffusion ellipsoid goes from being almost spherical to being more pointed. As seen in figure 9.8c and 9.8d, the underestimation of RD is somewhat larger than the overestimation of AD, which results in an overestimation of the FA value. Image noise produces a random perturbation of the diffusion tensor, which leads to an overestimation of AD and an underestimation of RD, hence, diffusion anisotropy is overestimated in general[41]. In the white matter, where the differences between RD and AD are larger, this is not an issue however.

In the CSF, diffusion is free, and FA is close to zero. As can be seen in figure 9.9a, the range of errors in the estimation of FA are substantial, ranging over several hundred percents. As discussed, the equation for FA becomes more sensitive to errors when the three eigenvalues are approximately equal in size, which is also the case in the CSF. Furthermore, when the relative error given by equation (7.9) is calculated, the denominator in the expression is a small number, as FA is close to zero in the CSF. Dividing by a small number may result in a very large quotient, which may be a possible explanation for the large relative errors observed in the CSF. Thus, the substantially large errors do not necessarily imply that the models are performing badly in the CSF, as dividing by a small number can have a dramatic effect on the relative errors. The inaccuracy of the DTI model in the CSF can also impact the accuracy of gray matter voxels containing a CSF component due to partial volume effects. Free diffusion also means that the diffusion signal decays rapidly, reaching the level of background noise at much lower b -values than in white matter, making

the estimations more sensitive to noise in the CSF. This makes it challenging to make an exact estimation of FA; even tiny contributions of noise to the signal in any direction may result in an FA larger than zero, as a small degree of anisotropy is measured. This results in an overestimation of FA. It should, however, be stressed that accurate measurements of FA in CSF is rarely relevant in clinical practice, and is mainly a concern with respect to partial volume effects; i.e. where part of a white matter or gray matter voxel contains CSF - which may contaminate the underlying DWI-signal.

The analytic DTI signal was included in the simulations to illustrate the difference between a completely synthetic, noise-free DTI signal, and the baseline DTI signals generated from DW-MRI data. The baseline signals have some inherent level of noise, and the tensor estimations might not be as accurate as the tensor estimations of a noise-free signal. The difference in the estimations between the baseline signal and the synthetic, noise-free signal can therefore be thought of as a margin of error, which indicates how accurate the baseline DTI signal is at estimating the diffusion tensor. When the differences in the estimations are small, as seen in the white and gray matter (figure 9.7 and 9.8), the margin of error is small, which substantiates the results of the baseline signals. When the differences in the estimations are more prominent, as in the CSF (figure 9.9), the margin of error is larger, suggesting that these results must be considered with some caution.

In general, there were no notable differences in the estimations between the synthetic DTI signal and the baseline DTI signal extracted from the clinical DWI-MRI data set (with the exception of in the CSF). This validates the results of the simulation experiments, as the inherent noise in the baseline signals did not affect the estimations to a large extent. Based on these results for DTI, one might make the assumption that the same results hold for DKI.

Finally, it should be emphasized that estimating kurtosis in settings of free diffusion serves no purpose. In the CSF, diffusion is assumed to be free and the water molecules have a Gaussian probability distribution. Thus, the kurtosis K should be equal to zero, which makes the second term in the expression for the DKI signal attenuation (5.2) superfluous. Nevertheless, DKI analyses in the CSF were included for completeness of the study. Figure 9.10 emphasizes this fact; the kurtosis estimations (MK, RK and AK) in the CSF are particularly inaccurate. Despite the fact that MK, RK and AK theoretically should be equal to zero in the CSF, the ground truth values are not, as seen in table 7.1. Since diffusion is assumed to be free in the CSF, the diffusion signal decays rapidly as a function of b . This leads to a monoexponential decay when the signal intensity is plotted in a logarithmic scale. The additional kurtosis term in the DKI model makes it a second-order polynomial as a function of b . Thus, the non-zero kurtosis values suggest that the DKI model fits badly to a monoexponential signal decay, which in consequence leads to inaccurate estimations of MK, RK and AK in the CSF. Moreover, a small denominator when calculating the relative error (7.9) might result in a large quotient and thereby

a large relative error, as discussed earlier.

In general, the parameter estimations seem to stabilize at an SNR of 50-60, as a further increase in SNR did not result in a noteworthy improvement in the estimations. The histograms in figure 9.4 also illustrate this fact; for high SNR, there was minimal variation between the parametric maps. At an SNR of 40 and 60, the estimations of FA and MK were almost equivalent to the estimations without added noise, while lower levels of SNR had a tendency to overestimate FA and MK. A previous study has shown that clinical b_0 -images typically have an SNR of approximately 40 in white matter at 3 T[42], which is positive considering the results of this experiment.

10.2 Parameter estimation as a function of NSA

As seen in the estimations of FA (figure 9.11), increasing the NSA from 1 to 2 had a positive effect on the parameter estimations for both DTI and DKI, with an improvement in both precision and accuracy. The improvements in precision appeared to be equivalent to a + 10 increase in SNR. Doubling the NSA also resulted in an increased precision in the estimations of MK (figure 9.12).

The improvements are, however, not as prominent as what might have been expected. The theory states that the SNR scales with $\sqrt{\text{NSA}}$ [9], which in this case means that the SNR should increase with a factor $\sqrt{2} \approx 1.4$. During the analysis part of the project, it was confirmed that the signals did, in fact, scale with $\sqrt{2}$. Yet, this relationship was not reflected in the parameter estimations. Thus, it can be speculated that the relationship $\text{SNR} \propto \sqrt{\text{NSA}}$ does not apply to the parameter estimations.

In the analyses, both models used all gradient directions with b -values up to $b = 1000$ for DTI, and $b = 3000 \text{ s/mm}^2$ for DKI. The effect of an increased NSA might have been more prominent in a smaller gradient set with fewer b -values and gradient directions, which is more sensitive to noise.

Nevertheless, the result of this experiment demonstrates the fact that the NSA and SNR are two closely related properties; when multiple signals are averaged, the signal becomes more prominent, as the background noise is being canceled out. This again leads to an increased SNR.

10.3 Parameter estimation as a function of number of gradient directions and b -values

Whereas image SNR is ultimately limited by MRI system hardware, coil design etc., the DWI sequence configuration can be directly modified by the user in order to obtain the best trade-off between quality and accuracy of resulting DWI maps and total scan time. The DWI sequence from which the baseline signals were extracted was a high-quality multi-shell sequence with an acquisition time of 8 minutes and 31 seconds (table 6.1). Reducing the number of directions for a given b -value will directly impact the total scan time. For instance, a DWI sequence with 1 b -value and 12 gradient directions had an acquisition time of 48 seconds, while a sequence with 1 b -value and 30 gradient directions had an acquisition time of 1 minute and 30 seconds.

One can generally assume that one diffusion measurement is approximately equal to one TR. As stated in table 6.1, TR = 2450 ms. Thus, by leaving out all 40 $b = 2000$ s/mm² measurements for instance, the total scan time will be reduced by 40×2450 ms = 98 seconds, which in this case is almost a 20 % reduction. Leaving out the 12 $b = 500$ s/mm² measurement reduces the total scan time by 48 seconds, or approximately 9 %.

Further, reducing the maximum b -value also allows for a shorter echo time, reducing signal loss and image distortion. Reduced scan time also reduces probability of motion induced artefacts. Overall, there is therefore a strong incentive to use the fastest DWI sequence with the lowest maximum b -value and number of b -values giving a sufficient quality for a given research question or clinical indication.

The results of the DTI analysis and the DKI analysis will be discussed in the following sections.

10.3.1 DTI analysis

The estimations of FA in white matter are illustrated in figure 9.13a. As expected, the gradient sets containing only 1 b_0 -image and either 15 or 30 gradient directions for $b = 1000$ s/mm² yielded the least precise results. These were the sets with the lowest number of gradient directions and b -values. Fewer gradient directions means fewer diffusion measurements and a less accurate estimation of the diffusion tensor, which in turn leads to less accurate estimations of the diffusion parameters. Furthermore, with only 1 b_0 -image, going from 15 to 30 gradient directions only resulted in a minor improvement in the estimations. In contrast, increasing the number of b_0 -measurements from 1 to 6 resulted in a drastic increase in precision, suggesting that the number of b_0 -measurements plays an important role in the parameter estimations. Increasing the number of b_0 -images could therefore be an effective way to

improve the SNR, since each b_0 -image only takes a few seconds to acquire.

Interestingly, there were no apparent differences between the set containing 6 b_0 and 30 $b = 1000$ s/mm², and the set containing 6 b_0 , 12 $b = 500$ and 30 $b = 1000$ s/mm². This might indicate that adding extra diffusion measurements at $b = 500$ s/mm² does not provide any new information to the DTI model fitting, as no major improvement was observed in the estimations. However, the signals from the selected voxels might also have an influence on the estimations. As seen in the white matter signals in subject 2 and 3 (figure 6.6), they are noticeable high at $b = 500$ s/mm², deviating from a monoexponential decay. The DTI model may therefore not be optimal for these selected signals. It might be speculated that the results would have turned out differently for another set of voxels.

Furthermore, it should be mentioned that adding the additional 50 gradient directions at $b = 3000$ s/mm² did not improve the estimations to a large extent in white matter. When diffusion is restricted or highly hindered, a deviation from a monoexponential decay occurs when higher b -values are used, due to the long diffusion time. The signal from the restricted or highly hindered water molecules elevates the tail of the signal decay. The monoexponential fit of the DTI model is not suitable when $b = 3000$ s/mm² is included in the gradient set, as it is not designed to pick up the deviation that occurs at higher b -values.

As figure 9.13c illustrates, adding $b = 3000$ s/mm² to the gradient set can have major negative effects on the estimations. In the CSF, diffusion is free and the signal decays rapidly, reaching the level of background noise at much lower b -values than in white and gray matter. Hence, the signal only consists of background noise at higher b -values, which makes the $b = 3000$ s/mm² measurements unnecessary. Also, even though diffusion is free in CSF and the signal decays monoexponentially, the curve shape will appear to be more biexponential when $b = 3000$ s/mm² is included, due to the noise floor. Gaussian diffusion characteristics can therefore appear non-Gaussian solely as a result of the noise floor present at higher b -values[43]. A log-conversion of the signal will therefore not become linear, which results in an underestimation of the diffusivity, and an overestimation of FA. However, accurate measurements of FA in CSF are rarely relevant in clinical practice. The estimations were included in the study for completeness.

When it comes to the estimations of MD, figure 9.14a indicates that MD is highly underestimated when $b = 3000$ s/mm² is included in the gradient set. The underestimation of MD in white matter might be due to the elevation of the signal curve that occurs at higher b -values, when signal from the restricted or highly hindered water molecules are dominating. As stated in equation (5.1), there is a linear relationship between the logarithmic DTI signal decay and the b -value, expressed as $-bD$, where the diffusion coefficient D represents the slope of the curve. Thus, with a monoexponential fit, the elevated tail at $b = 3000$ s/mm² might lead to a lower diffusion coefficient D and an underestimation of MD. This has also been shown in a previous study[44].

10.3.2 DKI analysis

The least accurate estimations of FA in white matter (figure 9.15a) occurred when the gradient set had a maximum b -value of only 1000 s/mm². As stated in equation (5.2), the logarithmic DKI signal decay can be expressed as $-bD + \frac{1}{6}b^2D^2K$, making it a second-order polynomial as a function of b . At higher b -values, the restricted or highly hindered water molecules are dominating the signal, due to the long diffusion time. This elevates the tail of the diffusion signal, and a deviation from the mono-exponential decay occurs. The second term in the DKI model is designed to pick up this deviation. However, at lower b -values, the diffusion signal is decaying mono-exponentially, which makes the DKI model unsuitable for gradient sets containing b -values up to only 1000 s/mm².

There were only minor differences between the gradient set containing 1 b_0 , 6 $b = 500$ and 25 $b = 3000$ s/mm², and the gradient set containing 1 b_0 , 12 $b = 500$ and 50 $b = 3000$ s/mm² in white matter. Thus, doubling the number of gradient directions did not improve the estimations to a large extent. Both sets had only 1 b_0 -measurement, however. Increasing the number of b_0 -measurements from 1 to 6 did improve both the precision and accuracy of the estimations. This might suggest that the number of b_0 -measurements has a large influence on the estimations. Increasing the number of b_0 -images is therefore an effective way of improving the SNR, due to the short acquisition time.

In general, step-wise improvements are observed as the number of gradient directions and b -values increases. The gradient sets with the highest number of gradient directions and b -values yielded the most precise results. However, there were no appreciable differences between the set containing 6 b_0 , 12 $b = 500$, 30 $b = 1000$ and 50 $b = 3000$ s/mm², and the full gradient set, containing 6 b_0 , 12 $b = 500$, 30 $b = 1000$, 40 $b = 2000$ and 50 $b = 3000$ s/mm². The result indicates that the additional measurements at $b = 2000$ s/mm² did not provide any new information to the DKI model fitting. A reduced acquisition time could therefore be achieved by leaving out the 40 $b = 2000$ s/mm² measurements.

Surprisingly, in the gray matter and CSF (figure 9.15b and 9.15c), estimating FA with the gradient set containing 6 b_0 , 12 $b = 500$ and 30 $b = 1000$ s/mm² was more precise than the gradient set containing 1 b_0 , 6 $b = 500$ and 25 $b = 3000$ s/mm². As discussed, the DKI model requires high b -value measurements to pick up the deviations in the signal from a monoexponential decay. However, the number of gradient directions also play an important role. At $b = 3000$ s/mm², the SNR is low due to the long diffusion time. Increasing the SNR by a sufficient number of diffusion measurements is therefore important for a precise estimation of the diffusion- and kurtosis tensor. It has also been shown in a previous study that DKI achieves minimum error in gray matter when using b -values less than 1000 s/mm², whereas maximal b -values of about 2500 s/mm² is optimal in white matter[45].

As can be seen in figure 9.16, the estimations of MD were particularly inaccurate

in the CSF, compared to in white and gray matter. In CSF, diffusion is free, and a rapid, monoexponential signal decay can be observed in figure 6.6. Thus, the kurtosis should be equal to zero in the CSF, which reduces the second term in the DKI signal expression (5.2) to zero. This again reduces the expression of the DKI model to that of the DTI model. As previously discussed, including the measurements at $b = 3000$ s/mm² might therefore result in an underestimation of the diffusion coefficient D , and hence an underestimation of MD. In contrast, the estimations of MD for the gradient set with b -values up to only 1000 s/mm² proved to be reasonably accurate in the CSF.

The gradient sets containing only 1 b_0 -measurement and no $b = 1000$ s/mm² measurements yielded the least accurate estimations of MD. Increasing the number of b_0 -measurements from 1 to 6 resulted in a noticeable improvement in both precision and accuracy. Furthermore, there were no clear distinction between the set containing 6 b_0 , 12 $b = 500$ and 50 $b = 3000$ s/mm², and the set containing, 1 b_0 , 12 $b = 500$, 30 $b = 1000$ s/mm² and 50 $b = 3000$ s/mm². This might can be interpreted as increasing b_0 from 1 to 6 is more or less equivalent to adding 30 $b = 1000$ s/mm² measurements.

The estimations of MK in white matter are shown in figure 9.17. The figure clearly illustrates the substantial errors in the estimations for the gradient set containing b -values up to only 1000 s/mm². Presumably, the large errors occur due to the fact that at lower b -values, the signal is decaying monoexponentially, and the DKI model might therefore not be suitable when $b = 1000$ s/mm² is the upper limit. Estimating kurtosis is therefore more suitable when higher b -values are included, which is where the signal starts deviating from a monoexponential decay.

All in all, only minor variations were observed across the gradient sets containing $b = 3000$ s/mm² (figure 9.18) when estimating MK. Surprisingly, removing the 40 $b = 2000$ s/mm² measurements from the full gradient set resulted in more accurate and precise estimations. This result suggests that diffusion measurements at $b = 2000$ s/mm² might be superfluous, as no apparent improvement was observed in the estimations of MK. It is, however, not certain that $b = 2000$ s/mm² is superfluous in DKI in general. Previous studies have shown that a maximum b -value of about 2000-2500 s/mm² is considered optimal for DKI[45, 46, 47]. It is important to stress that this analysis is based on the assumption that the full gradient set, containing all gradient directions and a maximum b -value of 3000 s/mm², is the optimal set. The ground truth values were therefore estimated using the full gradient set. This might have affected the results. If the ground truth values were estimated using a maximum b -value of 2000 instead of 3000 s/mm², the results could have turned out differently.

10.4 Study limitations

There are certain limitations to this study. First of all, the simulation experiments were limited to a small sample size of only 3 DW-MRI data sets, which may cause selection bias. This makes the results of the study only indicative, and therefore, no hard conclusions can be drawn. The small sample size was due to time limitations. To avoid selection bias, more data sets should be included in future studies.

Furthermore, the voxels were selected manually from each region which may have caused additional bias. However, the source data sets from which the baseline DWI signals were selected were of high quality using a state-of-the-art multi-shell DWI sequence acquired on a high-end 3T MRI system. Further, selection of baseline signals representing the three investigated tissues types were based on pre-segmentation of FA-values into three classes - ensuring that each baseline signal was a reasonable representation of the given tissue class (WM, GM and CSF).

Some of the white matter signals in subject 2 and 3 (figure 6.6) had surprisingly high signal intensities at $b = 500 \text{ s/mm}^2$, causing the signal to deviate from a monoexponential decay, which in turn may have affected the results. Nevertheless, this is a risk one must be willing to take when signals are extracted from experimental DWI data sets.

Another issue with experimental DWI data sets is the inherent level of noise in the images. The signals used in the analysis were therefore not completely noise-free. Also, we do not really know the ‘ground truth’, and our ‘apparent’ ground truth is the respective models with the full gradient sets, so all metrics of accuracy and precision are estimated relative to this. As of today, a direct interpretation of the individual kurtosis tensor elements has yet to be explored[23]. Thus, generating a representative kurtosis tensor from scratch is quite challenging. Completely synthetic DKI signals could therefore not be generated due to the mathematical complexity of the kurtosis tensor.

When it comes to the simulation part on the project, the number of iterations was set to $n = 1000$. In Monte Carlo simulations, the number of iterations is usually higher, and increasing n to 10 000 could have been advantageous. A smaller number of iterations results in larger variations within the distributions, which may have an influence on the results. But then, increasing the number of iterations would lead to a more time demanding simulation process, and a need to speed up the code with e.g. parallel programming.

Finally, any scientific work that relies on measurements is prone to uncertainties. The uncertainty in computer simulations is bounded to the floating point precision of the computer. Modern computers have a floating point precision in the order of 10^{-38} , and round-off errors may accumulate and become significant when multiple iterations is required for the calculation of a value[48]. However, in this study it was assumed that other uncertainties were of higher significance, considering the level of

noise in the baseline signals.

10.5 Conclusion

The presence of noise alters the parameter estimations considerably. In white matter, the DTI model was found to be most noise sensitive. In gray matter and CSF, the DKI model was most noise sensitive. In general, the DKI model performs badly in highly isotropic voxels, presumably due to the fact that the kurtosis is equal to zero in settings of free diffusion. The results suggest that DKI is preferred in white matter analysis, while DTI is preferred in gray matter analysis. The parameter estimations stabilized at an SNR of 50-60, as a further increase in SNR did not improve the estimations to a large extent. Increasing the number of signal averages can also have a positive effect on the parameter estimations in both models, as an increase in NSA also increases the SNR.

The number of b_0 -images can have an important influence on the parameter estimations. More than 1 b_0 -image is generally preferred. The results of the gradient sets containing 6 b_0 -images were more precise than gradient sets containing only 1 b_0 -image.

The DTI analysis showed that including diffusion measurements at $b = 3000$ s/mm² did not improve the estimations noteworthy. Measuring diffusion at $b = 3000$ s/mm² is therefore not necessary for DTI. 6 b_0 -images and 30 $b = 1000$ s/mm² measurements appears to be the optimal choice for DTI, with an acceptable error range and number of gradient directions and b -values. Leaving out the 12 $b = 500$ s/mm² measurements can therefore reduce the scan time by 9 %.

The DKI analysis showed that an upper bound of $b = 1000$ s/mm² in general should be avoided in DKI, due to the large errors that may arise in the kurtosis estimations. 6 b_0 -images, 12 $b = 500$, 30 $b = 1000$ and 50 $b = 3000$ s/mm² measurements seems to be the optimal gradient set for DKI, which makes the 40 $b = 2000$ s/mm² measurements redundant, reducing the scan time by almost 20 %.

All in all, this study has shown that the number of gradient directions and b -values can have a big impact on the parameter estimations in both DTI and DKI. However, the small sample size makes the results of the study only indicative. A more extensive study should be conducted in the future, including a larger range of DWI data sets to fully understand the effect of SNR, NSA, number of gradient directions and b -values on the parameter estimations.

Bibliography

- [1] NobelPrize.org. *The Nobel Prize in Physics 1952*.
<https://www.nobelprize.org/prizes/physics/1952/summary/>
(visited on 12/22/2018).
- [2] H. H. Mitchell, T. S. Hamilton, F. R. Steggerda, and H. W. Bean. “The chemical composition of the adult human body and its bearing on the biochemistry of growth”. In: *Journal of Biological Chemistry* 158.3 (1945), pp. 625–637.
- [3] Els Fieremans, Jens H. Jensen, and Joseph A. Helpert. “White matter characterization with diffusional kurtosis imaging”. eng. In: *Neuroimage* 58.1 (2011), pp. 177–188.
- [4] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. “MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders.” eng. In: *Radiology* 161.2 (1986), pp. 401–407.
- [5] Hui Zhang, Torben Schneider, Claudia A. Wheeler-Kingshott, and Daniel C. Alexander. “NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain”. eng. In: *NeuroImage* 61.4 (2012), pp. 1000–1016.
- [6] Jens H. Jensen, Joseph A. Helpert, Anita Ramani, Hanzhang Lu, and Kyle Kaczynski. “Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging”. eng. In: *Magnetic Resonance in Medicine* 53.6 (2005), pp. 1432–1440.
- [7] Brian Duignan. *Occam’s razor*. 2018.
<https://www.britannica.com/topic/Occams-razor>
(visited on 10/30/2019).
- [8] Stephen Hawking. *On the Shoulders of Giants*. eng. United States: Running Press, 2002, p. 731.
- [9] Atle Bjørnerud. *The Physics of Magnetic Resonance Imaging*. eng. Department of Physics, University of Oslo, 2008.
- [10] Igor N. Serdyuk, Nathan R. Zaccai, and Joseph Zaccai. In: *Methods in Molecular Biophysics: Structure, Dynamics, Function*. .Part J. Cambridge University Press, 2007. Chap. J1 and J2, pp. 971–1031.
- [11] David J. Griffiths. *Introduction to Quantum Mechanics*. eng. Upper Saddle River, N.J, 1995.

- [12] Trygve Holtebekk. *Zeemaneffekt*. 2009, February 15th. <https://snl.no/zeemaneffekt> (visited on 07/11/2019).
- [13] James Keeler. *Understanding NMR Spectroscopy*. 2nd ed. Wiley, 2010. Chap. 2.
- [14] Susumu Mori. *Introduction to Diffusion Tensor Imaging*. eng. 1st ed. Elsevier, 2007.
- [15] Tuva Roaldsdatter Hope. “MRI-basert analyse av diffusjonsparametere”. nor. Thesis. Department of Physics, University of Oslo, 2009.
- [16] Peter B. Kingsley. “Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors”. In: *Concepts in Magnetic Resonance Part A* 28.2 (2006), pp. 101–122.
- [17] L.J. Erasmus, D. Hurter, M. Naude, H.G. Kritzinger, and S. Acho. “A short overview of MRI artefacts”. In: *South African Journal of Radiology* 8.2 (2004).
- [18] O. Dietrich, S. Heiland, T. Benner, and K. Sartor. “Reducing motion artefacts in diffusion-weighted MRI of the brain: efficacy of navigator echo correction and pulse triggering”. In: *Neuroradiology* 42.2 (2000), pp. 85–91.
- [19] Heidi Johansen-Berg and Timothy E. J. Behrens. *Diffusion MRI: From Quantitative Measurement to In vivo Neuroanatomy*. eng. 2nd ed. Academic Press, 2013, pp. 24, 26.
- [20] P.J. Basser, J. Mattiello, and D. Lebihan. “Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo”. eng. In: *Journal of Magnetic Resonance, Series B* 103.3 (1994), pp. 247–254.
- [21] Christian Beaulieu. “The basis of anisotropic water diffusion in the nervous system - a technical review”. eng. In: *NMR in Biomedicine* 15.7-8 (2002), pp. 435–455.
- [22] Andrew J. Steven, Jiachen Zhuo, and Elias R. Melhem. “Diffusion kurtosis imaging: an emerging technique for evaluating the microstructural environment of the brain.” eng. In: *AJR. American journal of roentgenology* 202.1 (2014), pp. 26–33.
- [23] Ed X. Wu and Matthew M. Cheung. “MR diffusion kurtosis imaging for neural tissue characterization”. eng. In: *NMR in Biomedicine* 23.7 (2010), pp. 836–848.
- [24] Ileana O. Jelescu and Matthew D. Budde. “Design and Validation of Diffusion MRI Models of White Matter”. eng. In: *Frontiers in Physics* 5 (2017), p. 61.
- [25] Nan-Jie Gong, Chun-Sing Wong, Chun-Chung Chan, Lam-Ming Leung, and Yiu-Ching Chu. “Correlations between microstructural alterations and severity of cognitive deficiency in Alzheimer’s disease and mild cognitive impairment: a diffusional kurtosis imaging study”. eng. In: *Magnetic Resonance Imaging* 31.5 (2013), pp. 688–694.
- [26] Maria F. Falangola, Jens H. Jensen, James S. Babb, Caixia Hu, Francisco X. Castellanos, Adriana Di Martino, Steven H. Ferris, and Joseph A. Helpert. “Age-related non-Gaussian diffusion patterns in the prefrontal brain”. In: *Journal of Magnetic Resonance Imaging* 28.6 (2008), pp. 1345–1350.

- [27] Jens H. Jensen and Joseph A. Helpert. “MRI quantification of non-Gaussian water diffusion by kurtosis analysis”. In: *NMR in Biomedicine* 23.7 (2010), pp. 698–710.
- [28] Manjón José V., Pierrick Coupé, Luis Concha, Antonio Buades, D. Louis Collins, and Montserrat Robles. “Diffusion Weighted Image Denoising Using Overcomplete Local PCA”. eng. In: *PLOS ONE* 8.9 (2013), pp. 1–12.
- [29] Elias Kellner, Bibek Dhital, Valerij G. Kiselev, and Marco Reiser. “Gibbs-ringing artifact removal based on local subvoxel-shifts”. eng. In: *Magnetic Resonance in Medicine* 76.5 (2016), pp. 1574–1581.
- [30] Stephen M. Smith, Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E.J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, and Paul M. Matthews. “Advances in functional and structural MR image analysis and implementation as FSL”. In: *NeuroImage* 23 (2004). Mathematics in Brain Imaging, pp. 208–219.
- [31] Jesper L.R. Andersson, Stefan Skare, and John Ashburner. “How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging”. In: *NeuroImage* 20.2 (2003), pp. 870–888.
- [32] Jesper L.R. Andersson and Stamatios N. Sotiropoulos. “An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging”. In: *NeuroImage* 125 (2016), pp. 1063–1078.
- [33] Jelle Veraart, Jan Sijbers, Stefan Sunaert, Alexander Leemans, and Ben Jeurissen. “Weighted linear least squares estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls”. English. In: *Neuroimage* 81 (2013).
- [34] Ali Tabesh, Jens H. Jensen, Babak A. Ardekani, and Joseph A. Helpert. “Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging”. eng. In: *Magnetic Resonance in Medicine* 65.3 (2011), pp. 823–836.
- [35] Daniel C. Alexander and Gareth J. Barker. “Optimal imaging parameters for fiber-orientation estimation in diffusion MRI”. In: *NeuroImage* 27.2 (2005), pp. 357–367.
- [36] Jelle Veraart, Dirk H. J. Poot, Wim Van Hecke, Ines Blockx, Annemie Van Der Linden, Marleen Verhoye, and Jan Sijbers. “More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging”. eng. In: *Magnetic Resonance in Medicine* 65.1 (2011), pp. 138–145.
- [37] Mustapha Bouhrara, David A. Reiter, Hasan Celik, Jean-Marie Bonny, Vanessa Lukas, Kenneth W. Fishbein, and Richard G. Spencer. “Incorporation of rician noise in the analysis of biexponential transverse relaxation in cartilage using a multiple gradient echo sequence at 3 and 7 Tesla”. In: *Magnetic Resonance in Medicine* 73.1 (2015), pp. 352–366.
- [38] H. Gudbjartsson and S. Patz. “The rician distribution of noisy mri data”. In: *Magnetic Resonance in Medicine* 34.6 (1995), pp. 910–914.
- [39] David C Hoaglin, Mosteller Frederick, and John W. Tukey. *Understanding robust and exploratory data analysis*. eng. New York, 1983.

- [40] K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference*. 2002.
- [41] Adam W. Anderson. “Theoretical analysis of the effects of noise on diffusion tensor imaging”. In: *Magnetic Resonance in Medicine* 46.6 (2001), pp. 1174–1188.
- [42] Seongjin Choi, Dustin T. Cunningham, Francisco Aguila, John D. Corrigan, Jennifer Bogner, W. Jerry Mysiw, Michael V. Knopp, and Petra Schmalbrock. “DTI at 7 and 3 T: systematic comparison of SNR and its influence on quantitative metrics”. In: *Magnetic Resonance Imaging* 29.6 (2011), pp. 739–751.
- [43] Derek K. Jones and Peter J. Basser. ““Squashing peanuts and smashing pumpkins”: How noise distorts diffusion-weighted MR data”. In: *Magnetic Resonance in Medicine* 52.5 (2004), pp. 979–993.
- [44] Lise Linn Løkken. “MRI based detection of time-of-day variation in white matter microstructure. Are more complex diffusion models better?”. eng. Thesis. Department of Physics, University of Oslo, 2019.
- [45] Andrey Chuhutin, Brian Hansen, and Sune Nørhøj Jespersen. “Precision and accuracy of diffusion kurtosis estimation and the influence of b-value selection”. In: *NMR in Biomedicine* 30.11 (2017), e3777.
- [46] Brian Hansen and Sune N. Jespersen. “Recent Developments in Fast Kurtosis Imaging”. In: *Frontiers in Physics* 5 (2017), p. 40.
- [47] Yen-Shu Kuo, Shun-Chung Yang, Hsiao-Wen Chung, and Wen-Chau Wu. “Toward quantitative fast diffusion kurtosis imaging with b-values chosen in consideration of signal-to-noise ratio and model fidelity”. In: *Medical Physics* 45.2 (2018), pp. 605–612.
- [48] Ole Gunnar Johansen. “DCE-MRI Pharmacokinetic Model Optimization and Implications for Brain Cancer Imaging”. eng. Thesis. Department of Physics, University of Oslo, 2018.

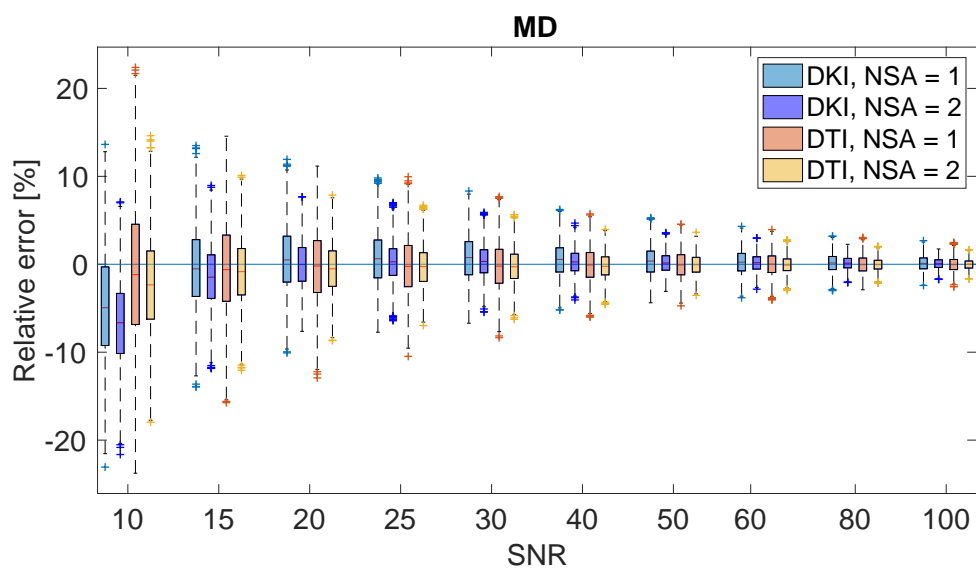
Appendices

Appendix A

Additional results

A.1 Parameter estimation as a function of NSA

MD



(a) MD in white matter

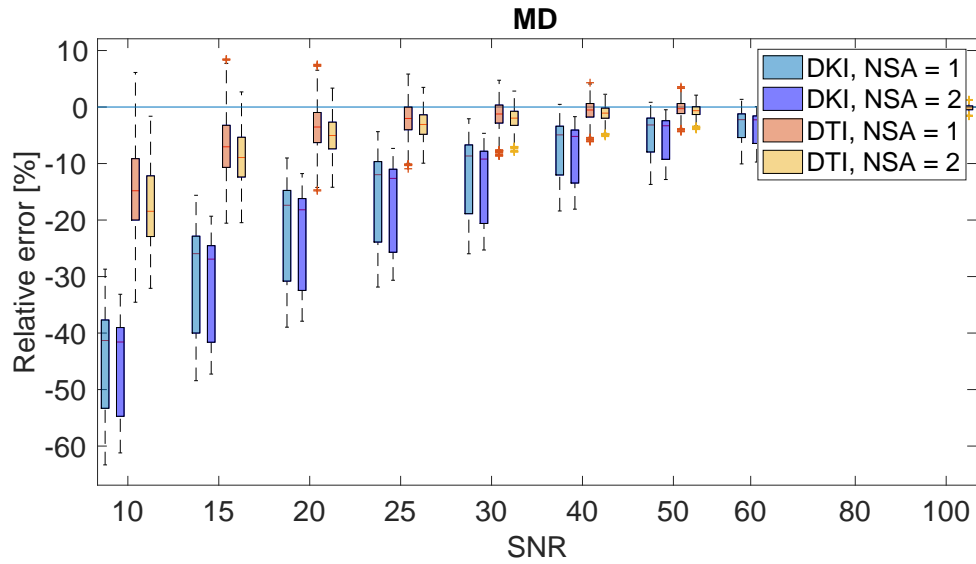
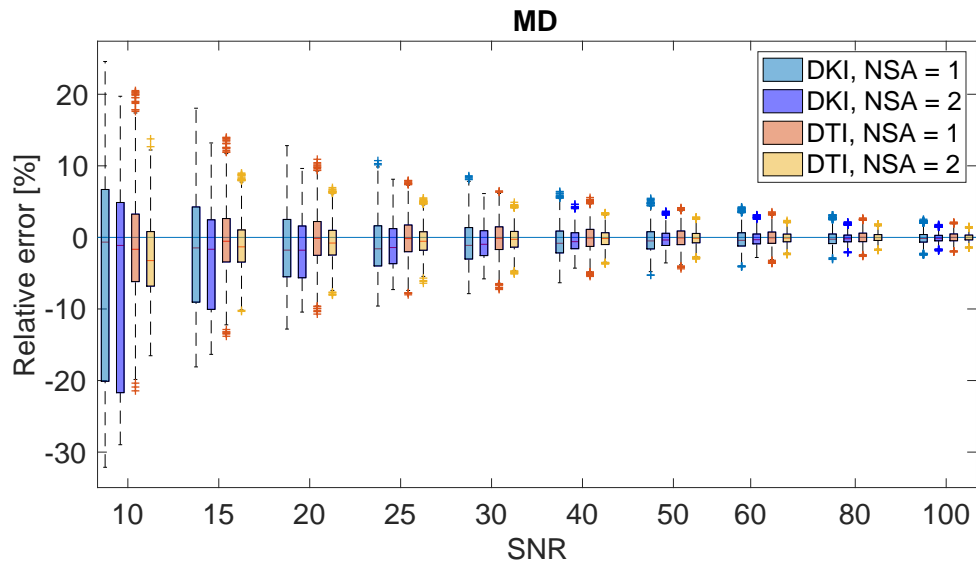
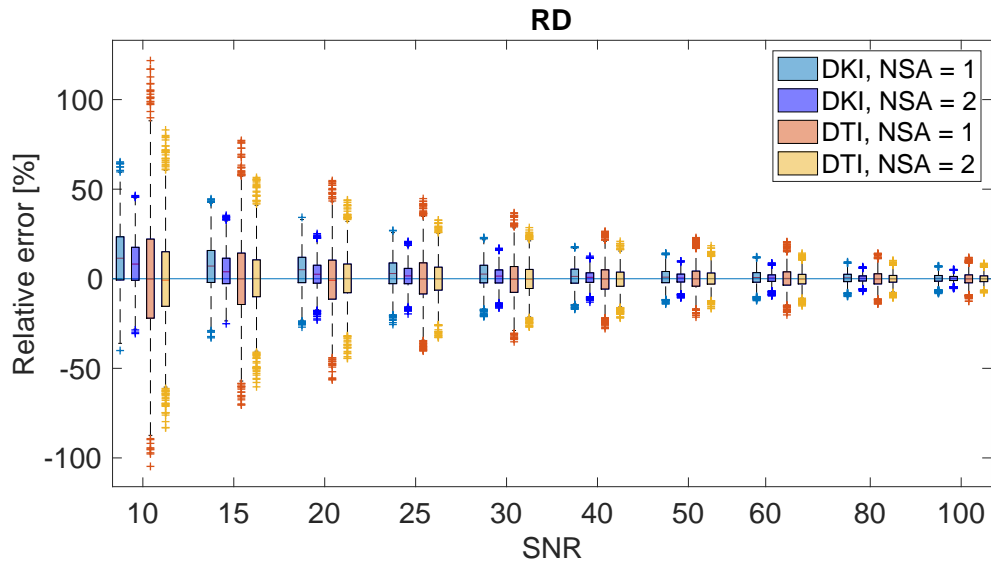
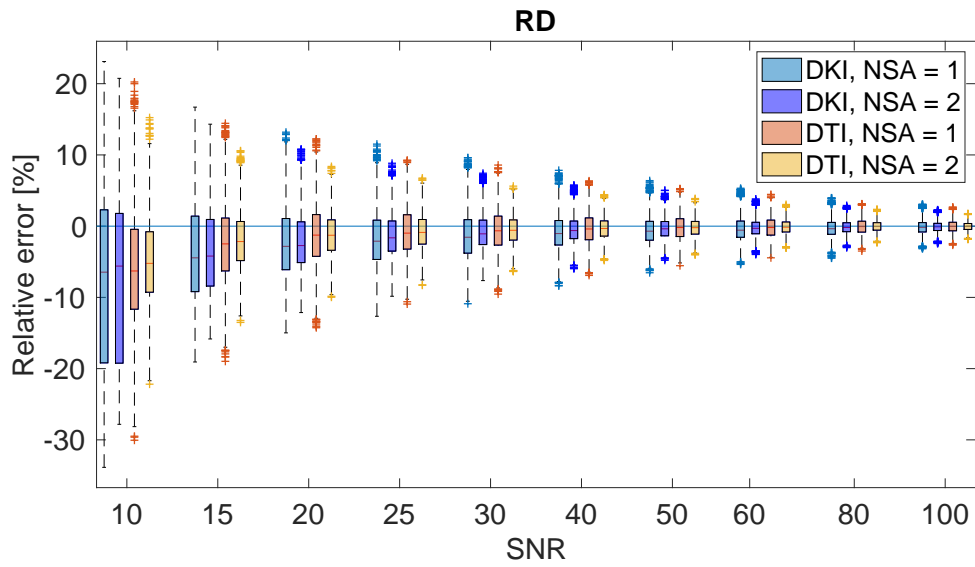


Figure A.1: Relative error in the estimations of MD from signals simulating white matter, gray matter and CSF, for NSA = 1 and NSA = 2.

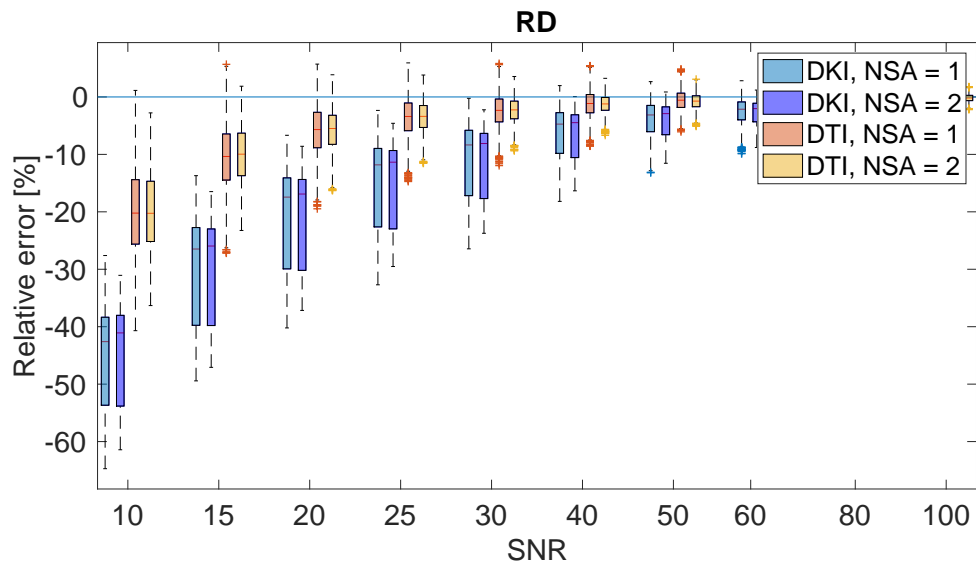
RD



(a) RD in white matter



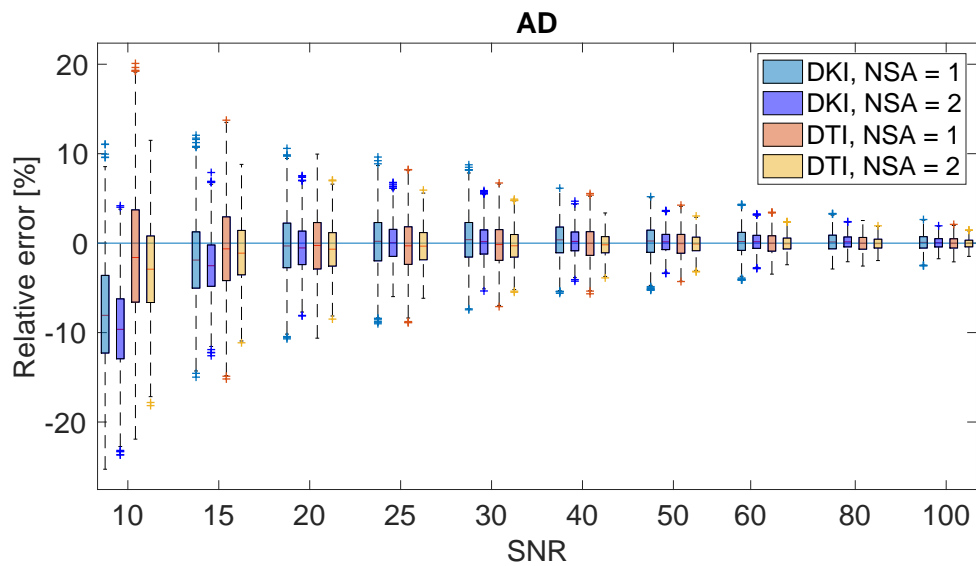
(b) RD in gray matter



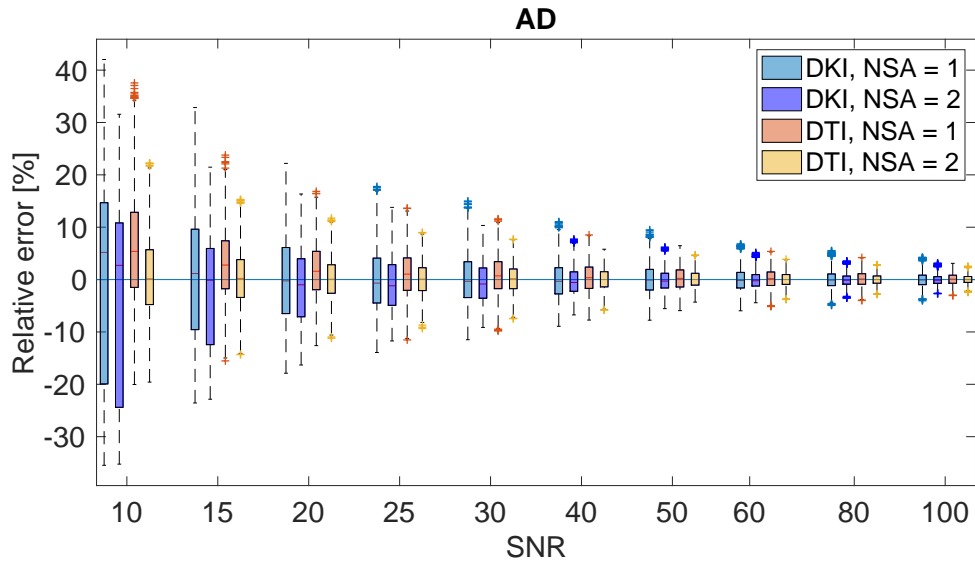
(c) RD in CSF

Figure A.2: Relative error in the estimations of RD from signals simulating white matter, gray matter and CSF, for NSA = 1 and NSA = 2.

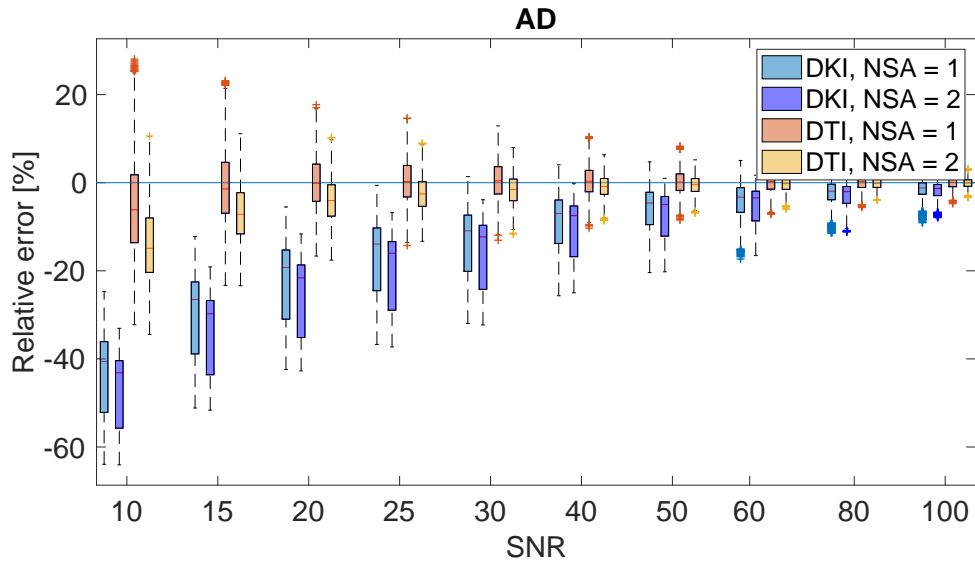
AD



(a) AD in white matter



(b) AD in gray matter



(c) AD in CSF

Figure A.3: Relative error in the estimations of AD from signals simulating white matter, gray matter and CSF, for NSA = 1 and NSA = 2.

RK

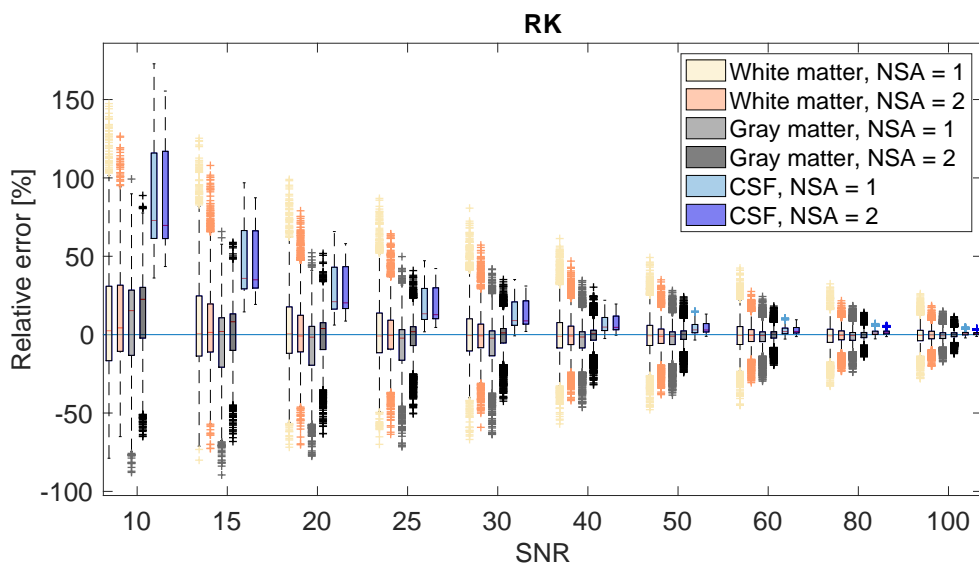


Figure A.4: Estimations of RK as a function of SNR for signals simulating white matter, gray matter and CSF.

AK

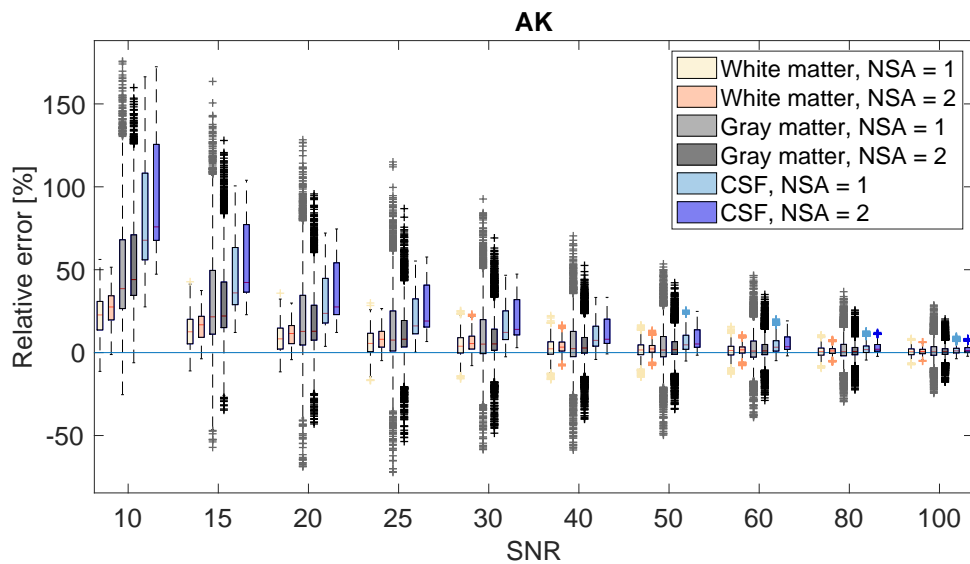


Figure A.5: Estimations of AK as a function of SNR for signals simulating white matter, gray matter and CSF.

A.2 Parameter estimation as a function of number of gradient directions and b -values

DTI analysis

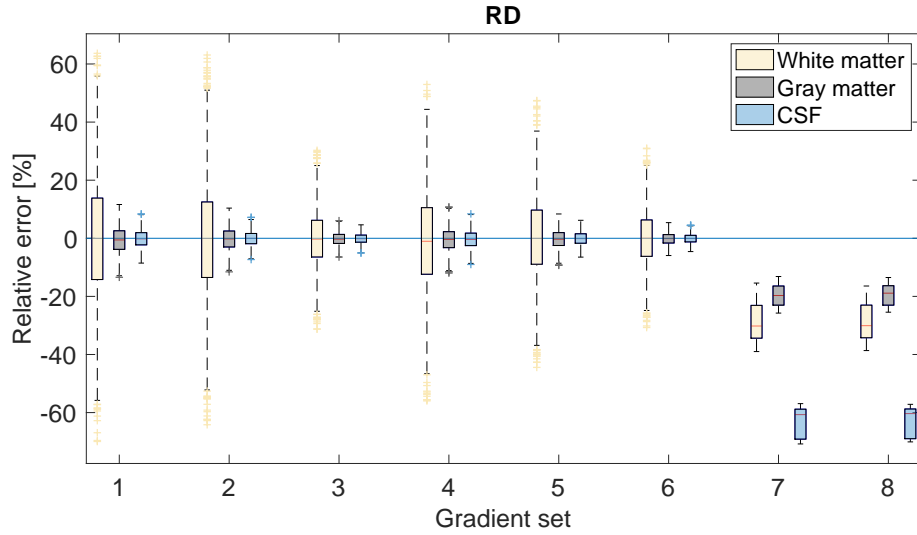


Figure A.6: Estimations of RD in white matter, gray matter and CSF for various gradient sets using the DTI model.

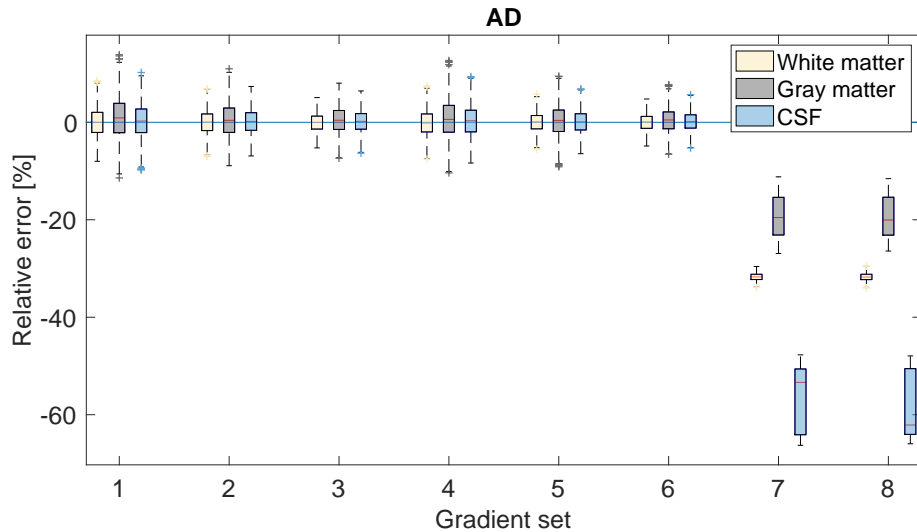


Figure A.7: Estimations of AD in white matter, gray matter and CSF for various gradient sets using the DTI model.

DKI analysis

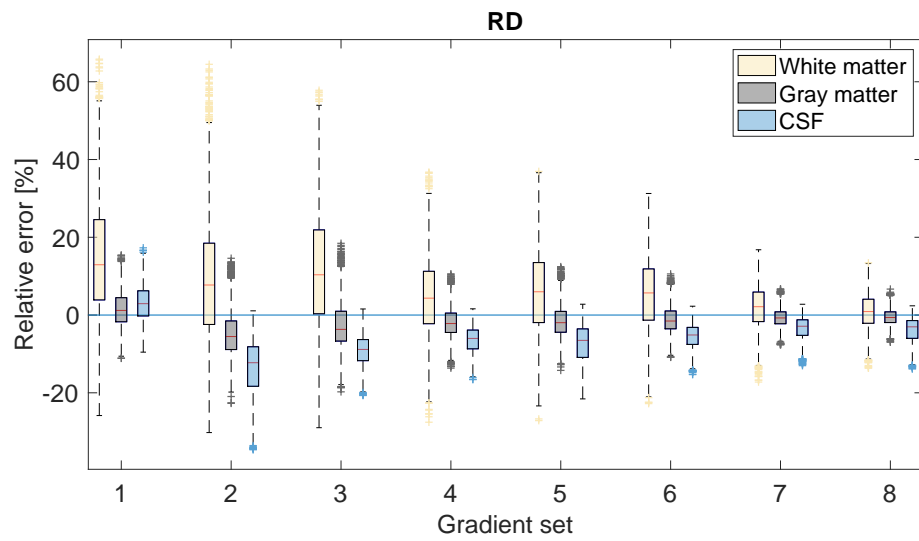


Figure A.8: Estimations of RD in white matter, gray matter and CSF for various gradient sets using the DKI model.

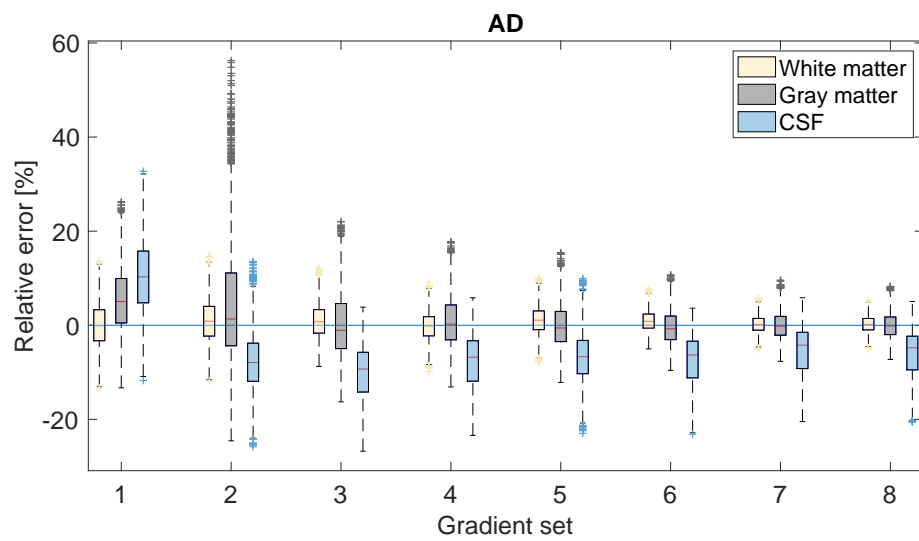
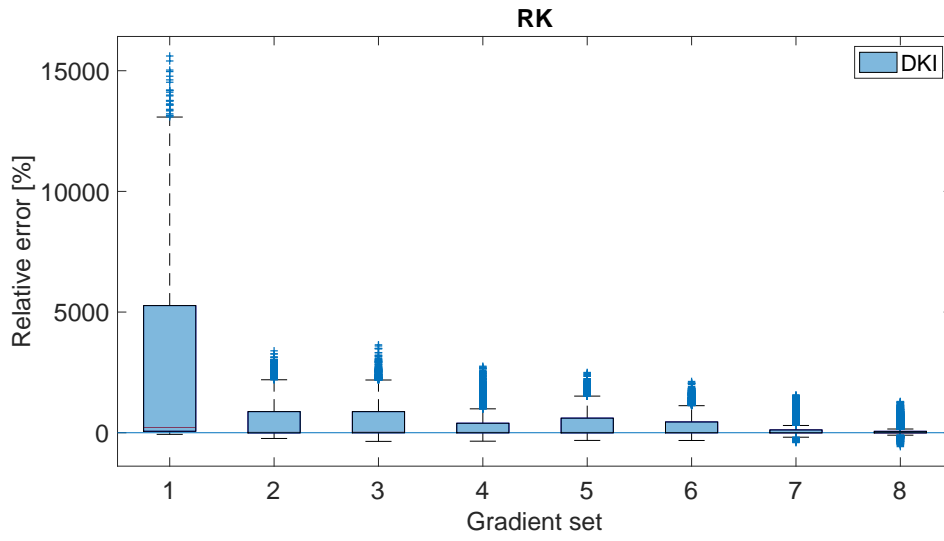
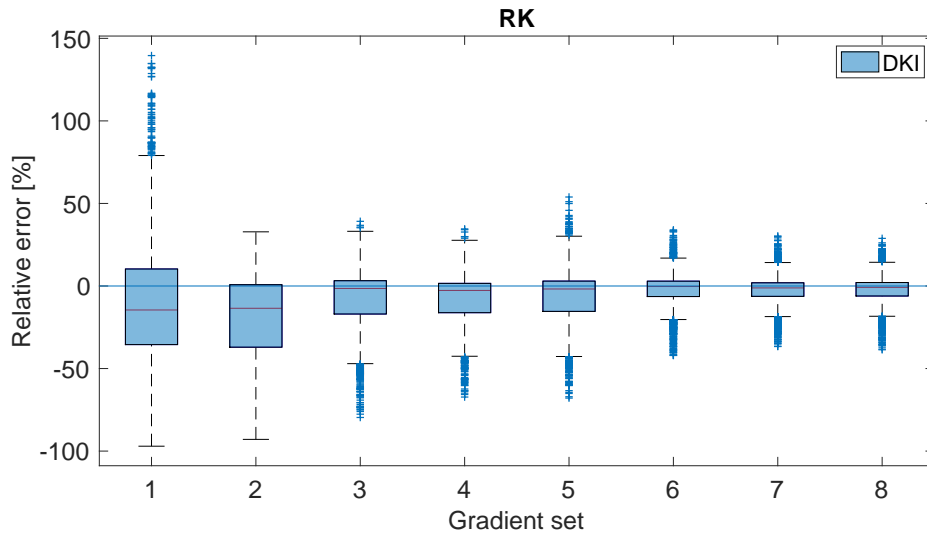


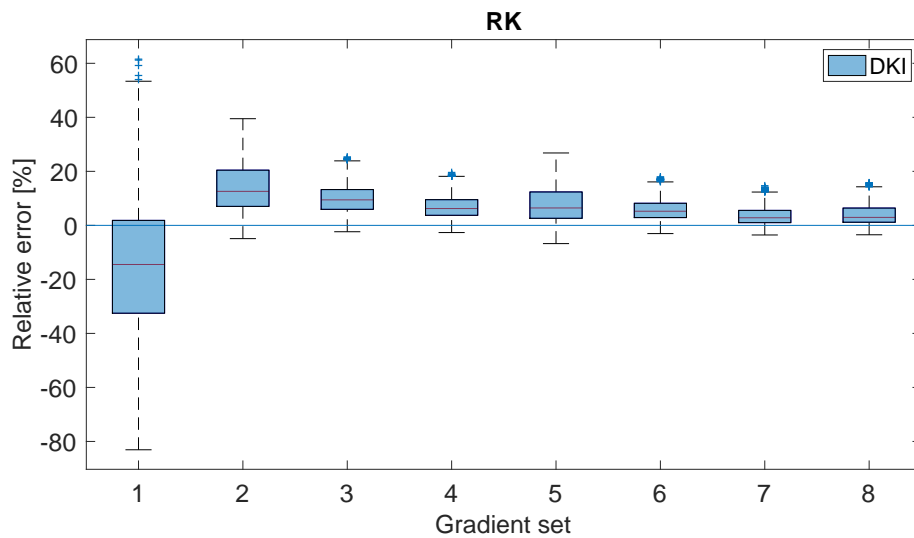
Figure A.9: Estimations of AD in white matter, gray matter and CSF for various gradient sets using the DKI model.



(a) RK in white matter



(b) RK in gray matter



(c) RK in CSF

Figure A.10: Estimations of RK in white matter, gray matter and CSF for various gradient sets using the DKI model.

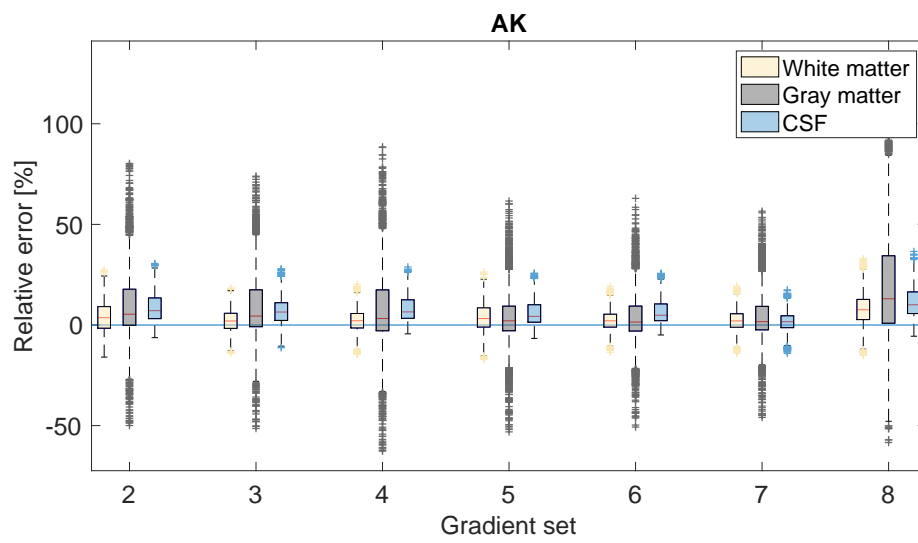


Figure A.11: Estimations of AK in white matter, gray matter and CSF for various gradient sets. Set 1 has been removed from the plot due to the large errors in the estimations.

Appendix B

Matlab code

The following sections contain snippets of the most relevant parts of the code. The code in its entirety can be found at <https://github.com/livewj/Masteroppgave>

B.1 dki_fit.m

Diffusion and kurtosis tensor estimation for the DKI signal.

```
function [b0, dt, w, w_wlls] = dki_fit(dwi, grad, mask,
constraints, outliers, maxbval)
% Diffusion Kurtosis Imaging tensor estimation using
% (constrained) weighted linear least squares estimation
% -----
% please cite: Veraart, J.; Sijbers, J.; Sunaert, S.;
% Leemans, A. & Jeurissen, B.,
% Weighted linear least squares estimation of
% diffusion MRI parameters:
% strengths, limitations, and pitfalls.
% Neurolmage, 2013, 81, 335–346
% -----
...
% Tensor fit
[D_ind, D_cnt] = createTensorOrder(2);
[W_ind, W_cnt] = createTensorOrder(4);

bS = ones(ndwis, 1);
bD = D_cnt(ones(ndwis, 1), :) .* grad(:, D_ind(:,
1)) .* grad(:, D_ind(:, 2));
bW = W_cnt(ones(ndwis, 1), :) .* grad(:, W_ind(:,
1)) .* grad(:, W_ind(:, 2)) .* grad(:, W_ind(:,
3)) .* grad(:, W_ind(:, 4));
```



```

b = [bS, -bval(:, ones(1, 6)).*bD, (bval(:, ones(1,
15)).^2/6).*bW]; % D and W tensor

% Unconstrained LLS fit
dt = b\log(dwi); % Moore–Penrose pseudoinverse
w = exp(b*dt); % Estimated signal

nvoxels = size(dwi,2);

% WLS fit initialized with LLS
if any(constraints)
    dir = [0.382517725304416 -0.748614094922528
0.541532202838631; ... ]; % Gradient directions
    ndir = size(dir, 1);
    C = [];
    if constraints(1)>0 % TRUE
        C = [C; [zeros(ndir, 1), D_cnt(ones(ndir, 1),
:).*dir(:, D_ind(:, 1)).*dir(:, D_ind(:, 2)),
zeros(ndir, 15)]];
    end
    if constraints(2)>0 %TRUE
        C = [C; [zeros(ndir, 7), W_cnt(ones(ndir, 1),
:).*dir(:, W_ind(:, 1)).*dir(:, W_ind(:,
2)).*dir(:, W_ind(:, 3)).*dir(:, W_ind(:, 4)]];
    end
    if constraints(3)>0 %TRUE
        C = [C; [zeros(ndir, 1),
3/max(bval)*D_cnt(ones(ndir, 1),
:).*dir(:, D_ind(:, 1)).*dir(:, D_ind(:, 2)), -
W_cnt(ones(ndir, 1), :).*dir(:, W_ind(:,
1)).*dir(:, W_ind(:, 2)).*dir(:, W_ind(:,
3)).*dir(:, W_ind(:, 4)]];
    end
    d = zeros([1, size(C, 1)]);
    options = optimset('Display', 'off', 'Algorithm',
'interior-point', 'MaxIter', 5000, 'TolCon', 1e-10,
'TolFun', 1e-10, 'ToIX', 1e-10, 'MaxFunEvals', 5000);
    parfor i = 1:nvoxels
        try
            in_ = outliers(:, i) == 0;
            wi = w(:, i);
            Wi = diag(wi(in_));
            dt(:, i) = lsqlin(Wi*b(in_,
:), Wi*log(dwi(in_, i)), -C, d,
[], [], [], [], [], options);
        catch
            dt(:, i) = 0;
        end
    end
else % Unconstrained
    parfor i = 1:nvoxels
        in_ = outliers(:, i) == 0;

```

```

        b_ = b(in_, :);
        if isempty(b_) || cond(b(in_, :))>1e15
            dt(:, i) = NaN
        else
            wi = w(:, i); Wi = diag(wi(in_));
            logdwii = log(dwi(in_, i));
            dt(:, i) = (Wi*b_)\(Wi*logdwii);
        end
    end
end

b0 = exp(dt(1,:)); % S0 signal
dt_0 = dt(1,:); % dt with S0

dt = dt(2:22, :);
D_apprSq = 1./((sum(dt([1 4 6], :), 1)/3).^2);
dt(7:21, :) = dt(7:21, :) .* D_apprSq(ones(15,1), :);

dt_wlls = [dt_0; dt];
w_wlls = exp(b*dt_wlls); % Signal estimated from WLLS

end

function [X, cnt] = createTensorOrder(order)
X = nchoosek(kron([1, 2, 3], ones(1, order)), order);
X = unique(X, 'rows');
for i = 1:size(X, 1)
    cnt(i) = factorial(order) / factorial(nnz(X(i, :) ==1))/
        factorial(nnz(X(i, :) ==2))/ factorial(nnz(X(i, :)
        ==3));
end
end

function [s, mask] = vectorize(S, mask)
if nargin == 1
    mask = ~isnan(S(:,:,:,1));
end
if ismatrix(S)
    n = size(S, 1);
    [x, y, z] = size(mask);
    s = NaN([x, y, z, n], 'like', S);
    for i = 1:n
        tmp = NaN(x, y, z, 'like', S);
        tmp(mask(:)) = S(i, :);
        s(:,:,:,i) = tmp;
    end
else
    for i = 1:size(S, 4)
        Si = S(:,:,:,i);
        s(i, :) = Si(mask(:));
    end
end
end
end

```

B.2 dti_fit.m

Diffusion tensor estimation for the DTI signal.

```
function [b0, dt, w, w_wlls] = dti_fit(dwi, grad, mask,
    constraints, outliers, maxbval)
% Diffusion Tensor Imaging tensor estimation using
% (constrained) weighted linear least squares estimation
% -----
% please cite: Veraart, J.; Sijbers, J.; Sunaert, S.;
% Leemans, A. & Jeurissen, B.,
% Weighted linear least squares estimation of
% diffusion MRI parameters:
% strengths, limitations, and pitfalls.
% NeuroImage, 2013, 81, 335–346
% -----
...
%% Tensor fit
[D_ind, D_cnt] = createTensorOrder(2);

bS = ones(ndwis, 1);
bD = D_cnt(ones(ndwis, 1), :) .* grad(:, D_ind(:,
    1)) .* grad(:, D_ind(:, 2)));

b = [bS, -bval(:, ones(1, 6)) .* bD]; %D tensor

% Unconstrained LLS fit
dt = b \ log(dwi); % = pinv(b) * log(dwi) Moore–Penrose
pseudoinverse
w = exp(b * dt); % signal

nvoxels = size(dwi, 2);

% Constrained WLLS fit initialized with LLS
if any(constraints)
    dir = [0.382517725304416 -0.748614094922528
        0.541532202838631; ... ];
    ndir = size(dir, 1);
    C = [];
    if constraints(1) > 0 %TRUE
        %C = [C; [zeros(ndir, 1), D_cnt(ones(ndir, 1),
            :) .* dir(:, D_ind(:, 1)) .* dir(:, D_ind(:, 2))),
            zeros(ndir, 15)]];
        C = [C; [zeros(ndir, 1), D_cnt(ones(ndir, 1),
            :) .* dir(:, D_ind(:, 1)) .* dir(:, D_ind(:, 2))]];
    end
    if constraints(2) > 0
```

```

        C = [C; [zeros(ndir, 7), W_cnt(ones(ndir, 1),
        :).*dir(:,W_ind(:, 1)).*dir(:,W_ind(:,
        2)).*dir(:,W_ind(:, 3)).*dir(:,W_ind(:, 4))]];
    end
    if constraints(3)>0
        C = [C; [zeros(ndir, 1),
        3/max(bval)*D_cnt(ones(ndir, 1),
        :).*dir(:,D_ind(:, 1)).*dir(:,D_ind(:, 2)), -
        W_cnt(ones(ndir, 1), :).*dir(:,W_ind(:,
        1)).*dir(:,W_ind(:, 2)).*dir(:,W_ind(:,
        3)).*dir(:,W_ind(:, 4))]];
    end

    d = zeros([1, size(C, 1)]);
    options = optimset('Display', 'off', 'Algorithm',
        'interior-point', 'MaxIter', 5000, 'TolCon', 1e-10,
        'TolFun', 1e-10, 'ToIX', 1e-10, 'MaxFunEvals', 5000);

    parfor i = 1:nvoxels
        try
            in_ = outliers(:, i) == 0;
            wi = w(:, i);
            Wi = diag(wi(in_));
            dt(:, i) = lsqlin(Wi*b(in_,
            :), Wi*log(dwi(in_, i)), -C, d,
            [], [], [], [], [], options);
        catch
            dt(:, i) = 0;
        end
    end
else % Unconstrained WLLS
    parfor i = 1:nvoxels
        in_ = outliers(:, i) == 0;
        b_ = b(in_, :);
        if isempty(b_) || cond(b(in_, :))>1e15
            dt(:, i) = NaN
        else
            wi = w(:, i); Wi = diag(wi(in_));
            logdwii = log(dwi(in_, i));
            dt(:, i) = (Wi*b_)\(Wi*logdwii);
        end
    end
end

b0 = exp(dt(1,:));
dt_0 = dt(1,:); % dt with S0

dt = dt(2:7, :);
%D_apprSq = 1./(sum(dt([1 4 6], :), 1)/3).^2;
%dt(7:21,:) = dt(7:21,:) .* D_apprSq(ones(15,1), :);

dt_wlls = [dt_0; dt];
w_wlls = exp(b*dt_wlls); % Signal estimated from WLLS

```

```
b0 = vectorize(b0, mask);
dt = vectorize(dt, mask);
```

B.3 dki_parameters.m

Parameter estimations of MD, FA, RD, AD, MK, RK and AK used in the DKI analysis.

```
function [fa, md, rd, ad, fe, mk, rk, ak, ev1, ev2, ev3] =
dki_parameters(dt, mask)
% diffusion and kurtosis tensor parameter calculation
% please cite: Veraart et al.
%           More Accurate Estimation of Diffusion Tensor
%           Parameters Using Diffusion Kurtosis Imaging,
%           MRM 65 (2011): 138–145.
% -----

% DTI parameters

for i = 1:nvoxels
    DT = dt([1:3 2 4 5 3 5 6], i);
    DT = reshape(DT, [3 3]);
    try
        [eigvec, eigval] = eigs(DT);
        eigval = diag(eigval);
    catch
        eigvec = NaN(3, 3);
        eigval = NaN(3, 1);
    end
    [eigval, idx] = sort(eigval, 'descend');
    eigvec = eigvec(:, idx);
    l1(i) = eigval(1,:);
    l2(i) = eigval(2,:);
    l3(i) = eigval(3,:);

    e1(:, i) = eigvec(:, 1);
end
md = (l1+l2+l3)/3;
rd = (l2+l3)/2;
ad = l1;
fa = sqrt(1/2) .* sqrt((l1-l2).^2+(l2-l3).^2+(l3-l1).^2) ./ ...
    sqrt(l1.^2+l2.^2+l3.^2);

% DKI parameters
dirs = get256dirs();
akc = AKC(dt, dirs);
```

```

mk = mean(akc);
ak = zeros([1, size(e1,2)]);
rk = zeros([1, size(e1,2)]);

parfor i = 1:nvoxels
    dirs = [e1(:,i), -e1(:,i)]';
    akc = AKC(dt(:,i), dirs);
    ak(i) = mean(akc);
    dirs = radialsampling(e1(:,i), 256)';
    akc = AKC(dt(:,i), dirs);
    rk(i) = mean(akc);
end

```

B.4 dti_parameters.m

Parameter estimations of MD, FA, RD and AD used in the DTI analysis.

```

function [md, fa, rd, ad, l1, l2, l3] = dti_parameters(D)
% Calculates mean diffusion (md), fractional anisotropy (fa)
% radial diffusivity (rd), axial diffusivity (ad)
% from diffusion tensor D

[eigvec, eigval] = eigs(D);
eigval = diag(eigval);

l1 = eigval(1);
l2 = eigval(2);
l3 = eigval(3);

md = (l1+l2+l3)/3;
rd = (l2+l3)/2;
ad = l1;

A = (l1 - md)^2 + (l2 - md)^2 + (l3 - md)^2;
B = l1^2 + l2^2 + l3^2;
fa = sqrt(3*A/(2*B));

end

```

B.5 percentile.m

Tissue segmentation based on 10th and 90th percentile of the FA values.

```

% Load data

```

```

dwi = niftiread();
[nx, ny, nz, np] = size(dwi);
bvecs = importdata();
bvals = importdata();
grad = [bvecs' bvals'];

% Generate FA map
[b0, dt_brain] = dki_fit(dwi(:,:,slice,:), grad, mask_whole);
[fa_brain] = dki_parameters(dt_brain);

% Select all numbers that are not NaN
select = ~isnan(fa_brain);
fa_values = fa_brain(select);

% Sort in ascending order
fa_sorted = sort(fa_values);

index10 = length(fa_sorted)*0.10;
index90 = length(fa_sorted)*0.90;

if mod(index10,1) == 0 %index=integer
    percentile10 = (fa_sorted(index10) + fa_sorted(index10 +
        1))/2;
else
    index10 = ceil(index10);
    percentile10 = fa_sorted(index10);
end

if mod(index90,1) == 0
    percentile90 = (fa_sorted(index90) + fa_sorted(index90 +
        1))/2;
else
    index90 = ceil(index90);
    percentile90 = fa_sorted(index90);
end

% 10% smallest and 90% largest FA values:
fa_10 = fa_sorted(1:index10); % Gray matter + CSF
fa_90 = fa_sorted(index90:end); % White matter

% Plot pixels
for i = 1:length(fa_10)
    voxels_10(i) = find(fa_brain == fa_10(i));
end

for i = 1:length(fa_90)
    voxels_90(i) = find(fa_brain == fa_90(i));
end

mask_10 = zeros(1, nx*ny);
mask_10(voxels_10) = fa_10;
mask_10 = reshape(mask_10, size(fa_brain));

```

```

mask_90 = zeros(1,nx*ny);
mask_90(voxels_90) = fa_90;
mask_90 = reshape(mask_90, size(fa_brain));

% Generate mask
white_matter_mask = boolean(mask_90);
gray_matter_mask = boolean(mask_10);

```

B.6 sim_exp1_DKI.m

Parameter estimation as a function of SNR and NSA for the DKI model

```

% Load data
dwi = niftiread();
bvecs = importdata();
bvals = importdata();
grad = [bvecs' bvals'];

% Generate signal and establish ground truth from selected voxel
for v = 1:size(S_voxel,1)
    S = S_voxel(v,:);
    S = reshape(S, [1 1 1 length(S)]);
    [~, dt_real, w] = dki_fit(S, grad); % full gradient set
    S_DKI(:,v) = w; % Estimated signal

    % Estimate parameters
    S = S_DKI(:,v);
    S = reshape(S, [1 1 1 length(S)]);
    [~, dt] = dki_fit(S, grad); % b < 3000
    [fa(v), md(v), rd(v), ad(v), ~, mk(v), rk(v), ak(v)] =
        dki_parameters(dt); % Ground truth
end

% Normalize signal
inds0 = find(bvals < 10);
S_DKI_norm = S_DKI./nanmean(S_DKI(inds0,:));

S = S_DKI_norm;
nvox=size(S,2);

n = 1000;
SNR = [10 15 20 25 30 40 50 60 80 100];
NSA = 1;

% MAIN LOOP
for v = 1:nvox % ~ 60 sec per voxel
    for k = 1:NSA
        for i = 1:n
            for j = 1:length(SNR)

```



```

        % Add rician noise
        N_real = randn(1, size(S,1))*(1/SNR(j));
        N_imag = randn(1, size(S,1))*(1/SNR(j));
        x = N_real + S(:,v);
        y = N_imag;
        S_noise_data(:, i, j, k) = sqrt(x.^2 + y.^2);
    end
end
end
end

% Average signals
S_noise_mean = mean(S_noise_data, 4);

% Calculate parameters from averaged signal
for i = 1:n
    for j = 1:length(SNR)
        S_noise = S_noise_mean(:, i, j);

        % Calculate tensor from signal with noise
        S_noise_reshaped = reshape(S_noise, [1, 1, 1,
            length(S_noise)]);
        [~, dt_app] = dki_fit(S_noise_reshaped, grad(:,:));

        % Parameter estimation
        [fa_data(i, j), md_data(i, j), rd_data(i, j), ad_data(i, j),
            ~, mk_data(i, j), rk_data(i, j), ak_data(i, j)] =
            dki_parameters(dt_app);

    end
end

% Remove outliers: Tukeys method
k = 1.5;
fa_data( fa_data < ( prctile(fa_data, 25) - k*iqr(fa_data)) ) =
    NaN;
fa_data( fa_data > ( prctile(fa_data, 75) + k*iqr(fa_data)) ) =
    NaN;
...

% Compute relative error
fa_rel = (fa_data - fa)./fa*100;
...

```

B.7 sim_exp1_DTI.m

Analytic DTI signal generation.

```
% Load data
```

```

dwi = niftiread();
bvecs = importdata();
bvals = importdata();
grad = [bvecs' bvals'];

inds1500 = find(bvals < 1500);
G = grad(inds1500,1:3);
b = grad(inds1500,4);
S0 = 1;

% Generate signal based on real tensor
%d_real = [1.6053, 0.1482, -0.0631, 0.1654, -0.0197,
           0.1202]*1e-3; % White matter
%d_real = [0.6940, 0.0090, -0.0606, 0.6893, 0.0854,
           0.7667]*1e-3; % GM
%d_real = [2.1055, 0.0064, 0.0443, 2.5200, -0.0087,
           2.3436]*1e-3; % CSF

for i = 1:length(G)
    H(i,1) = G(i,1)^2;
    H(i,2) = G(i,2)^2;
    H(i,3) = G(i,3)^2;
    H(i,4) = 2*G(i,1)*G(i,2);
    H(i,5) = 2*G(i,1)*G(i,3);
    H(i,6) = 2*G(i,2)*G(i,3);
end

Y = H*d_real';

% Calculate signal
S_DTI_2 = S0*exp(-b.*Y);

% Estimate D from S_DTI_2
Y = -(1./b).*log(S_DTI_2./S0);
H_ps = inv(H'*H)*H';
d = H_ps*Y;

% Parameter estimation
D = [d(1) d(4) d(5); d(4) d(2) d(6); d(5) d(6) d(3)]*1000;
[MD_2, FA_2, RD_2, AD_2] = dti_parameters(D);

```

B.8 sim_exp3_DKI.m

Parameter estimation as a function of number of gradient directions and b -values for the DKI model.

```

% Load data
dwi = niftiread();

```

```

bvecs = importdata();
bvals = importdata();
grad = [bvecs' bvals'];

% Create gradient sets
inds0 = find(bvals < 10);
inds500 = find(400 < bvals & bvals < 600);
inds1000 = find(800 < bvals & bvals < 1200);
inds2000 = find(1900 < bvals & bvals < 2100);
inds3000 = find(2900 < bvals & bvals < 3100);

% DKI gradient sets
set1 = [inds0 inds500 inds1000];
set2 = [inds0(1) inds500(1:2:end) inds3000(1:2:end)];
set3 = [inds0(1) inds500 inds3000];
set4 = [inds0 inds500 inds3000];
set5 = [inds0(1) inds500(1:2:end) inds1000(1:2:end)
        inds3000(1:2:end)];
set6 = [inds0(1) inds500 inds1000 inds3000];
set7 = [inds0 inds500 inds1000 inds3000];
set8 = [inds0 inds500 inds1000 inds2000 inds3000];

% ENSURE ISOTROPICALLY DISTRIBUTED GRADIENT DIRECTIONS
% Plot gradient directions
inds = inds1000(1:2:end)
figure, hold on,
for i=1:length(inds)
    plot3([-grad_roi(inds(i),1) grad_roi(inds(i),1)],
          [-grad_roi(inds(i),2) grad_roi(inds(i),2)],
          [-grad_roi(inds(i),3) grad_roi(inds(i),3)]);
    pause;
end

% Generate signal and establish ground truth
...

% Add noise
nvox=size(S_DKI_norm,2);
n = 1000;
SNR = 50;
NSA = 1;
set_nr = [{set1} {set2} {set3} {set4} {set5} {set6} {set7}
          {set8}];

% MAIN LOOP
for s = 1:length(set_nr)
    inds = set_nr{s};
    S = S_DKI_norm(inds);
    for v = 1:nvox % ~ 60 sec per voxel
        for k = 1:NSA
            for i = 1:n
                for j = 1:length(SNR)
                    %Add rician noise

```

```

N_real = randn(1, size(S,1))*(1/SNR(j));
N_imag = randn(1, size(S,1))*(1/SNR(j));
x = N_real + S(:,v);
y = N_imag;
S_noise = sqrt(x.^2 + y.^2);

% Calculate tensor from signal with noise
S_noise_reshaped = reshape(S_noise, [1, 1,
    1, length(S_noise)]);
[~, dt_app] = dki_fit(S_noise_reshaped,
    grad(inds, :));

[fa_n, md_n, rd_n, ad_n, ~, mk_n, rk_n,
    ak_n] = dki_parameters(dt_app);

fa_data(i, j, k, v) = fa_n;
md_data(i, j, k, v) = md_n;
rd_data(i, j, k, v) = rd_n;
ad_data(i, j, k, v) = ad_n;
mk_data(i, j, k, v) = mk_n;
rk_data(i, j, k, v) = rk_n;
ak_data(i, j, k, v) = ak_n;

fa_set{s} = fa_data;
md_set{s} = md_data;
rd_set{s} = rd_data;
ad_set{s} = ad_data;
mk_set{s} = mk_data;
rk_set{s} = rk_data;
ak_set{s} = ak_data;
    end
  end
end
end
end
end

```