CHALMERS

Developing a Method for Distortion Correction in High b-Value Diffusion-Weighted Magnetic Resonance Imaging

Master's thesis in Complex Adaptive Systems

HENRIK HANSSON

Department of Applied Physics
Division of Complex Adaptive Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2013
Master's thesis

MASTER'S THESIS IN COMPLEX ADAPTIVE SYSTEMS

Developing a Method for Distortion Correction in High b-Value Diffusion-Weighted Magnetic Resonance Imaging

HENRIK HANSSON

Department of Applied Physics
Division of Complex Adaptive Systems
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2013

Developing a Method for Distortion Correction in High b-Value Diffusion-Weighted Magnetic Resonance Imaging
HENRIK HANSSON

Master's thesis
ISSN 1652-8557
Department of Applied Physics
Division of Complex Adaptive Systems
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 1000

Cover:
Magnetic resonance image of the author's brain

Chalmers Reproservice
Gothenburg, Sweden 2013

Developing a Method for Distortion Correction in High b-Value Diffusion-Weighted Magnetic Resonance Imaging
Master's thesis in Complex Adaptive Systems
HENRIK HANSSON
Department of Applied Physics
Division of Complex Adaptive Systems
Chalmers University of Technology

Abstract

Diffusion-weighted magnetic resonance imaging (MRI) is a medical imaging technique that utilizes strong magnetic field and radio waves to measure the speed of diffusion of water in the human body. It has been a growing field in recent years, with new methods continually being developed to further the knowledge of the human body, and especially the brain. Most of these methods are used to calculate tissue characteristics based on information from multiple images showing the same area but showing the speed of diffusion in different directions. The calculations require thousands of images to hold information regarding the same tissue volume in the same place. These methods are improved by the use of more powerful magnets and larger magnetic field gradients in the imaging sequences. These stronger fields unfortunately also introduce distortions in the images, caused by undesired induced eddy currents and the heating of the main magnetic field, causing it to drift from its ideal strength. These distortions can make a part of the subject end up in different pixels for different images, while they should in fact be in the same place. Such a distortion makes it impossible to correctly calculate tissue characteristics based on multiple images. An example of this kind of characteristic is the fractional anisotropy that tells how much the speed of diffusion of water differs in different directions.

Different correction methods can be applied to counter distortions. Various methods are available to correct distortions for images that use clinical strength diffusion gradients ($b < 1000$ s mm^{-2}), but none are available for research sequences where the diffusion gradients can have b-values larger than 4000 s mm^{-2}. This thesis outlines and implements a completely new correction method for such high b-value sequences. The method is a post-processing method that can be applied without any special requirements on the imaging sequence that is used. It is designed to correct for the image distortions caused by eddy current and magnetic field drift in high b-value diffusion MRI.

A new post-processing method has been developed to correct for these distortions in high b-value diffusion imaging. The method requires multiple b-values to be captured for an image series and works by registering these images to each other, while the low b-value images can be registered to a non-diffusion weighted image, resulting in a global correction of the whole set of images. It uses local correlation for image comparison and particle swarm optimization to find the maximum of the sum of local correlations.

The new method has been tested on simulated data, on data from an imaging phantom and on real data from brain scans on volunteers. It is able to find most of the distortions, being 10-20% off from the true parameters on the simulated data in the presence of a large level of noise. On the phantom data and the in vivo data, it is shown to correct all of the distortions that are visually present in the images. It greatly improves the alignment of the images in the data sets when large distortions are present, while it does not negatively affect the images when no distortions are present. While the method improves images in its current configuration, further work is required to perfect the method. It does currently not handle patient movement, and optimization method that tries to find the best correction parameters could be improved to make sure that all distortions are removed.

Keywords: Magnetic resonance imaging, diffusion weighted MRI, eddy currents, high b-value

Acknowledgements

This thesis has been completed over a period of almost four years, taking me many and long hours to finally complete. I would not have been able to do this without the help of the many people to whom I owe my gratitude.

I would like to thank the whole of the MR Physics group at Lund University who made my time in Lund a pleasure with their comradeship and support. I owe the greatest thanks to my supervisors Jimmy Lätt and Markus Nilsson, who spent hours and days explaining the concepts of MRI to me, and who have taught me everything that I know of the subject. They were great fun to work along, and I deeply thank them for still showing interest in this work four years later. Special thanks also goes to Freddy Ståhlberg for introducing me to MRI and planting the seed for this work. I would also like to thank him for all the reminders he kept sending me to get it finished.

I am also grateful to all the staff at the MRI department at Skåne University Hospital in Lund, who welcomed me into their company and allowed me to spend many evenings using their MRI scanners.

Thanks to my family who has helped me by continually harassing me to get this done, and also helped me with proof reading this text.

Finally, my greatest thanks goes to my future wife and best friend, Jessica Jönsson, for supporting my work on this thesis and never growing tired of all the time that I unfortunately failed to spend with her.

This study was supported by the The Swedish Cancer Society, grant number CAN 2009/1076.

Nomenclature

α	Shearing parameter applied to an image
β	Scaling parameter applied to an image
γ	Both translation correction parameter and Gyromagnetic Ratio
δ	Gradient duration
Δ	Time between diffusion gradients
ADC	Apparent Diffusion Coefficient
$\mathbf{B}_0$	External static magnetic field
b-value	Strength of diffusion weighting
CC	Cross Correlation
CSF	Cerebrospinal fluid
CT	Computer Tomography
D	Diffusion coefficient
DICOM	Digital Imaging and Communications in Medicine
DT	Diffusion Tensor
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Imaging
EPI	Echo Planar Imaging
FA	Fractional Anisotropy
FID	Free Induction Decay
fMRI	Functional Magnetic Resonance Imaging
$\mathbf{G}$	Gradient vector
k-space	Frequency space for imaging
LC	Local Correlation
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
PCA	Pricipal Component Analysis
PGSE	Pulsed Gradient Spin Echo
PSO	Particle Swarm Optimization
RF	Radio Frequency
S	Signal level
SD	Standard Deviation
SNR	Signal to Noise Ratio
T	Tesla
T_d	Effective time of diffusion
TE	Time of echo
TR	Repetition time
T_1, T_2, T_2'	Time for signal relaxation of different types
Voxel	Image pixel generated from a volume

Contents

1 Introduction

Magnetic resonance imaging (MRI) is an imaging technique made possible by the physical phenomenon of magnetic spin resonance in atomic nuclei. This effect was discovered in the early 20th century and determines the way the magnetic spin of the nuclei precess in a static magnetic field. Manipulation of the nuclear spin can then be utilized to retrieve information about the materials containing the atoms. It was first used for spectroscopy under the name of nuclear magnetic resonance (NMR), and resulted in several Nobel prizes, such as the one awarded to Felix Bloch and Edward Mills Purcell in 1952. This technique was further developed to allow actual imaging in the 1970s, thus creating the branch of physics known as MRI, and introducing a completely new non-invasive imaging method for use in health care [1].

The basic idea of nuclear resonance is to excite magnetic spins using radio frequency pulses and then measure the time it takes for the spins to return to their equilibrium state. Different materials and tissues will have different relaxation times, making it possible to differ between them. This is what creates contrast between fat, bone and other tissue types in the human body.

The advancement of MRI during the last thirty years have made new functional techniques possible, which allows not only morphological imaging of tissues but also the measurement of properties that link directly to the functionality of the brain. The most well known such technique is functional MRI (fMRI), which measures small changes in blood flow in the brain as a response to different activities, thus connecting an activity, such as raising an arm, to a certain part of the brain.

This study concerns another such technique, diffusion-weighted imaging. Since the brain, as the rest of the human body, mainly consists of water, diffusion is a constantly present physical process. Measuring the speed of diffusion of this water can give much information regarding the properties of tissue in different parts of the brain.

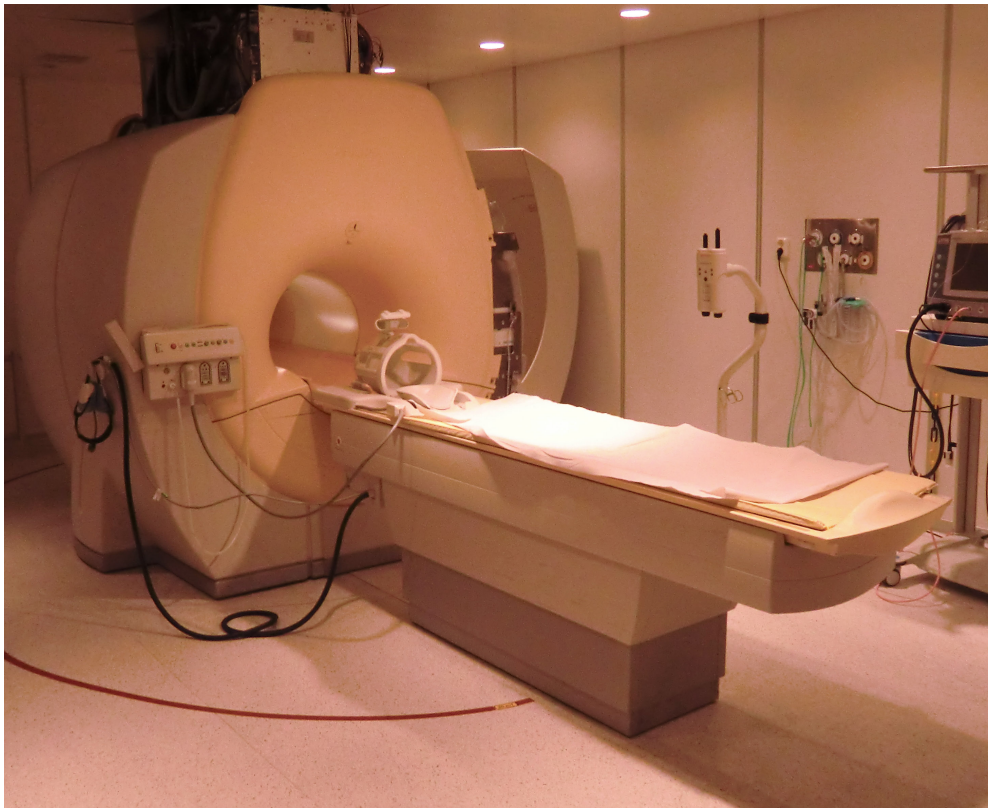


Figure 1.1: *The 3T Philips MRI scanner that was used to generate all the images used in this thesis. The scanner is fitted with the receiving head coil that was used for imaging of both the phantom and the volunteers.*

1.1 Diffusion

Diffusion of particles, called Brownian motion for larger particles, is the random movement of particles that occur at all time in all matter due to the thermal energy that creates constant motion and collisions of molecules. It is named after the botanist Robert Brown who first described it in a scientific paper in 1828, where he discusses the way small particles contained in plant pollen moves on a water surface [2]. The mathematics and physics behind this process were not known at the time and was first fully described almost a century later, when Albert Einstein published a paper on Brownian motion in 1905 [3]. In this paper he showed that the mathematics describing this movement leads to an equation controlling diffusion, known as Fick's law, that had previously been discovered by studies of the way concentrations even out in fluids and gases. Einstein showed that the spatial distribution of particle displacements due to diffusion can be calculated by a simple formula if the diffusion coefficient and the time of diffusion are known. The diffusion coefficient is measured when there is no concentration difference present in the subject, and is then called the self-diffusion coefficient. In water this coefficient at room temperature is 2.3 $\mu m^2/ms$ [4].

1.2 Usage of Diffusion Imaging

Measuring diffusion of water in the body using MRI was proposed soon after the first medical scanners were built in the late 1970s [1]. It was realized that the speed of diffusion could be measured in different tissues, giving another property to differentiate tissue types. The property that is found by measuring the diffusion in this way is the average rate of the diffusion, known as the apparent diffusion coefficient (ADC), in m^2/s. It is only the apparent coefficient, since there are other effects affecting the signal, such as the flow of blood in the brain and also the structures of the tissue that hinders diffusion. A real diffusion coefficient D is measured as free diffusion where the moving particles are not hindered by barriers during the diffusion time. The first practical usage of diffusion-weighted imaging (DWI) arrived in 1990, when Moseley et al. showed that diffusion imaging could be used to diagnose ischemic stroke in cats in the first few hours after onset, something that is not possible with neither computer tomography (CT) nor normal MRI [5].

The next evolution in diffusion MRI was to study the diffusional anisotropy in white matter in the brain. It was found that the ADC was considerably higher in the direction of the neural fibers than in the perpendicular directions. A method to calculate this anisotropy was presented by Basser et al. in 1994 [6]. They captured several diffusion images that each showed the speed of diffusion in a different direction, but for the same volume of the brain. These values were then combined to calculate a tensor for each voxel (an image pixel containing signal value from a volume), called the diffusion tensor (DT). This tensor makes it possible to describe the diffusion in the voxel as an ellipsoid. A spherical shape shows that the diffusion is similar in all directions, while anisotropy is indicated by an elongated ellipsoid. This can be used to make probabilistic calculations of the paths of white matter fiber bundles. An example of such an image is found in figure 1.3. Such images can be used to detect damage to white matter in patients.

Diffusion is also used to make measurements on a microscopic scale. A typical diffusion image shows the result of the diffusion that has occurred during about 50 ms. In this time, unhindered water molecules will travel a root mean square distance of 15 μm, in a single dimension. This value is far smaller than the typical volume that is contained within an image pixel, which is typically 1-10 mm^3. This means that diffusion measurement can potentially contain information that has far better resolution than the voxels of an ordinary MRI image. This has been used in more recent years, when diffusion imaging has been utilized to measure the size of very small compartments in a sample [7]. Another measurable diffusion parameter is the kurtosis, which indicates by how far the diffusion distribution deviates from a Gaussian curve. This parameter has the potential of offering further information regarding tissue [8].

1.3 Distortions and Corrections

As with any other imaging method, MRI is not free from distortions. The distortion that is most troublesome for diffusion imaging is the effect of eddy currents. Diffusion is measured in MRI by the application of strong magnetic field gradients that encode the position of molecules and lead to signal loss in places where the encoded spins have changed their position through diffusion. A real world system can never apply a magnetic field without ramping up to the desired level. Changing a magnetic field strength creates current through induction in the system, and these induced eddy currents will influence the following magnetic field gradients

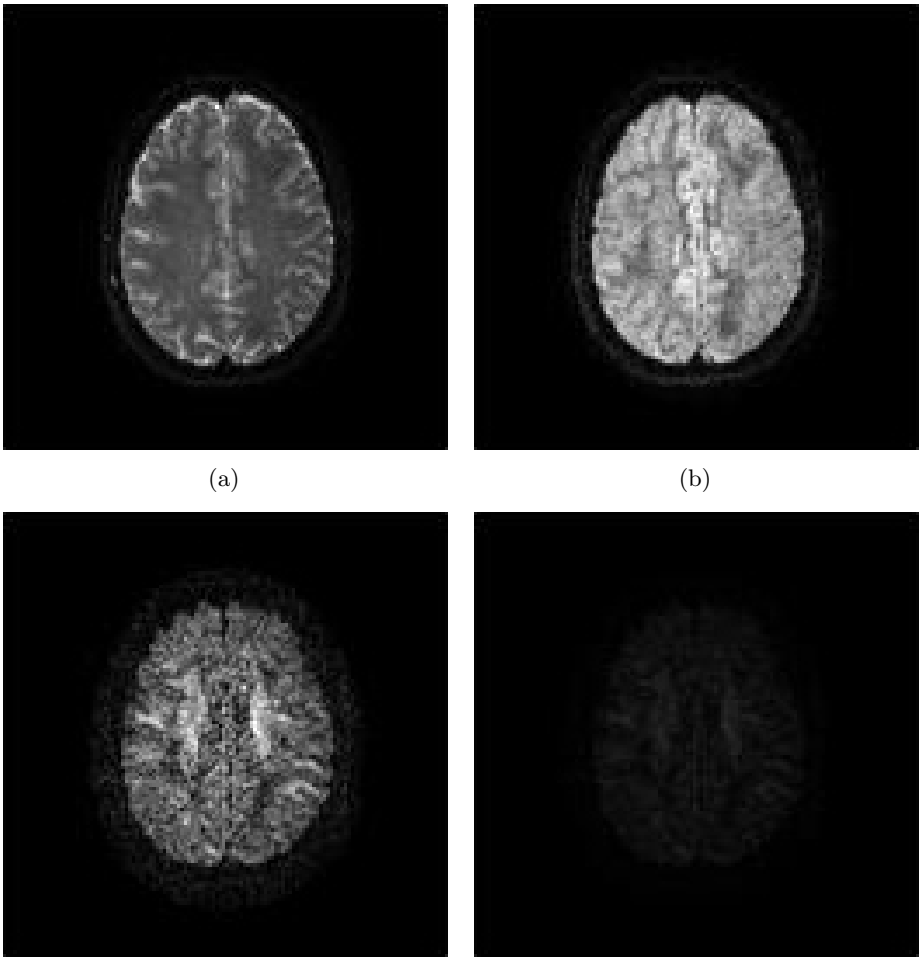


Figure 1.2: *MRI images from a diffusion set showing the same slice, but with different contrast scale. (a) T_2-weighted image without any diffusion gradient applied. (b) Diffusion image, b = 500. (c) Diffusion image, b = 3000. (d) The same image as (c), but shown using the same scale as in (a). This shows the loss of signal caused by the strong diffusion gradient.*

by their own associated magnetic fields. This leads to distortion during the image sampling and results in transformations in the final image. This kind of distortion is the central question in this thesis, and is discussed in detail in section 3.2. Figure 3.9 show how this distortion affects a diffusion image.

Another effect in MRI is that the patient can move during the imaging sequence. Creating a single image using MRI takes from a few milliseconds to several seconds depending on the type of sequence being used, and patient movement can affect the image in the longer sequences. Movement can be even more of a problem in diffusion, where images with different diffusion directions are combined to calculate the values of *ADC* and other derived diffusion-related parameters. These images can be taken many seconds apart, making it possible for movement to heavily affect the results, or even render the whole image set useless.

1.4 Purpose

This thesis is focused on the correction of eddy current distortions in diffusion images. The MRI physics group at the Department of Medical Radiation Physics at Lund University studies cutting edge diffusion methods that require the strongest possible magnetic field gradients to be used [9, 10]. These strong magnetic fields create eddy currents that could influence and potentially ruin studies of kurtosis and other parameters that are calculated using multiple images. Few research groups focus on diffusion in MRI with strong diffusion gradients, and there has been a lack of established methods to correct distortions in these images. The existing methods

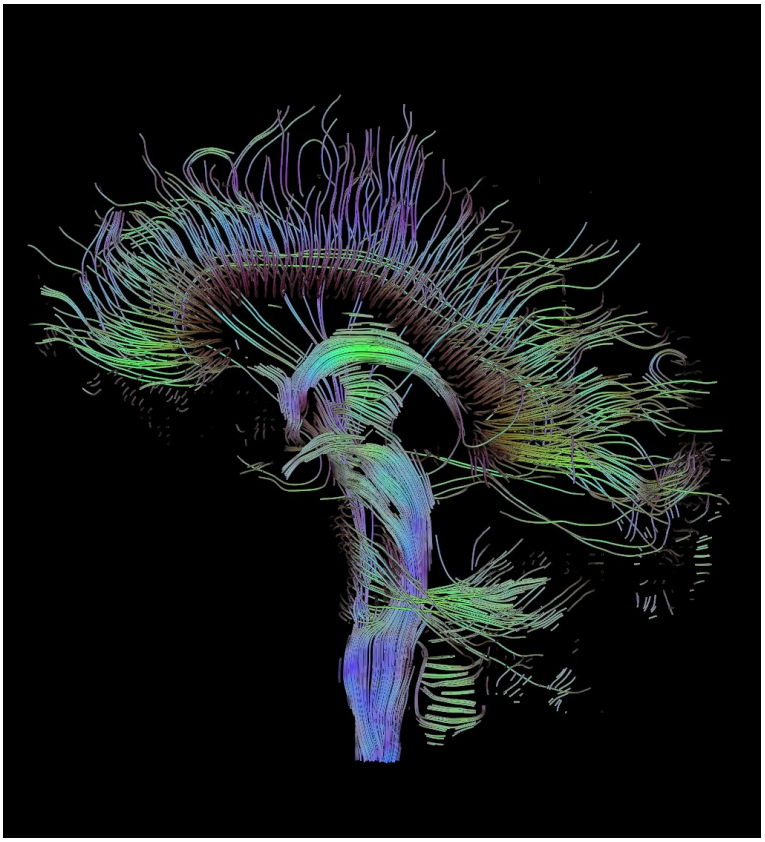


Figure 1.3: *Example of fiber tracking using diffusion imaging. The paths are the result of a probabilistic tracking using the calculated diffusion tensor of each voxel of a diffusion imaging set. Image courtesy of Thomas Schultz under the Creative Commons Attribution-Share Alike 3.0 Unported license.*

work well for ordinary diffusion images, but rest on principles that do not hold for these new sequences. To limit or eliminate the effect of distortions, a new method was needed.

The purpose of this work is to try to develop a method that makes it possible to correct for image distortions caused by eddy currents and magnetic field drift in diffusion weighted MRI images taken with strong diffusion gradients. The method should preferably not require any modification to the existing diffusion sequences, making it possible to use it for post correction of already captured images. Another requirement is that the method should have a solid base in the known physics of MRI and avoid the type of corrections that lack such a foundation, which are common in many methods. Finally, the method should possibly be combined with some kind of correction of patient movement, thus giving the group a complete suite of tools for distortion correction.

1.5 Goals

- To develop a model for correction of eddy current distortions in diffusion images.
- To use this model in a post-processing correction algorithm that can be applied to diffusion image sets.
- To evaluate the percentage of distortions that can be removed using the new method.

2 MRI Theory

This chapter aims to give a short explanation of the physics involved in generating MRI images. It demonstrates the principles of MRI physics and the most basic technique to generate images. The reader that is familiar with the concept and theory of MRI can move on to the next chapter since this chapter does not contain any theory that is specific to this thesis.

2.1 Creating a Signal

The basics of MRI begin with creating a signal that can be detected. This first part explains how this is possible to do by affecting the atomic spins in the sample using radio frequency pulses. The affected spins are excited to a state from which they return to equilibrium. The excited spins will precess at a high frequency in a strong magnetic field, which creates an induced electric current in coils surrounding the sample. The resulting voltage can be measured to generate a signal from the sample.

2.1.1 Spins and Magnetization

Magnetic resonance is based on the properties of the atomic spin. All atomic nuclei with a non-zero magnetic spin have a magnetization and can be utilized for resonance. This is the case with the hydrogen atom, which is found in abundance in all tissues in the human body, thus making it a perfect agent for medical resonance imaging. The spins of the hydrogen nuclei behave as randomly distributed microscopic magnetic vectors in the body, having a total magnetization of zero due to the large number of randomized spins.

To create a state of equilibrium, suitable for imaging, an external static magnetic field ($\mathbf{B}_0$) is applied along the z axis. This field invokes a peculiar quantum effect on the spins. It creates a torque on the magnetic moments, making them all precess along the z axis (figure 2.1a) with the same constant angular frequency, thus creating an observable angular momentum.

All these spins can now be treated using classical physics, where each spin acts as a precessing magnet with a random orientation. The external field will however affect these magnets slightly, giving them a slight tendency to align along the z axis. This tendency is enough to create the requirement of MR, which is a total magnetization vector behaving as one large predictably spinning magnet. Figure 2.1b illustrates how this combined magnetization $\mathbf{M}$ behaves. The size of this magnetization is proportional to the external magnetic field. A stronger field yields a larger net magnetization which is desirable, since this magnetization is what creates signal in MRI, as will be shown next. A stronger field results in a better signal to noise ratio, and is the reason why higher field strength is continually in demand in MRI [11].

2.1.2 Larmor Frequency

The frequency of the precession discussed above is called the Larmor frequency. It is commonly presented in the form of the Larmor equation:

$$\omega_0 = \gamma \mathbf{B}_0 \tag{2.1}$$

The nuclei dependent part γ is known as the gyromagnetic ratio and is measured in rad/(sT). In the case of hydrogen, the value is $\gamma/2\pi = 42.58$ MHz/T. Note that the definition in the Larmor equation gives the Larmor frequency as an angular frequency. It is more common to refer to the Larmor frequency in Hz instead. This value is found by dividing the value of ω_0 by 2π. Thus, the frequency required for resonance in an MRI scanner with a static field of 3 Tesla is $3 \cdot 42.58 = 127.5$MHz. This value is in the radio frequency (RF) range, meaning that MRI scanners can affect the magnetization using radio pulses [12].

The Larmor equation shows that spins will precess about the z axis with a very high frequency. To simplify later calculations, it helps to use a rotating frame of reference. Instead of x, y and z, we introduce a coordinate system of x', y' and z where the plane spanned by x' and y' rotates with angular frequency w_0 in relation to x and y. This creates a system where the main spinning magnetization is static instead of rotating. This system will be used in later sections.

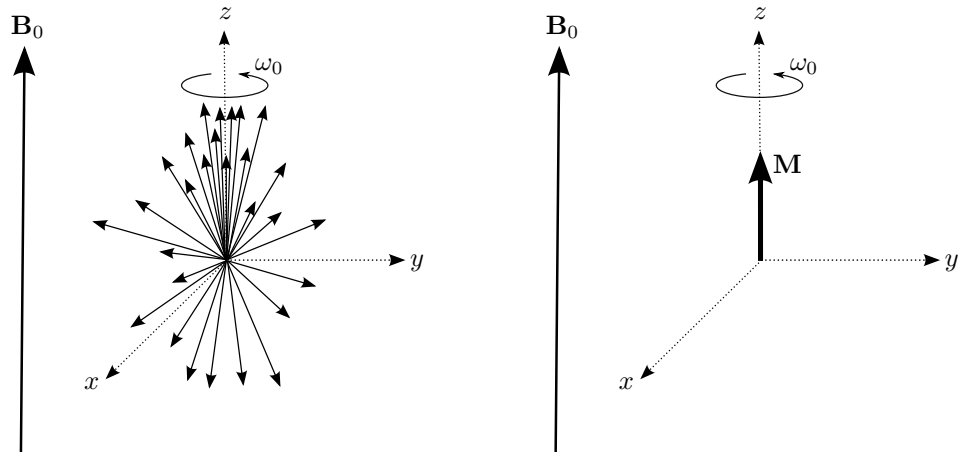


Figure 2.1: *(a) Spins in an external field $\boldsymbol{B}_0$ will precess around the z axis with the Larmor frequency ω_0. The spins will be randomly distributed, but with a slight tendency to align with the z axis due to the presence of the external field. (b) The sum of all these spins result in a total magnetization $\boldsymbol{M}$ along the z axis. The x and y projections of the spins sum to zero due to their completely random orientations.*

2.1.3 RF Pulses and the MR Signal

As seen above in figure 2.1b, the sum of all spins can be seen as single total magnetization vector. Due to the statistical properties of quantum mechanics, this allows us to ignore the individual spins and focus on this magnetization instead. This will behave according to classical mechanics, and can be affected by other magnetic fields. A second external field will apply a torque to the magnetization as given by

$$\frac{\partial \mathbf{M}}{\partial t} = \gamma(\mathbf{M} \times \mathbf{B}) \tag{2.2}$$

where γ is the gyromagnetic ratio, $\mathbf{M}$ is the magnetization and $\mathbf{B}$ is a new external field [13]. This is known as the Bloch equation and the effect is used to excite the magnetization, which is shown in figure 2.2.

To enable this effect, a magnetic field has to be applied that is perpendicular to both the total magnetization and the static field. This is accomplished by using an RF pulse that is circularly polarized with regard to the z axis. Since the magnetization vector will rotate with the Larmor frequency in the x-y plane, the RF pulse along the z axis must have the same frequency. This will create a magnetic field $B\cos\omega_0 t$ along the y axis, and hence a constant field B along the y' axis. Thus, according to equation (2.2), we can achieve any desired rotation of the magnetization vector by enabling an RF pulse a suitably long time. It is most common to rotate the magnetization vector 90 degrees to excite the system. The degrees of rotation is known as the flip angle.

With the system excited, the magnetization vector will be precessing around the z axis at the angle determined by the flip angle. With the magnetic vector constantly changing in the x-y plane, induction will occur according to Faraday's law of induction

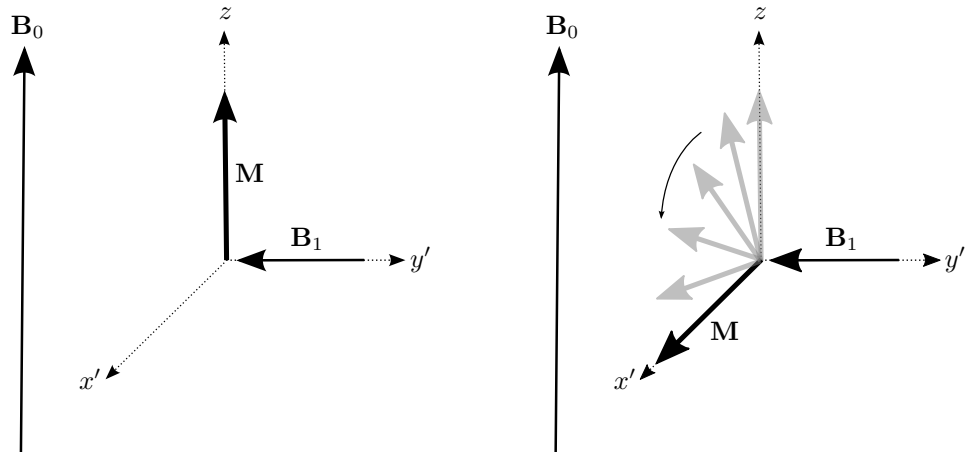
$$\mathcal{E} = -\frac{d\Phi}{dt} \tag{2.3}$$

where $\mathcal{E}$ is the electromotive force induced in the coil in volts and Φ is the magnetic flux through the coil [14]. By surrounding the spinning magnetization with coils, a current will be induced and will have the frequency ω_0 of the rotating magnetization. In this way, the most basic MR signal is created, known as the free induction.

The signal induced in this way consists of contributions from each spin in the whole sample. At each moment t, the total signal can be calculated as

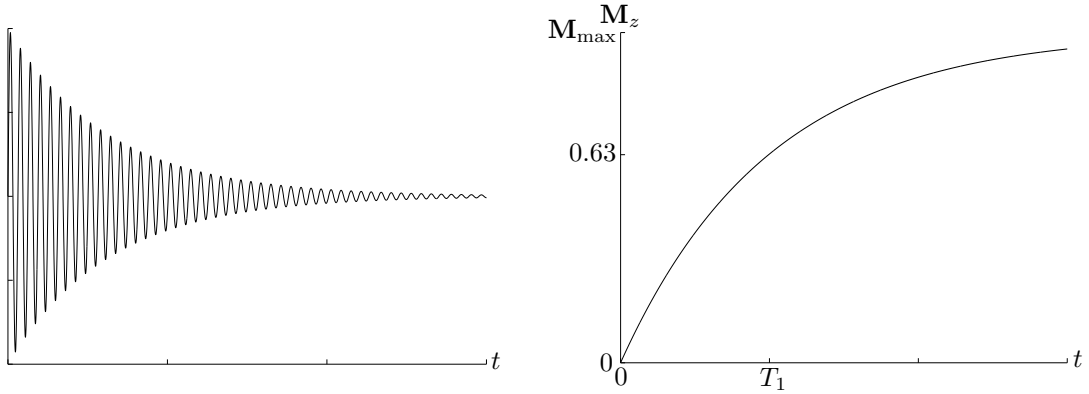
$$s(t) = \int_{\mathbf{r}} \rho(\mathbf{r}) e^{i\phi(\mathbf{r},t)}\, d\mathbf{r} \tag{2.4}$$

where $\rho(\mathbf{r})$ is the spin density in each point, and $\phi(\mathbf{r}, t)$ is the phase of the spinning magnetization. In the most simple case, where ϕ is independent of the spatial location and the sample has a uniform spin density, the result is a pure complex sinusoidal signal.



(a) (b)

Figure 2.2: *(a) A second magnetic field $\boldsymbol{B}_1$ is applied along the y' axis by an RF pulse. This will create a torque on the magnetization $\boldsymbol{M}$, leading to the situation in (b) after a time t. The magnetization will keep rotating around the y' axis as long as the $\boldsymbol{B}_1$ field is applied. The system will return to the equilibrium state when $\boldsymbol{B}_1$ is switched off.*



(a) *Free induction decay* (b) T_1 *relaxation*

Figure 2.3: *(a) The free induction decay current that is induced during relaxation. It is a sinusoidal signal with frequency ω_0. Due to dephasing spins, the frequency spectrum is slightly widened as the signal decays. (b) The recovery of magnetization due to T_1 relaxation. After T_1 seconds, it has grown to 63%, or -3 dB from the original signal. After $3T_1$ seconds, the signal is back to 95% of its original strength.*

2.1.4 Relaxation

The free induction signal induced by the precessing magnetization vector will not last long after the RF pulse has been shut off. Without its effect the system will revert to the equilibrium state, a process known as relaxation. During a free relaxation, the signal induced in the receiving coils is known as the free induction decay (FID). An example of such a signal is shown in figure 2.3a. The signal is lost over time through two different relaxation processes.

The signal drops as the original magnetization along the z axis is recovered, which occurs when there is no longer a force present that pushes the magnetization away from its initial state. This process is called spin-lattice relaxation, and is also commonly referred to as T_1 relaxation. The latter name comes from the definition of the time T_1, which is the time it takes for the magnetization to recover 63% of its initial signal strength along the z axis. The recovery of magnetization is described by the following equation:

$$\mathbf{M}_z(t) = \mathbf{M}_z(0)(1 - e^{t/T_1}) \tag{2.5}$$

where $\mathbf{M}_z$ is the projection of the magnetization vector on the z axis. This process is illustrated in figure 2.3b. T_1 will differ depending on the tissue or material where the spins are located. It varies between approximately half a second to several seconds in different human tissue.

Relaxation in the x-y plane occurs when the precessing magnetization is dephased. When the magnetization is flipped into the plane, the spins that make up the total magnetization will initially all precess with the Larmor frequency ω_0. However, as time pass they will experience slight differences in the magnetic fields, making some spins increase their frequency of precession, while other decreases theirs. This process will make the total magnetization in the plane decrease (figure 2.4).

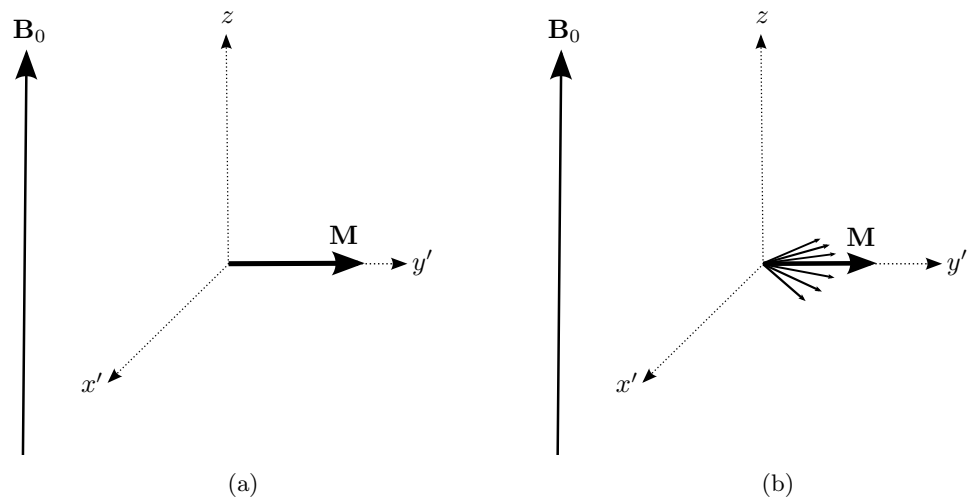


Figure 2.4: *Illustration of T_2 and T_2' relaxation. In (a), the magnetization has just been flipped to the x'-y' plane and all spins are precessing with the Larmor frequency. (b) shows the effect of T_2 and T_2' relaxation. Some spins have increased their precession frequency while other have decreased their frequency. The total magnetization projected onto the y' axis will then be smaller than in (a).*

The decrease of signal in the plane is caused by two different processes that have the same effect on the spins. Each individual spin will affect the magnetic fields surrounding its neighboring spins, making it differ slightly from $\mathbf{B}_0$. This will make the different spins have different precession frequency as described above, and is called T_2 relaxation. The other process that work in the same way as the T_2 effect is the much stronger T_2' relaxation. It is the effect of local fluctuations in the magnetic field and it quickly dephases the magnetization. Fortunately, this effect can be mostly removed by the use of spin echoes, which will be discussed later. The combined effect of T_2 and T_2' is called T_2^* and can be calculated as

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \tag{2.6}$$

The loss of magnetization in the x'-y' plane due to the dephasing spins is calculated by

$$\mathbf{M}_{x'y'}(t) = \mathbf{M}_{x'y'}(0)e^{-t/T_2^*} \tag{2.7}$$

The relaxation process in the plane is much faster than T_1, and T_2 will thus always be smaller than T_1. Typical values for T_2 are 100 ms in gray matter, 80 ms in white matter and 50 ms in muscle, for a 1.5 T scanner. This can be compared to T_1, where the relaxation times for the same tissue types are 950 ms for gray matter, 600 ms in white matter and 900 ms in muscle [12]. Note that the relaxation time for muscle is only half that of gray matter for T_2, but about the same as gray matter for T_1. This kind of difference makes it possible to get different contrast in images depending on which kind of signal is sampled.

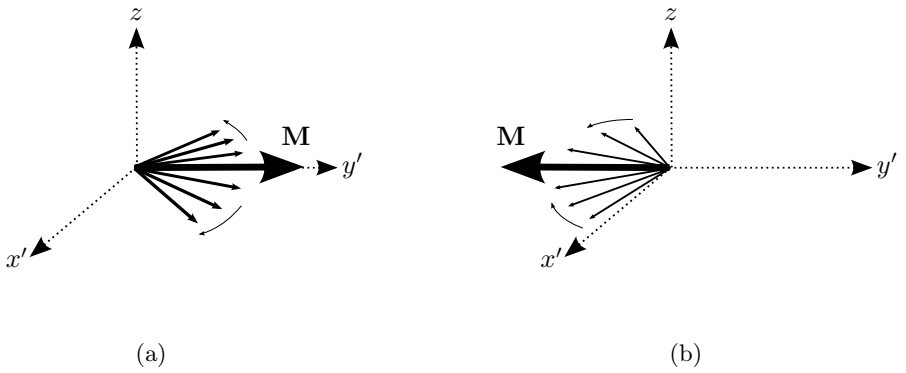


Figure 2.5: *Principle of a spin echo. In (a), the spins are dephasing due to the T_2' effect. (b) shows the situation after a 180 degree RF pulse has been applied. Spins that were previously ahead of the main magnetization are now behind, and vice versa. Since each spin will still change phase in the same way as compared to $\boldsymbol{M}$, they will now refocus and cause a spin echo.*

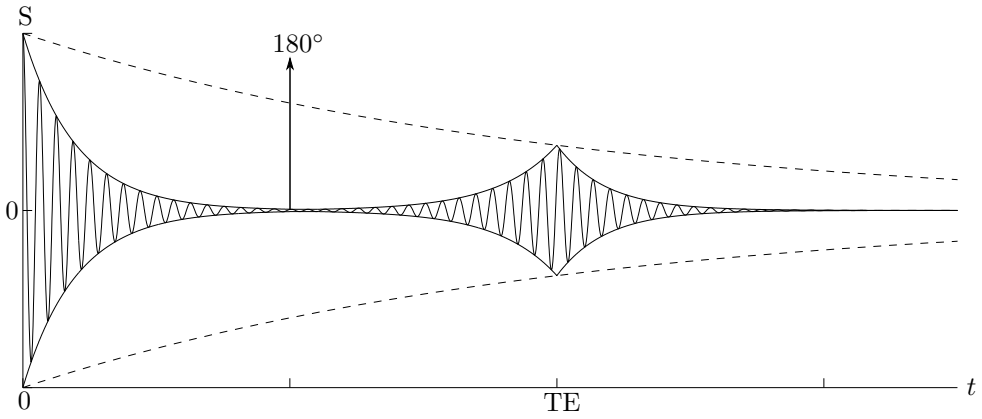


Figure 2.6: *Induced signal as a function of time for a spin echo. The dotted line represents the signal decrease that would result from T_2 relaxation only. The solid line shows the signal strength of the combined relaxation. The signal is quickly lost from T_2' relaxation, but this signal loss is recovered by the use of a 180 degree RF pulse, making a spin echo appear at TE, after twice the time from the application of the original exciting RF pulse at $t = 0$. The arrow is a common way to represent RF pulses in MRI literature.*

2.1.5 Spin Echoes

As mentioned before, the very fast process of T_2' relaxation can be countered by usage of a so called spin echo. T_2 and T_2' has the same effect and are both caused by the local magnetic field differing slightly from $\mathbf{B}_0$. T_2 relaxation is irreversible since it is caused by the magnetic fields of nearby spins, whose nature is unpredictable. The local fluctuations of the $\mathbf{B}_0$ field that cause T_2' relaxation is instead easily predictable, since it is due to fixed factors such as the distribution of different tissue in the sample. A spin in a certain location will continuously either fall behind or be ahead of the main magnetization, as illustrated in figure 2.4b. When an RF pulse is applied to flip all spins by 180 degrees, the situation is reversed. All spins that were behind the main magnetization are now ahead of it, and all spins that were ahead of the magnetization are all behind it

instead (see figure 2.5). Thus, if the signal is sampled exactly when the spins refocus, it will be void of any loss from the T_2' effect, and the only relaxation losses will be from T_2 [15]. The effect on the signal from a spin echo is illustrated in figure 2.6.

2.2 Imaging

The MR signal discussed so far would enable spectroscopy of a sample. To create actual images, more elaborate signal sampling is required. This second part of the chapter demonstrates how it is possible to find how the spin density varies within the subject, and how this variation is turned into an image. The key to this is the usage of magnetic field gradients to first enable imaging of a specific desired area, and then make it possible to encode the spatial positions in the final sampled signal.

2.2.1 Slice Selection

So far the signal discussed has been the free induction decay signal where a single frequency is induced in the receiving coil. The initial RF pulse has excited the whole sample, meaning that the collected signal is an average over all spins. This is not very useful for actual imaging, where spatial information is required to generate an image. Since the signal in MRI is an induced frequency, connecting frequency to spatial position is the natural way to create a positional difference. This can be achieved by using the fact that the resonance frequency is proportional to the external magnetic field, as established in section 2.1.2. By making sure that each location experiences a unique external magnetic field, a correlation between position and signal can be created. This is done in two different ways: by selecting a slice for imaging and by frequency encoding that slice.

Slice selection is made in order to make it possible for the signal to only include the thin slice of the sample that should end up in the image. This is accomplished by making sure that only the spins in the desired slice are excited when the RF pulse is applied to flip the spins. To make this possible, a magnetic field gradient is applied along the z axis, which makes each position along the axis experience different static magnetic fields, hence having different resonance frequency. When the RF pulse is applied, it can be adapted to only affect the resonance frequencies that are present in the slice that is to be imaged. Since we want to excite a slice, we need to limit the excitation to the thin range of frequencies that correspond to that slice, i.e. to a rectangular function in the frequency domain. This function corresponds to a sinc function in the time domain, meaning that this is the shape that an RF pulse should have in order to limit the excitation. Previously, excitation has been shown using an RF pulse in the shape of a sine function of Larmor frequency. To end up with a rectangular function in the frequency domain holding the needed frequencies and excite the desired slice, the sine function is multiplied with the sinc function to get the following Fourier transform:

$$\mathcal{F}\big(\sin(f_0 t)\cdot \mathrm{sinc}(at)\big) = \begin{cases} 0, & \omega < f_0 - a \\ 1, & f_0 - a < \omega < f_0 + a \\ 0, & \omega > f_0 + a \end{cases} \tag{2.8}$$

where $2a$ is the width of the selected slice, centered around the frequency f_0 [16]. When a pulse like this is used, only the selected slice will generate any signal, as illustrated in figure 2.7.

2.2.2 Frequency Encoding and the Fourier Transform

With slice selection in use, only the excited slice of the sample in the scanner will contribute to the generated signal. It does however not create any spatial awareness in the generated data. To do this, magnetic field gradients are put to use again, in a process known as frequency encoding.

The principle of frequency encoding is the same as the basis for slice selection, namely to make different parts of the sample experience different magnetic fields in order to make them unique. When parts of a sample experiences different external magnetic fields, the resonance frequencies are affected as shown before. The resonance frequency in the presence of a gradient along the x axis will be

$$\omega(x,t) = \omega_0 + \omega_G(x,t) \tag{2.9}$$

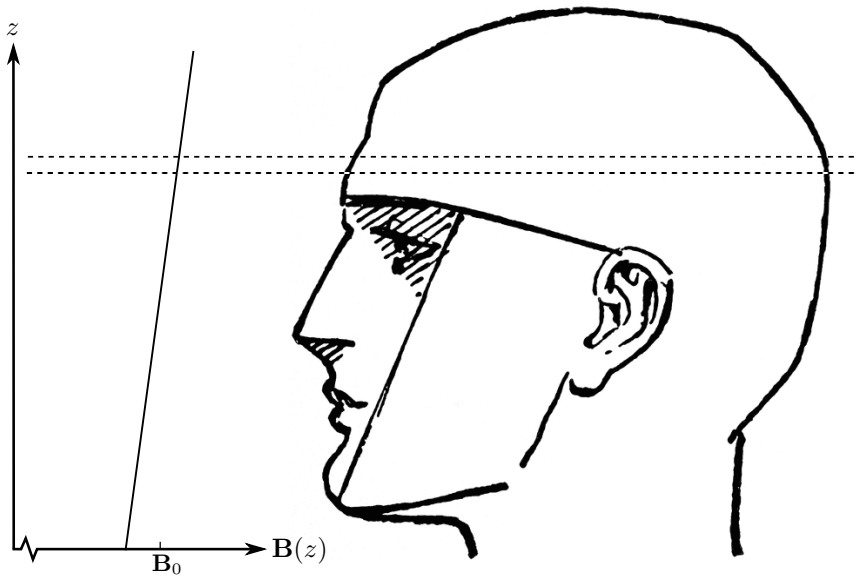


Figure 2.7: *A single slice (dotted lines) can be selected for imaging by applying a gradient on the static external field $\boldsymbol{B}_0$. By having the resonance frequency depend on z, it is possible to make only the desired slice to be excited by the following RF pulse.*

where ω_G is the deviation caused by the gradient $G(x)$. This difference in resonance frequency carries over to a difference in phase for the spinning magnetization along x. The accumulated phase due to the gradient for each position will be

$$\phi_G(x,t) = -\int_0^t \omega_G(x,t')\,dt' = -\gamma x \int_0^t G(t')\,dt' \tag{2.10}$$

where the second step assumes that the gradient is linear along x, giving $\omega_G(x,t) = \gamma x G(t)$. This difference in phase along the affected axis is illustrated in figure 2.8.

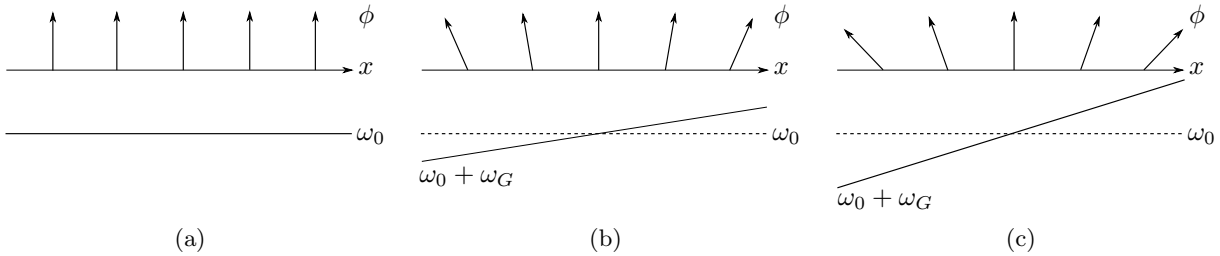


Figure 2.8: *The effect of a linear gradient on the accumulated phase difference after the gradient has been on for some time t. In (a), no gradient is present and all spins have the same resonance frequency ω_0 and the same phase. (b) shows the application of a gradient. The spins that experience a negative magnetic field have a positive phase difference against the spins in the center that still have the nominal resonance frequency. In (c), an even larger gradient creates a situation with larger phase differences. These differences in phase translates to a larger span of frequencies ending up in the induced signal.*

We can now imagine that we have a one dimensional object that is to be imaged, with the object placed along the x axis. The signal equation from (2.4) can now be written as

$$s(t) = \int \rho(x) e^{i\phi_G(x,t)}\,dx = \int \rho(x) e^{-ix\gamma \int_0^t G(t')dt'}\,dx \tag{2.11}$$

where only the phase difference is considered, since the frequency induced by the Larmor frequency is constant for all spins, and will be removed in filters before the signal is considered. If we use the convenient variable

replacement

$$k(t) = \gamma \int_0^t G(t')\,dt' \tag{2.12}$$

the expression in (2.11) is turned into

$$s(k) = \int \rho(x) e^{-ikx}\,dx \tag{2.13}$$

which is analogous to the Fourier transform as defined in [16]:

$$\hat{f}(\xi) = \int f(x) e^{-i\xi x}\,dx \tag{2.14}$$

with k in place of ξ. Thus, the sampled signal is in fact the Fourier transform of the spin density $\rho(x)$, which is what we want to find. This value can now be found by simply performing an inverse Fourier transform on the measured signal [12].

This can be demonstrated using a simple example where the one dimensional object from before is replaced by two single spots of non-zero spin density at $-x_0$ and x_0. A linear gradient Gx is applied along the x axis from time $0 < t < t_1$, while the signal is sampled during the same time interval. The spin at x_0 will now have a phase $\phi_G(x_0, t) = -\gamma G x t$ and the spin at $-x_0$ has instead arrived ahead by the same phase so that $\phi_G(-x_0, t) = \gamma G x t$. The signal that is sampled will be:

$$s(t) = s_0 e^{-i\gamma G x_0 t} + s_0 e^{i\gamma G x_0 t} = 2s_0 \cos(\gamma G t x_0) \tag{2.15}$$

or transformed to the k variable:

$$s(k) = 2s_0 \cos(k x_0) \tag{2.16}$$

with s_0 as the basic amplitude of each spot. From here, the inverse Fourier transform should result in the original spin density. k can be considered to be valid for negative values with an inverted sign on the gradient. This gives

$$\begin{aligned} \rho(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} 2s_0 \cos(k x_0) e^{ikx}\,dk \\ &= \frac{s_0}{2\pi} \int_{-\infty}^{\infty} e^{ik(x+x_0)} + e^{ik(x-x_0)}\,dk \\ &= s_0\big(\delta(z+z_0) + \delta(z-z_0)\big) \end{aligned} \tag{2.17}$$

which is two spikes at positions x_0 and $-x_0$, exactly as expected.

2.2.3 k-Space Sampling and Generating an Image

In the previous section, it was shown that the spin density can be recovered by performing an inverse Fourier transform on the sampled signal where the time and gradient dependent variable k from (2.12) is used. The key to being able to find the spin density in the earlier example was to sample the signal at times when k took on different values. In the final transformation, it was assumed that we had values for all values of k from negative to positive infinity. This is of never true in a real situation, but the principle still holds, and the finite number of samples available makes it possible to perform a discrete inverse Fourier transform to find the spin density. This way of using k is called k-space sampling, and it was first described in two separate articles in 1983, by Stig Ljunggren and Donald B. Twieg [17, 18].

To determine the spin density, the signal is sampled at as many different k values as possible. A closer examination of (2.12) shows that it is possible to reach virtually any value of k by applying negative or positive gradients for a suitable amount of time. This is used to create a grid of values of $s(k)$ that is used for the reverse transform. The only limitation is that the signal will experience relaxation and lose strength during the sampling. It is therefore important to sample as many points as possible close to the peak of the spin echo.

k-space is not limited to one axis, since the same principle that was demonstrated for the x axis is equally valid for further dimensions. It is most common to use gradients along the x and y axis, to create two dimensional images, but it is also possible to image in all three dimensions at once. In the two dimensional case, points (k_x, k_y) are sampled, and a two dimensional inverse Fourier transform is then performed to generate the actual image. An example of the way k-space is traversed and sampled is found in figure 2.9, where as much as

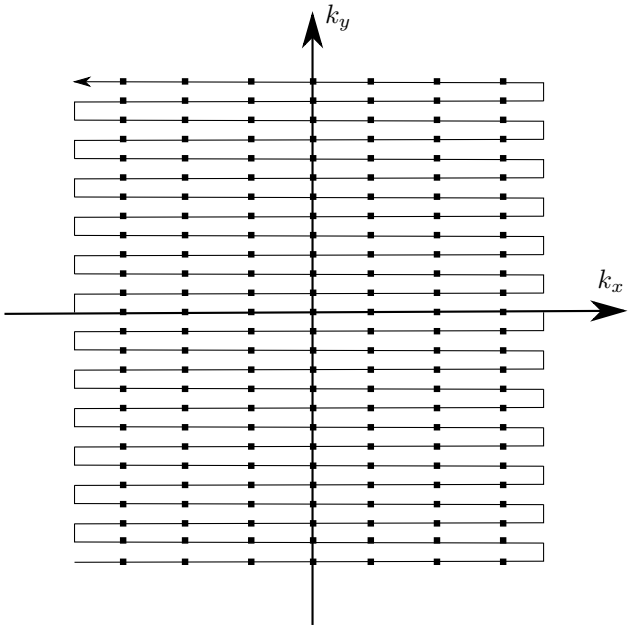


Figure 2.9: *The sampling of k-space during a full image slice readout. Each black dot represents a point where the signal is sampled. The path through k-space begins in the bottom left corner and sweeps back and forth along k_x, with small gradient blips creating the movement upwards along k_y. It is very important to sample the center of k-space exactly at the maximum of the spin echo, since this sample point holds information regarding the zero frequency, which holds the general signal level. In reality, the number of sampling points and lines in k-space are much higher than illustrated here. In general, between 64 and 256 lines are sampled.*

possible of k-space is scanned starting in the bottom left corner. Positive and negative gradients along x are switched on to move to the right and left in the figure, while positive gradients applied for a short time in the y direction creates the steps upwards. The x direction is generally known as the read direction, since the gradient in this direction is switched on during the actual sampling while the y direction is often called the phase direction. In the figure, the whole of k-space is sampled during a single excitation. This method is called Echo Planar Imaging (EPI), and is most common in the kind of images discussed in this thesis. Figure 2.10 gives an example of a gradient sequence that gives rise to this kind of sampling. It is also possible to excite the same slice several times and sample just one line in k-space during each excitation (i.e. keeping k_y unique for each excitation). This requires a longer scanning time, but can result in a better signal to noise ratio (SNR).

It should be noted that it is possible to sample k values in any desired order, and not just the one shown in figure 2.9. This order is commonly used since it makes the following Fourier transform quite simple, but many other patterns are possible. One such pattern is to sample k-space in a spiral [19]. Another is to use a sequence where the same slice is imaged several time, with the center of k-space sampled during each excitation, while different sections of the outer reaches are sampled each time [20]. This has the advantage of making it possible to get better SNR, while at the same time making it possible to adjust for patient movement that might otherwise affect the image quality.

The discussion here has mainly considered two dimensional images as being placed in the x-y plane. It is equally possible to excite a slice in an arbitrary plane of the sample. In this case, all the gradients described will simply have to be adjusted by an ordinary vector transformation.

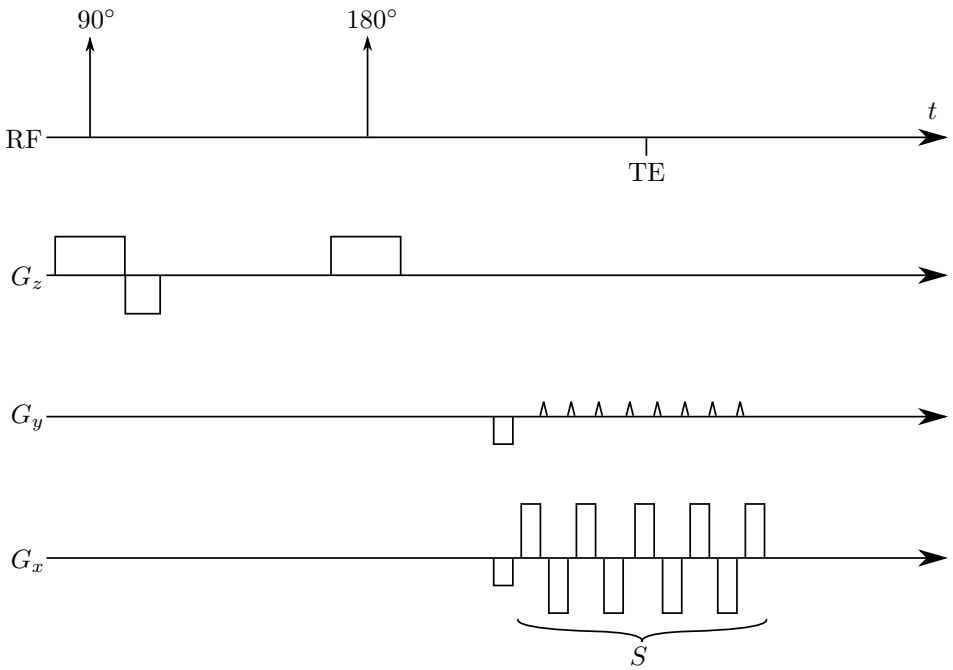


Figure 2.10: *Example of a full sequence that give the k-space pattern showed in figure 2.9. It shows the RF components and the different gradients played out on the three different axes. Only nine sweeps along the x axis is shown here for simplicity. The signal is sampled during the period indicated by S, and will result in a two dimensional image in the x-y plane. Note that the centre of k-space will be sampled at the echo time TE. The negative gradient shown for G_z is required to refocus the spins. This refocusing is not required after the 180 degree pulse, since the gradient is self refocusing for this flip angle. The negative gradients along x and y before the sampling begins are required to make the k-space path during sampling start in the lower left corner as in figure 2.9. The larger gradients along the x axis move the path back and forth, while the smaller triangular blips along the y axis moves the k-space position upwards along k_y. The actual sampling is only done when the x gradients are active.*

3 Diffusion Theory and Distortion Correction

This chapter explains the physics of diffusion and how it makes the field of diffusion MRI possible. The specifics of diffusion sequences and their sensitivity to distortions are introduced, followed by the concept of eddy current distortion. Finally, it summarizes the previous scientific efforts concerning eddy current compensation.

3.1 Diffusion Theory

As discussed in the introduction, diffusion measurements is an important tool in MRI. Its use was first shown in diagnosing acute ischemic stroke, but its field of use has since grown to many other applications. In the last two decades, several methods have been developed that use diffusion to probe into microscopic properties of tissue that are otherwise not possible to image with the resolution available using MRI scanners. These methods have in common that they measure diffusion in several different directions, and then combine these images into a value or vector for each voxel. This section shows how such images are created, and why they are susceptible to certain image distortions, creating the requirements for such correction methods as the one outlined in this thesis.

3.1.1 The Diffusion Equation

While Albert Einstein explained the physics and mathematics underlying diffusion [3], the macroscopic effect was already well known at that time. Diffusion is a process where molecules randomly change position. It is most noticeable in fluids and gases, but is also present in solids, especially in metals [21]. The observable macroscopic effect of diffusion is that differences in concentration disappear given time. That is, molecules move from an area with a higher concentration to areas where they have less presence until the same concentration is present everywhere. This effect was discovered in the first half of the 19th century, and was further described by Adolf Fick in 1855 [22]. Fick put salt in one end of a horizontal tube filled with water and measured how the salt spread in the tube. This led to the formulation of the diffusion equation:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{3.1}$$

where C is the concentration of the diffusing substrate and D is the diffusion coefficient in m^2/s. This coefficient is dependent on both the diffusing particles and the viscosity of the medium they diffuse in. Is is also dependent on the squared velocity of the particles, which in turn depend on the temperature. It is thus quite a complex coefficient. The relationship can easily be extended to three dimensions, which turn it into:

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} + \frac{\partial^2 C}{\partial z^2}\right) = D\nabla^2 C \tag{3.2}$$

This is in fact identical to the heat equation (with C replaced by temperature) that governs how heat spreads in a material. That is not very surprising giving the similar nature of heat transfer and diffusion.

The equations above describes the speed of diffusion. In diffusion MRI, it is the distance traveled by the diffusing particles that is relevant to the measurements. It is useful to consider the probability distribution of distance moved by the diffusing particles after a certain time. This distribution starts out as a delta function at $t = 0$ and then turns into a Gaussian distribution as time goes by, assuming that the diffusion is unhindered in all directions. The mean displacement of a freely diffusing particle can be calculated as

$$\lambda = \sqrt{2n_d D T_d} \tag{3.3}$$

where T_d is the diffusion time and n_d is the number of dimensions in which diffusion is possible [3].

The version of the diffusion equation given above in (3.2) works when the diffusion is unhindered in all directions. This is not the case in the human body, where tissue structures can work as barricades for the molecules. Since these barriers make the diffusion dependent on direction, it has to be described using a tensor with nine terms. These make up the diffusion tensor:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{3.4}$$

With the tensor, (3.2) becomes a summation over the nine different coefficients so that

$$\frac{\partial C}{\partial t} = \sum_{i,j} D_{ij} \frac{\partial^2 C}{\partial i \, \partial j} \tag{3.5}$$

3.1.2 Diffusion Sensitive Imaging

The idea of diffusion MRI is to make the image sensitive to the rate of diffusion of the spins contained in each voxel during the imaging. Any movement of molecules containing excited spins would be to lower the signal value in the voxel. This turns out to be quite simple to achieve using magnetic field gradients that cause phase dispersion in a similar way to the T_2 effect. In each diffusion image sequence, two identical gradients are placed on each side of the 180 degree RF pulse present in the normal imaging sequence from figure 2.10. The resulting sequence is found in figure 3.1. This basic diffusion sequence was first created by O. E. Stejskal and J. E. Tanner in 1965 and was first used for spectroscopy. It is called a pulsed gradient spin echo (PGSE) [23].

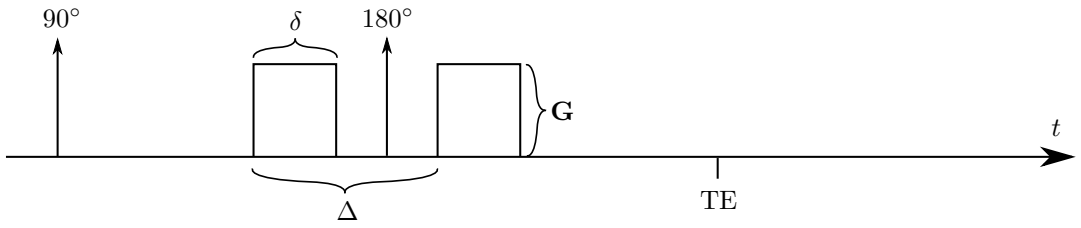


Figure 3.1: *Composition of a diffusion sequence of the pulsed gradient spin echo type, showing gradients and the common denotion of different parts of the sequence. δ is the duration of the gradients, while the gradient strength is marked $\boldsymbol{G}$. The separation of the gradients, Δ, is the time during which the movement of the particles affects the measurement. The size of the gradients shown here compared to the length of the sequence is for illustrative purposes and is not representative for the real world values of δ, Δ and $\boldsymbol{G}$.*

The effect of the PGSE sequence comes from the two gradients surrounding the 180 degree pulse. The first gradient will change the phase of each particle's magnetization a certain amount. The second gradient will have exactly the opposite effect due to the 180 degree pulse, meaning that the phase will change in the other direction. A particle that is in the same position during both gradients will thus have its magnetization left completely unaffected by the two gradients. Any movement during or between the gradients will however make the them have different effect on the moving spin, leaving it with a net phase change. The effect of these diffusion sensitizing gradients on particles moving in different ways is illustrated in figures 3.2 – 3.5. An important point to notice is that each diffusion measurement is only sensitive to motion along the direction of the applied gradients, as shown in figure 3.3. This makes it possible to detect differences in diffusion in different directions, as discussed regarding DTI in the previous section.

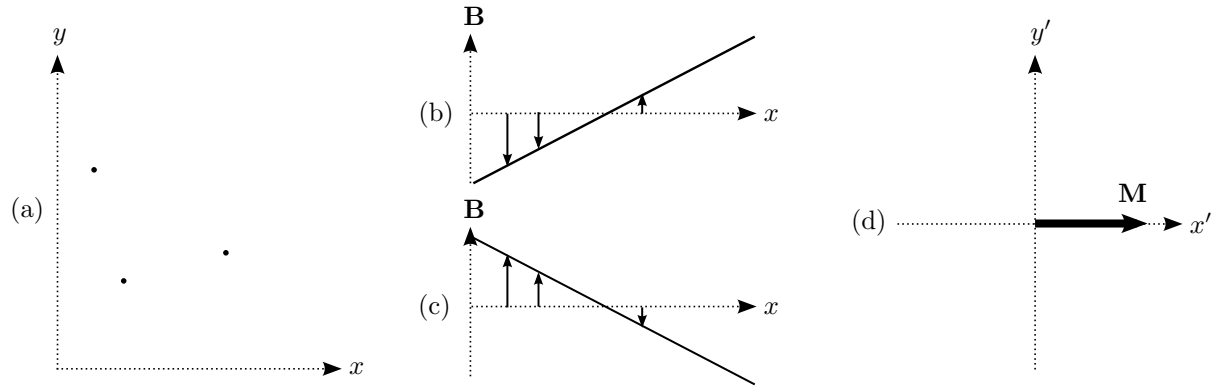


Figure 3.2: *The effect of diffusion sensitizing gradients on stationary particles. When the particles are still as in (a), the gradient effect on them during the first diffusion block (b) is exactly the same as during the second block (c). The two diffusion gradients will cancel each other, leaving the total magnetization unaffected (d).*

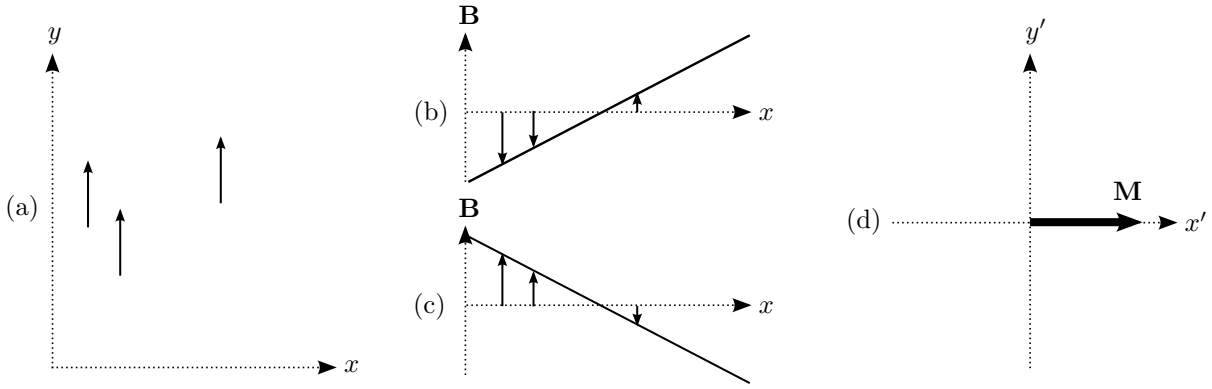


Figure 3.3: *The case of particles moving along the y axis (a), perpendicular to the diffusion sensitizing gradients, is identical to the effect on stationary particles. Since the particles don't change their position along the x axis, the first gradient (b) and the second gradient (c) will affect the particles in exactly the same way and will cancel each other. The total magnetization is still unaffected (d).*

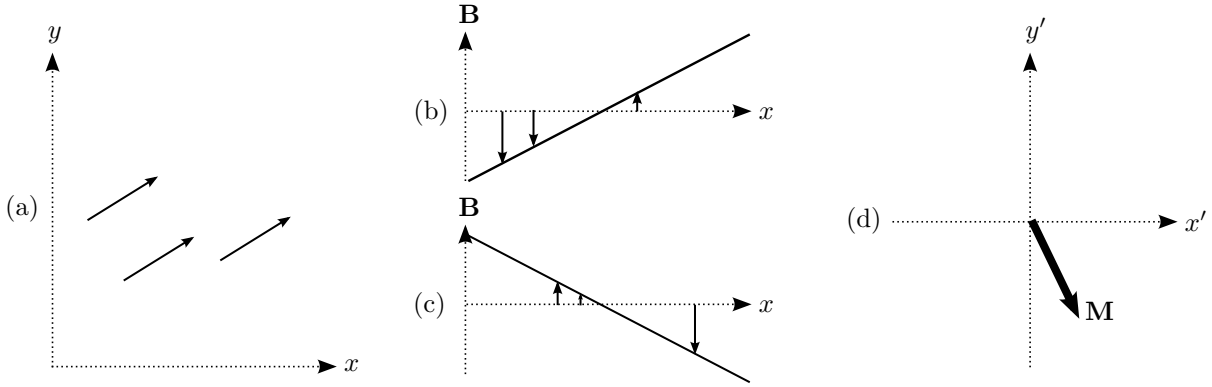


Figure 3.4: *The effect of diffusion sensitizing gradients on particles that experience constant flow during the imaging. All the particles move the same distance along the x axis between the two diffusion gradients (a). This means that the strength of the first gradient on each particle (b) will be less than the strength of the second gradient (c). The effect is that each particle will experience a phase change by an equal amount, hence inducing a total phase change on the magnetization, but not decreasing the total strength (d), leaving the image unaffected.*

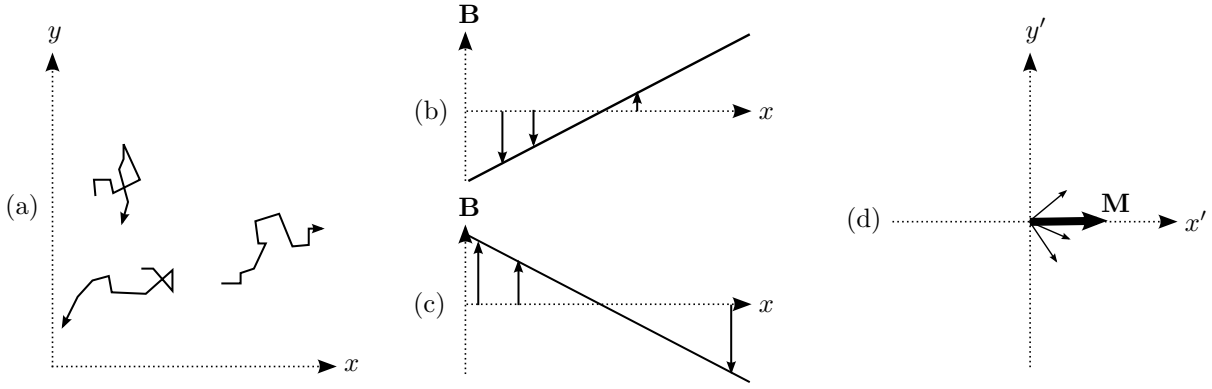


Figure 3.5: *The effect of diffusion gradients on actually diffusion particles (a) is profound. The effect of the first gradient (b) and the second gradient (c) are totally different since the particles have moved a random distance. The particles experiences phase changes that differ in both direction and magnitude. The combined effect is that the total magnetization* $\boldsymbol{M}$ *in (d) is lower than if the particles had not moved. This results in a loss of signal, meaning that pixels will have a lower value the more diffusion the particles contained in it experiences.*

Stejskal and Tanner also derived the effect that the applied gradients have on the signal level in the images [23]. They defined the commonly used diffusion strength parameter b, having the unit sm^{-2}:

$$b = \gamma^2 \delta^2 (\Delta - \frac{\delta}{3}) |\mathbf{G}|^2 \tag{3.6}$$

where Δ, δ and $\mathbf{G}$ are defined as shown in figure 3.1 and γ is the gyromagnetic ratio. The value $\Delta - \delta/3$ is often written as T_d, and is called the effective diffusion time. b is most often referred to using the using s mm^{-2}. With this definition, the b-value can be used as an effective way to indicate the loss of signal strength, now given by

$$\frac{S}{S_0} = e^{-bD} \tag{3.7}$$

where S is the signal value with diffusion gradients present and S_0 is the signal value without them. D is the diffusion coefficient of the medium. Hence, the remaining signal is

$$S(b) = S_0 e^{-bD} \tag{3.8}$$

This equation combines all the relevant diffusion parameters and gives an easy way to describe the effects of the performed measurement. The signal loss experienced for different b-values can be seen in figure 1.2. It should be noted that the b-value is only a total sum of the effect of the diffusion gradients on the image signal. Two measurements with the same b-value can still differ a lot if they have different values for δ, Δ and $\mathbf{G}$.

3.1.3 Diffusion Tensor Imaging

A diffusion MRI image shows the speed of diffusion in one specific direction only. The first applications of diffusion weighted imaging (DWI) used a single such image of the diffusability in some direction to allow detection of larger disruptions in tissue. It was then realized that information from several different images of diffusion in different directions could be combined to give even more information about the body. In 1994, Basser et al. showed how measuring the coefficients from (3.4) could be used to calculate the fractional anisotropy (FA) for each voxel in the MR image [6]. Fractional anisotropy indicates how much the rate of diffusion varies in different directions in each voxel. It can be calculated if diffusion images are taken in at least seven different directions. Since the coefficients $D_{ij} = D_{ji}$ for water molecules, these seven images are enough to find all six unique coefficients for the tensor (seven are required since two data points are needed for each coefficient). The tensor can then be diagonalized to find three eigenvalues λ_1, λ_2 and λ_3 and three eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$. The three eigenvectors are scaled with their respective eigenvalues to create the ellipsoid spanned by $\lambda_1\mathbf{v}_1$, $\lambda_2\mathbf{v}_2$ and $\lambda_3\mathbf{v}_3$, illustrated in figure 3.6, where an elongated ellipsoid shows that the speed of diffusion is higher in the direction where λ is largest. Using these ellipsoids, it is possible to create images such as the one in figure 1.3. In this image, the paths of nerve fibers in the brain have been calculated by linking voxels together when they have clearly elongated diffusion ellipsoids that line up. This is due to the fact that the speed of diffusion is higher along a nerve fiber than perpendicular to it. Such images can be used to locate damaged nerves for certain medical conditions [24].

3.1.4 Distortion Sensitivity

Diffusion images are rarely used on their own as a single image. Instead, clinically useful values such as FA and kurtosis are calculated from whole image set, as shown in the case with Diffusion Tensor Imaging (DTI) above. In the case of DTI at least seven images must be taken to find the six parameters that are required for each voxel. When the analysis is done, these images are used as a single set, and the image value s_{ij} is taken from all the images to be used for the calculation of the tensor for that voxel. When used like this, it is extremely important that all seven images actually show the exact same tissue in the exact same voxels. If the patient has moved, or the image has been otherwise distorted, the calculation will have a certain error since the values compared do not actually belong to the same area. This is the reason why distortions can be very problematic in diffusion imaging, and is why methods such as the one proposed in this work are required to counter the distortions.

Another common use case is the calculation of the apparent diffusion coefficient, ADC. This is can be achieved when one has an image without any diffusion weighting and one or more images of the same area with diffusion gradients applied. Looking at (3.7) above, it is apparent that knowing two different signal values S

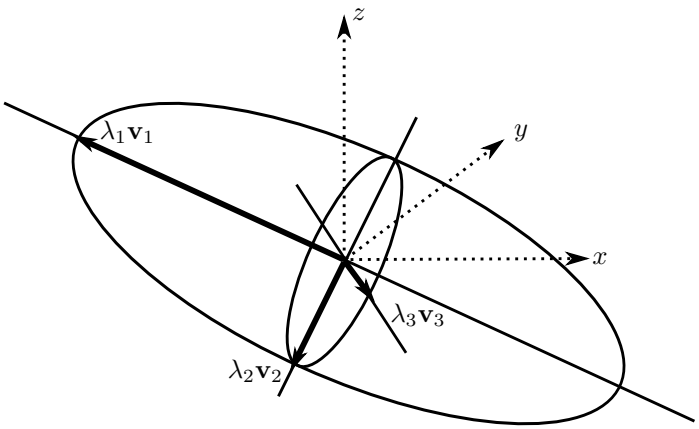


Figure 3.6: *Using diffusion tensor imaging, a diffusion ellipsoid can be created for each voxel. This ellipsoid shows the probability distribution of spin movements during the imaging. It becomes a sphere when the diffusion is equal regardless of direction. Such ellipsoids can be calculated for each voxel using diffusion tensor imaging, and can give further information about the tissue. The ellipsoid is created by the eigenvectors* $\mathbf{v}_i$ *and eigenvalues* λ_i *found by diagonalizing the diffusion tensor.*

and S_0 and all the coefficients involved in calculating the b-value makes it possible to find the value of D if it is unknown. This apparent diffusion coefficient can be calculated voxel by voxel to create a whole map of the ADC, which has proved useful in detecting differences in brain tissue [25]. The important realization in this case is that a correct ADC map requires the measurements from the two compared voxels to have imaged exactly the same area. Any offset introduced by a distortion such as patient movement or a measurement deviation causes a misregistration artifact in the calculated image. The correct comparison of images to produce a final result is thus a major hurdle in producing good results in diffusion MRI.

The possible effect of any distortions grow with the complexity of the comparisons made. Measurements that require comparisons of many different diffusion images are much more susceptible to wrongly aligned images. The situation grows even worse when the measurements depend on images generated with large b-values, due to the much lower SNR available. A measurement of that kind is the kurtosis, which has been found to add valuable information regarding brain tissue [26]. Kurtosis is the variance of of ADC present in a voxel. This variance is a Gauss distribution in the case of completely free diffusion, since the distance that particles have traveled will be normally distributed. This is not the case in the human body, and kurtosis can be used to characterize different types of tissue. This kurtosis value K can be found from

$$S(b) = S_0 e^{(-bD+\frac{1}{6}D^2K)} \tag{3.9}$$

where D is the diffusion coefficient [8].

Kurtosis is studied by the MRI physics group at the Department of Medical Radiation Physics at Lund University, often using the largest possible b-values [9, 27]. This makes the group especially interested in countering the effects of distortions.

3.2 Correcting Distortions

As has been outlined above, distortions can be a sever problem for diffusion weighted MRI measurements. This thesis is focused on the problems caused by eddy currents, and a possible solution for these. Below is found the theory of eddy currents, along with descriptions of previous efforts to create correction methods.

3.2.1 Eddy Currents

Eddy currents are currents that are induced in a conductor whenever the magnetic field passing through the conductor changes. These circular currents are named after the term eddy in fluid dynamics, since they are very similar to these circular flows. Examples of eddies are the swirls surrounding the tip of an oar during rowing and the circular patterns present in weather systems.

Eddy currents arise from the induction of an electromagnetic force when a time-varying magnetic flux is present, in accordance with Faraday's law from (2.3). This force creates a current and an associated magnetic field of its own. Eddy currents normally decay by causing heating of the material, which is the principle of induction heating used in stoves [14].

In MRI, eddy currents are induced in the coils by the fast switching of gradients that is present during much of the imaging cycle. In theory, the perfect gradient is a square wave that switch immediately from off to maximal strength. As with any model, this is not true in a real world system where a certain time, however small, is required to ramp the gradients from zero to the desired level. During this ramping up or down of the magnetic field, the field is time-varying and eddy currents are induced. This process of induction is illustrated in figure 3.7.

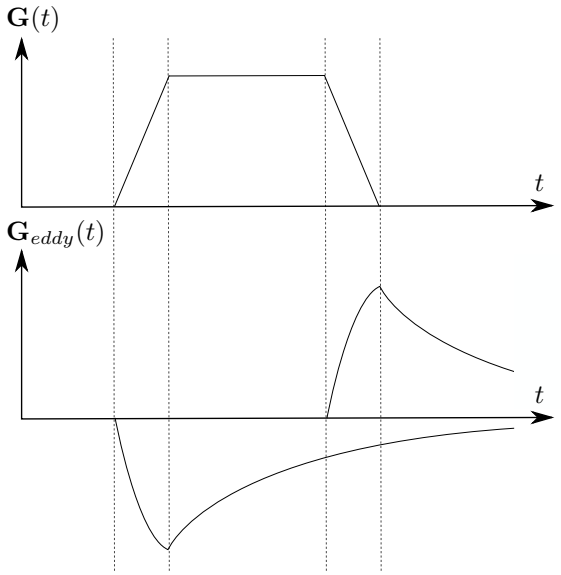


Figure 3.7: *Undesired eddy currents are generated when the magnetic field gradients are switched on and off. The upper part shows how a positive gradient $\boldsymbol{G}$ is applied. Since the ramping of the gradient is not instantaneous, the value $d\boldsymbol{G}/dt$ will be non-zero in the periods indicated by the dotted lines, causing the induction of eddy currents shown in the lower part of the figure. The eddy currents give rise to a magnetic field gradient $\boldsymbol{G}_{eddy}$ that then decays as the current is turned into heat. These undesired extra gradients will cause problems if they remain when other gradients are used, especially during the readout phase. The growth of eddy currents is non-linear due to the decay that begins as soon as the induction starts. Note that this figure illustrates the principle and that the y axes do not have the same scale. The induced gradient is in reality much smaller than the diffusion inducing gradient.*

With eddy currents present, a residual undesired gradient remains after the intentional gradient has been switched off. This remaining gradient will affect any further gradients applied so that the effect is now $\mathbf{G}+\mathbf{G}_{eddy}$, which can give rise to a major change in the imaging sequence when $\mathbf{G}_{eddy}$ is not insignificant compared to $\mathbf{G}$. This is often the case in diffusion imaging, where the whole k-space is scanned for each image. The small gradient blips that are used to change the k-space position, illustrated in figure 2.10, are easily affected by

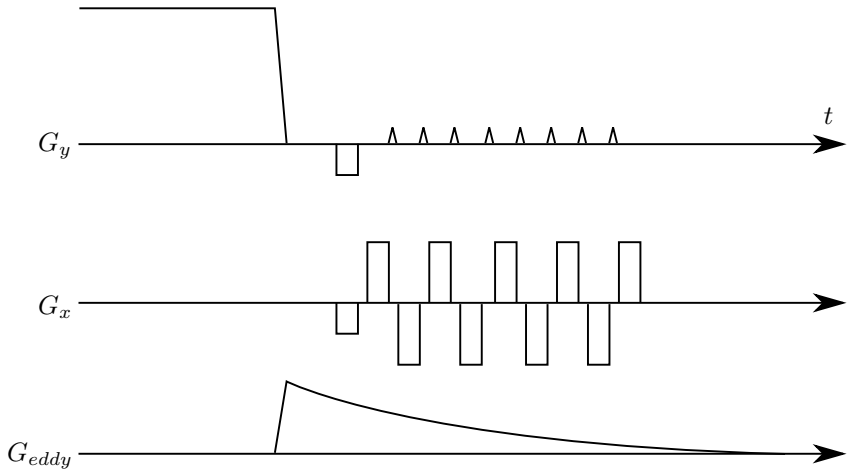


Figure 3.8: *The readout sequence affected by an induced eddy current. A diffusion gradient applied along the y axis at the top gives rise to an eddy current gradient that stretches into the readout phase during which the k-space is scanned by applying smaller gradients along x and y. These smaller gradients that move the k-space position are altered by the gradient caused by the eddy current. In this case, x is the read direction and y is the phase direction, as can be seen by the larger read gradients present for G_x.*

residual eddy current gradients. This principle is shown in figure 3.8 where a diffusion gradient is switched off, inducing an eddy current that cause a gradient that slowly degenerates during the actual readout period. The effect of an eddy current gradient being applied on top of the readout gradients is that the movements in k-space will not be as expected. Instead of the even sampling pattern shown in figure 2.9, the actual sampling will be offset in some direction by the eddy current gradient. Since the reconstruction algorithm will still treat the data points as having been sampled at the expected locations in k-space, the image will be generated based on an incorrect Fourier transform. The resulting effect on the image is a distortion of the imaged substance [28]. Such an eddy current distorted diffusion image is shown in figure 3.9, where the image that is affected has experienced significant transformation. From this figure, it is immediately apparent that calculating values per voxel based on overlaying these two images would be fraught with errors.

Eddy currents have experimentally been found to be an issue in diffusion imaging using EPI readout. The residual gradients caused by eddy currents can be modeled by the following equation:

$$\boldsymbol{\varepsilon} = \sum_i \boldsymbol{\varepsilon}_i \, e^{-t/\tau_i} \tag{3.10}$$

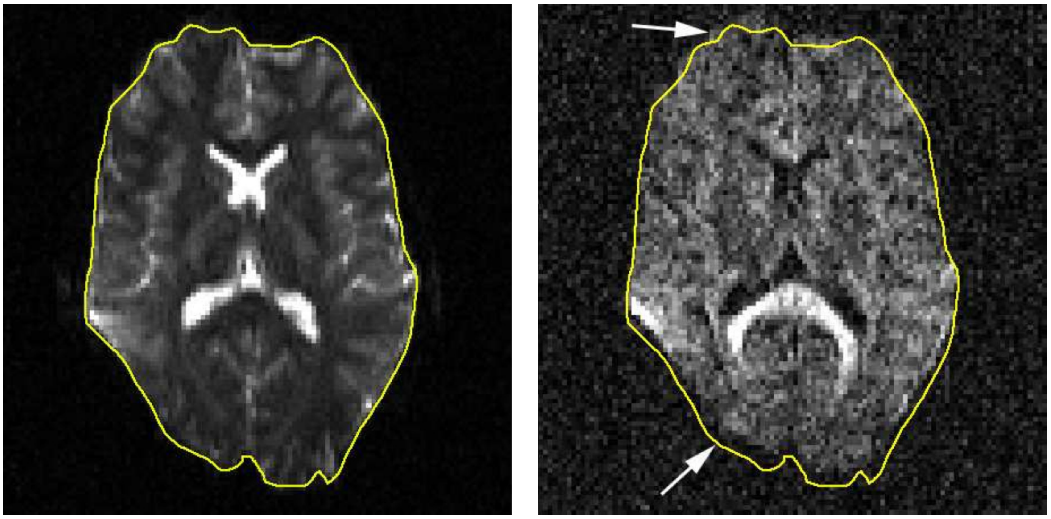
where $\boldsymbol{\varepsilon}$ is the eddy current gradient and $\boldsymbol{\varepsilon}_i$ and τ_i are the individual factors that are used to model the decline as a sum of exponentially declining gradients. The time factors have been found to be in the order 1-100 ms, which means that the declining currents will have a significant impact for a few hundred milliseconds. Since a typical EPI readout is about 40-50 ms long, and a full image cycle is in the order of 100 ms, this indicates that eddy current gradients are present during the EPI readout and need to be taken into account [12, 29].

3.2.2 The Effect on Images

It has previously in (2.11) been established that the MRI signal after filtering depends on the spin density $\rho(\mathbf{r})$ and the phase accumulated by the gradients $\phi_G(\mathbf{r}, t)$. It has also been shown that we can use the k-space representation to describe the signal (2.13) and that this can be expanded to two or more dimensions. Standard MRI images are created by sampling the signal in a two-dimensional k-space and calculating the spin density from the signal using a two-dimensional discrete Fourier transform. This means that the signal will be

$$s(k_x, k_y) = \sum_x \sum_y \rho(x, y) \, e^{-i(k_x x + k_y y)} \tag{3.11}$$

The steps k_x and k_y taken in k-space are determined by the phase change that occurs between adjacent sample points. To calculate the effect of the induced eddy currents, we need to take a closer look at how this phase



(a) (b)

Figure 3.9: *Illustration of eddy current distortion in a diffusion-weighted image. Baseline image without diffusion weighting (a) and the same slice imaged using diffusion sensitizing gradients (b) with the border of the brain from the left image overlaid on both images. This clearly indicates how the shape of the subject has been distorted by the eddy current effect. The areas with clear such effects are indicated by the arrows. The diffusion weighted image was captured using a diffusion gradient of 29.3 mT/m with duration $\delta = 20$ ms. The low resolution of these images is due to the use of EPI for readout. Images courtesy of Markus Nilsson.*

change is affected. If we assume the normal linear EPI readout gradients G_r in the read direction x and G_p in the phase direction y, this phase change is

$$\Delta\phi(x,y) = \gamma\left(\int_0^t G_r(t)x\,dt + \int_0^t G_p(t)y\,dt\right) \tag{3.12}$$

This phase change governs a normal EPI readout that is unaffected by eddy currents [30]. When eddy currents appear, a new term will have to be introduced. To start with, we can define the diffusion gradient as a vector on the form

$$\mathbf{G}_{diff} = c_x\hat{i} + c_y\hat{j} + c_z\hat{z} \tag{3.13}$$

where c_i is the gradient strength in the given direction, measured in T/m. This gradient will give rise to an eddy current induced gradient on the form

$$\mathbf{G}_{eddy}(t) = \varepsilon_x(t)\hat{i} + \varepsilon_y(t)\hat{j} + \varepsilon_z(t)\hat{z} \tag{3.14}$$

where ε_i is a gradient strength in the same way as in the previous equation. This gradient is proportional to the characteristics of the diffusion gradient, with larger gradients giving rise to a larger eddy current gradient. It is, however, not linearly proportional to the gradient strength in the same direction, i.e. ε_x is not necessarily proportional to c_x. The induction of eddy currents in the scanner is a complex process where currents may appear in many different surfaces. It has been shown that cross-terms where c_x leads to ε_y, and so on, are a significant part of the eddy current gradient. There is also a potential effect from eddy currents on the $\mathbf{B}_0$ field, which is expressed as a term $\varepsilon_0(t)$ in the direction of the static field. The combined eddy current gradient effect is thus

$$\mathbf{G}_{eddy}(t)\cdot\mathbf{r} + \varepsilon_0(t)\hat{B}_0 \tag{3.15}$$

where $\mathbf{r}$ is a positional vector measured from the center of the magnet bore. This new term can now be added to the phase change (3.12) that becomes

$$\Delta\phi(\mathbf{r}) = \gamma\left(\int_0^t G_r(t)x\,dt + \int_0^t G_p(t)y\,dt + \int_0^t \mathbf{G}_{eddy}(t)\cdot\mathbf{r} + \varepsilon_0(t)\,dt\right) \tag{3.16}$$

This expression combined with the definitions given in figure 3.10 makes it possible to calculate the eddy current effect on the phase changes in each direction. We also make the assumption that the eddy current gradient will

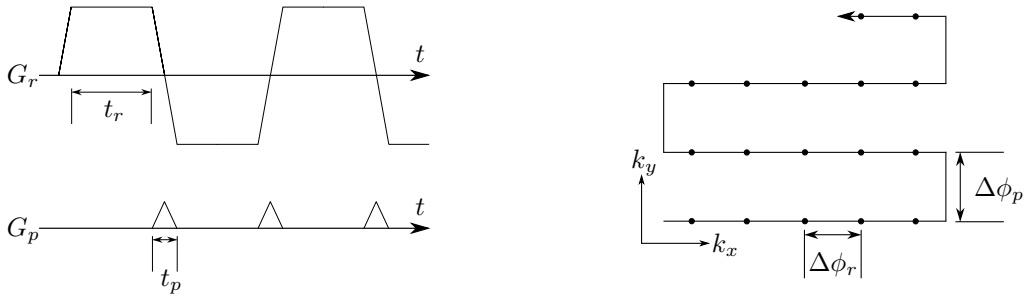
be constant during the EPI readout, which is reasonable since the eddy currents last several times longer than the time a readout takes. The total phase change between two sample points along the read (x) axis is then

$$\Delta\phi_r(\mathbf{r}) = \gamma\left(G_r x\frac{t_r}{N} + (\mathbf{G}_{eddy}\cdot\mathbf{r})\frac{t_r}{N} + \varepsilon_0\frac{t_r}{N}\right) \tag{3.17}$$

where t_r/N is the time period between the two samplings, since N samplings are made during t_r when the read gradient is constant. The phase change during the time t_p when the phase gradient is active becomes

$$\Delta\phi_p(\mathbf{r}) = \gamma\big(G_p y\, t_p + (\mathbf{G}_{eddy}\cdot\mathbf{r})\, t_r + \varepsilon_0\, t_r\big) \tag{3.18}$$

where it should be noted that the phase direction gradient does only alter the phase during t_p, while the eddy current effect is active and accumulates phase drift during the whole period t_r between two adjacent phase blips.



(a) (b)

Figure 3.10: *Illustration showing readout gradients and the desired path through k space. (a) Real shape of the readout gradients in an EPI readout. The gradients along the read axis are trapezoidal while the smaller phase direction gradients are triangular. The phase direction blips are applied while the readout gradients change direction. We define this period as t_p and the period during which the read gradients are constant as t_r. Note that the scale is not correct, the amplitude of the read gradients is about a hundred times larger than the amplitude of the phase gradients, and the difference between t_r and t_p is in reality much larger than it appears here. (b) shows the k space path generated by the gradients in (a), with the sampling points indicted by the circles. The phase change induced during a read blip is defined as $\Delta\phi_p$ and the phase change between two samplings along k_x as $\Delta\phi_r$. The number of sampling points N are in reality much higher, normally between 64 and 256. The definitions given here means that the time between two sampling points is t_r/N.*

Using these equations, it is now possible to estimate the potential impact of the eddy currents on the readout sequence. The read gradients used in EPI are often as strong as possible to make the readout scan fast enough. This means that the gradient G_x can be about 50 mT/m, compared to the eddy current gradient that is normally less than 1 mT/m [31]. This means that in the read phase equation (3.17) we have

$$G_r x\frac{t_r}{N} \gg (\mathbf{G}_{eddy}\cdot\mathbf{r})\frac{t_r}{N} + \varepsilon_0\frac{t_r}{N} \tag{3.19}$$

The total impact of eddy current effects on the phase in the read direction is thus very small and can be ignored in the total calculation of the eddy currents effect.

The situation is unfortunately very different for the phase change in the phase direction. The time-average of the gradient blips that move the readout in the phase direction in k space is very small, and can thus be of approximately the same strength as the eddy current gradient. This is definitely probable when the eddy current inducing diffusion gradients can be up to 100 mT/m, as is the case with kurtosis sequences. We can then conclude that the eddy current impact is substantial in the phase direction, since

$$G_p y\, t_p \approx (\mathbf{G}_{eddy}\cdot\mathbf{r})\, t_r + \varepsilon_0\, t_r \tag{3.20}$$

and the eddy current part can even be larger than the intended phase change.

When it comes to any potential phase change in the z direction due to eddy currents, this can safely be ignored. Any gradients in this direction will not change the EPI sequence dramatically since there are no

z gradients present in the readout sequence. It would only amount to a minor change in the slice selection, resulting in a small effect on the SNR [31].

The combined effect from all eddy currents that need to be included in any correction method is thus the terms given in (3.18), which can be expanded into

$$\Delta\phi_p(x,y,z) = \gamma\big(G_p y\, t_p + (\varepsilon_x x + \varepsilon_y y + \varepsilon_z z)\, t_r + \varepsilon_0\, t_r\big) \tag{3.21}$$

The total impact on the generated images can be determined by analyzing the different eddy current terms and their effect on the normal phase change separately. With no eddy currents present, the expected phase change would be

$$\Delta\phi_p(x,y,z) = \gamma\, G_p y\, t_p \tag{3.22}$$

The effect on this phase change with an eddy current effect from ε_x present, would be an added x dependent term

$$\Delta\phi_p(x,y,z) = \gamma\big(G_p y\, t_p + \varepsilon_x x\, t_r\big) \tag{3.23}$$

This effect on the phase change carries over into the step Δk_y so that the real k_y becomes $k_y + \varepsilon_x x\, t_r$. This new term will end up in the inverse Fourier transform that generates the final image and will change the y transform so that

$$\sum_y \rho(x,y)e^{-ik_y y} \Rightarrow \sum_y \rho(x,y)e^{-i(k_y + t_r\varepsilon_x x)y} \tag{3.24}$$

This additional term will cause a shearing along the x axis since the effect on the inverse Fourier transform by an added term in the exponential is

$$\mathcal{F}^{-1}\left[\hat{f}(\xi)e^{-ia\xi}\right] = f(t-a) \tag{3.25}$$

In our case we will end up with an offset of $\gamma t_r \varepsilon_x x$ in the y direction for each pixel. This shearing effect is illustrated in figure 3.11b.

Analogously, if only ε_y is present, (3.21) is reduced to

$$\Delta\phi_p(x,y,z) = \gamma\big(G_p y\, t_p + \varepsilon_y y\, t_r\big) = \gamma G_p y\, t_p\big(1 + \frac{\varepsilon_y\, t_r}{G_p\, t_p}\big) \tag{3.26}$$

which shows that this will result in a scaling of k_y, as illustrated in figure 3.11c, due to the Fourier transform scaling rule

$$\mathcal{F}^{-1}\left[\frac{1}{|a|}\hat{f}\left(\frac{\xi}{a}\right)\right] = f(at) \tag{3.27}$$

The effect of a potential ε_z gradient can be completely ignored. This is due to the slice selection gradient explained in section 2.2.1. This gradient is played out at the beginning of the imaging sequence and makes sure that only the desired slice has its spins at the correct resonance frequency. At the same time this adjustment makes it possible for us to regard z as zero for each image. Thus, the ε_z term from (3.21) will always disappear.

Finally, the eddy current effect on the static magnetic field has to be considered. With only ε_0 present, the remaining phase change is

$$\Delta\phi_p(x,y) = \gamma\big(G_p y\, t_p + \varepsilon_0 t_r\big) \tag{3.28}$$

which is identical to the effect that caused the shearing, except that there is no x dependence present. Hence, instead of a shearing, the effect on the image is an offset along the y axis that is the same for all pixels, i.e. a translation of the image. This effect is illustrated in figure 3.11d.

The complete effect on an image caused by eddy current gradients consist of distortions along the phase axis in the shape of shearing, scaling and translation of the image [31, 32]. The full set of distortions are illustrated in figure 3.11, while an example on an actual image is shown in figure 3.9.

3.2.3 Existing Correction Methods

The existence of eddy currents is a problem that has been known since the early times of NMR. As the potential of diffusion measurements was made apparent, studies of active measures to counter this problem were initiated. These counter measures can be roughly divided into three different categories.

The first type consists of methods that are integrated in the MRI/NMR scanners at hardware or software level. They are completely general and are applicable to all sequences, and can be seen as preprocessing

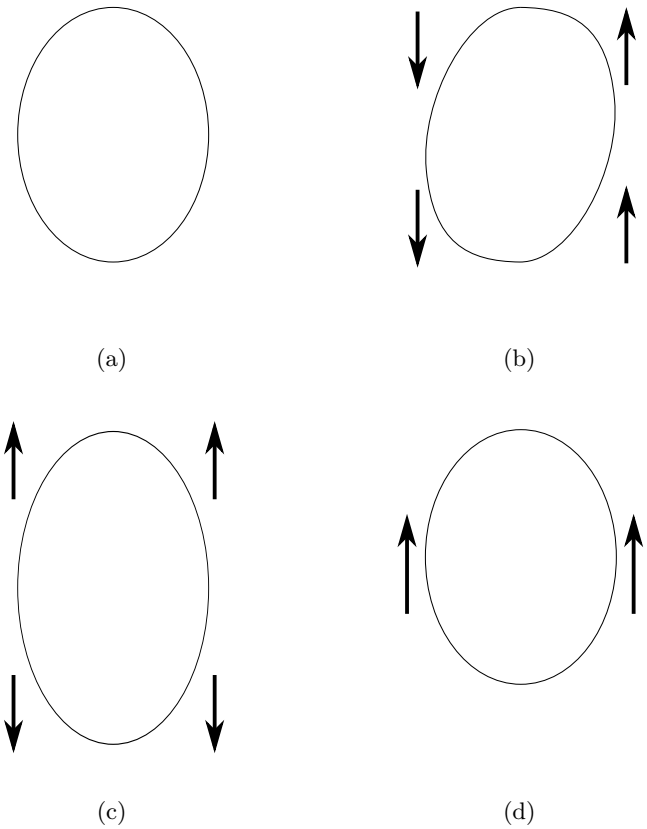


Figure 3.11: *Illustrations of the different distortions that can occur due to eddy currents. An unaffected imaged shape is shown in (a) for comparison. The read direction is horizontal, while the phase direction is vertical in all images. Shearing (b) is an effect where each pixel is offset depending on its position along the read axis. The next effect is scaling (c), where the image is either compressed or extended. Finally, eddy currents affecting the static magnetic field can cause translation (d). All these three effects can also appear in the opposite direction to the arrows shown here, depending on the direction of the eddy current gradients.*

methods. The second category consists of methods that are pure post-processing methods. They are applied only after the image sets are generated, and do not require any sequence modifications. The final type of methods require special sequences, or modification of existing imaging sequences to function. They work by either canceling eddy currents straight away, or by supplying data that makes it possible to remove the eddy current effects in the image generation step.

The first improvements that were made were in the early days of MRI when the hardware was quickly evolving. It was discovered that shielding the gradients and rearranging them could lower the electromagnetic coupling between the gradients and other parts of the scanner, resulting in substantially less induced eddy currents [33]. Such easy gains were quickly incorporated into the scanners by the manufacturers. In the early 1990s, attention turned to improving the sequences being used. The eddy currents induced were mathematically described and sequences were developed to counteract the eddy current induction [34, 35]. These countermeasures work by slightly modifying the gradients applied during readout to get the desired path in k-space even with eddy currents being present. Hence the new sequence would have a term that cancels the results function of the eddy currents, known by calibration. They were found to work well and have been integrated into modern scanners as they lower the impact of eddy currents for all sequences.

Early diffusion sequences only sampled one row in k-space per excitation. The capabilities of MRI scanners improved quickly, and in the early 1990s it was possible to use EPI readouts where the whole k-space is sampled during one echo (explained in 2.2.3). This made DTI possible, and it became an established method in the middle of the decade. The much higher eddy current sensitivity of the EPI scans required for DTI gave rise to a burst of articles on distortion correction, both concerning sequence modification and post-processing methods.

The first such method was presented by Haselgrove and Moore [36]. This method was the first eddy current correction that was built on co-registration of images, which means that one image (or whole volume) is altered in some way until it becomes as similar as possible to a base image that is supposedly more correct. In diffusion imaging, this is most commonly done by comparing a diffusion-weighted image to a normal image that is not diffusion-sensitized, since the latter does not include any diffusion related distortions. Co-registration is done by maximizing or minimizing some kind of comparison function, of which there are many. The strategy outlined by Haselgrove and Moore was to correct the diffusion images by applying shearing, scaling and translation as discussed in the previous section. This was done by changing column after column in the image to maximize the comparison against the same column in the base image. Since the diffusion should only change the image in the phase direction, shearing, scaling and translation can be combined into scaling and translation of each column. The study used a comparison function called cross correlation for the co-registration. Cross correlation is dependent on images having a comparable overall intensity, meaning that it is only usable for diffusion images taken with a low b-value. The reason for this is apparent in the huge difference in signal strength between figures 1.2a and 1.2d. Haselgrove and Moore corrected images taken with $b = 160$ in this way. By visual inspection, they deduced that the distortion was linear with the b-value and that the correction parameters found from the first image could be extrapolated to images taken with higher diffusion strength. While this method was found to be acceptable, it has the large disadvantage that a whole extra set of images with low diffusion strength has to be added to the sequence just to find correction parameters, since the images used for DTI normally has a b-value of about 1000. The distortion parameters found are also not restricted by the physics involved, meaning that accidental similarity of two columns could lead to mis-registrations.

The principle of correction as outlined by Haselgrove and Moore has been the basis of most post-processing methods developed since. An early improvement was to use it with sequences that suppress the intense signal from cerebrospinal fluid (CSF), thus lessening the negative effect of the signal difference that is a problem to the cross correlation function [37, 38]. This makes it possible for the original correction algorithm to be effective for DTI sequences without the need to collect an additional image set with a low b-value. It does however come at the cost of lower SNR in the diffusion images, which has a negative impact on calculation of DTI.

The problem with signal intensity and cross correlation has since been circumvented by not using cross correlation at all, but instead using the comparison function known as mutual information. This function is very robust towards signal differences, and is shown in a study to be consistent over separate imaging sets of the same subject, which was often not the case with the method as proposed by Haselgrove and Moore [39]. Another method to handle this question was suggested by two separate studies in 2004. They show that instead of comparing diffusion weighted images to non-diffusion weighted images it is possible to compare diffusion images taken with completely opposite diffusion gradients. These images have the characteristic of having opposite eddy current distortions while the corrected images should be identical. This is shown to work very well, but does come at the cost of longer sequences since twice the number of diffusion images has to be collected [40, 41].

Another negative aspect of these post-processing algorithms is that they are time consuming to apply on an image set. A processing time of several hours for each patient is not uncommon in these studies. Several articles suggests that this could be solved by running the post-processing algorithms on an image set collected from specially made phantoms. This could then be done less frequently than for each patient, such as once a week. The most promising such method is the one presented by Horsfield in 1999, which is based on a physical model where the theoretical effect of eddy currents is taken into account and then adjusted with mutual information to maximize similarity to non-diffusion weighted images. While it seems that the eddy currents are consistent over time, the method as suggested leaves some remaining errors in the parameters for shearing and scaling [42].

The post-processing methods presented so far have one big negative aspect in common - they don't take patient movement into account. While they all reduce the impact of eddy current distortions, there is always a risk that the distortion correction is wrongly calculated due to patient motion. This possibility exists because the diffusion weighted images are compared to the non-diffusion weighted image takes first in the imaging sequence. This means that the time period between image captures can be several minutes in the worst case, enough time for the patient to have moved a substantial distance. To combat the effect of this, more advanced correction algorithms were created in the early 2000s. These work by comparing whole volumes of images and finding patient motion and eddy current distortion at the same time. One of the most popular such algorithms is the one used in the medical image processing software package FLIRT, which is general method for co-registering medical images [43, 44]. This method is especially used for correction of patient motion in fMRI imaging, where motion can cause severe artifacts since very small differences in blood flow are measured. The FLIRT method finds the affine transformation that maximized the mutual information between two image sets,

using twelve dimensions of freedom. In this way, it can find both patient motion (rotations and translations) and the three dimensions of eddy current distortion. There is one big drawback using this method, which is that there is no correlation between the diffusion gradients in use and the correction applied, and hence no base in the physics involved. It simply finds the best possible match of images, regardless of the reason. Due to this, there is criticism against the method and the unrealistic transformations is can lead to [45].

Some more specific eddy current correction methods have also been created along the same lines. Andersson and Skare showed in 2002 that the remaining residual in the diffusion tensor is a good fitness function that can be used to find both patient motion and eddy current distortion. They used separate models for the two types of transformations, and applied them in the same calculation step as two different transformation matrices. Using this, they convincingly showed that the method works well and that the residual from the diffusion tensor [46] is a good fitness function. They also importantly used limits for the allowed transformations to avoid any unrealistic corrections.

The most thorough post-processing method is probably the one presented by Rohde et al. in 2004 [47].Their method models eddy current effects from the applied diffusion gradients, giving a model soundly based in the physics involved. They use mutual information to co-register non-diffusion weighted image volumes and image volumes taken with $b = 1120$. The method is evaluated and found sound by showing that the method maximizes the information captured in the first few components of a principal component analysis (PCA). The method considers parameter limitation and and also signal change and rotation of the diffusion gradient matrix due to the applied transforms. The medical image software NordicICE uses this method for image registration.

The final type of methods emerged as the same time as the first post-processing methods. This class consist of methods that require some sort of special sequence to counter the eddy current effect. The first idea for such a method was to counter the eddy current inducing gradients by applying gradients in the opposite direction in close proximity to the diffusion sensitizing gradients. This can be done either by applying the opposing gradients before the slice excitation (prepulse) or by putting focusing and refocusing gradients on each side of the 180 degree pulse (bipolar gradients). While especially bipolar gradients have shown good results towards eddy current correction, they both come at the cost of longer imaging time and lower SNR [48].

Another option is to measure the effect of eddy currents using special navigational phase maps. These are special low resolution samplings of the phase that can be used to determine how far out of phase the signal is due to image correction. The images can then be corrected afterwards using the parameters determined in this way. Jezzard et al. developed such a method in 1998 [31]. They measure the error in the phase direction and successfully correct the eddy currents effects. This works well on interleaved images when each line in k-space is collected during a separate excitation. When used on single image EPI, however, it comes with a penalty that makes the sequence three times longer, which is clearly unacceptable.

This method has since been improved [49] upon by Chen et al. They developed a method where the diffusion sequence for DTI was applied on a phantom, and the eddy current effects from scaling and shearing were then measured through field maps as in the previous method. The parameters found in this way could then be applied later on the patient images. The translation was determined separately for each patient by running an additional two minute scan that measured the $\mathbf{B}_0$ shift using the same type of phase measurements. This vastly reduced both the time and the SNR penalty from the method as suggested by Jezzard et al.

The final method to be described here is by many seen as the general solution to the problem of eddy current distortion. In 2003, Reese et al. developed a completely new diffusion sequence, known as twice-refocused spin echo [50]. This imaging sequence differs from the ordinary spin echo sequence described here by using two different refocusing pulses instead of one. This allows the diffusion gradients to be split into four different blocks that can be adjusted to suitable lengths to cancel out the eddy current effects while still being combined into the desired diffusion strength. The study shows that this configuration is able to completely cancel out the eddy current effect from a single exponentially declining eddy current. This makes the sequence able to remove an overwhelming majority of the eddy current effect, which has made it perfect for common diffusion measurements such as DTI. It does however induce a penalty of an SNR loss of up to 20%, which is an acceptable cost in most applications [49].

3.2.4 Applicability in High *b*-value Sequences

Since much diffusion research effort has been focused on DTI and fMRI, most eddy current correction methods are designed to correct images generated with a b-value of up to 1000. High b-value diffusion sequences is an emerging subject and has thus so far not given rise to any research into eddy current correction. Many of the correction methods described above would not work for the kind of diffusion images used for kurtosis, where

b-values up to 4000 are used. These sequences apply very strong diffusion gradients that can be close to 100 mT/m, compared to the values of 20-30 mT/m used in most correction studies. As discussed previously, these strong gradients mean that the images will have a much lower SNR than normal DTI images (see figure 1.2). This creates huge problems for most established post-processing methods. The lower SNR means that these images will have much less similarity to the $\mathbf{B}_0$ images normally used as reference for comparison, making it impossible to register the images as intended. It also derails Andersson and Skare's method that uses the diffusion tensor for correction, since a correct diffusion tensor cannot be calculated using images with very high b-values. The only study that discusses correction for higher b-values is the article by Rohde et al. who figures that their method could be applicable.

Sequence dependent methods do not have the same problems. Twice refocused spin echo works for high b-value imaging, but the SNR penalty that goes with it it not seen as viable for the kurtosis sequences where the SNR is already very low. The methods, such as the one by Chen et al. that use navigational phase maps could potentially work, but at the cost of additional scan time. Since the scan time for kurtosis sequences is already 30-45 min, all additional time is very undesirable due to the problems inherent with requiring patients to remain completely still for a long time.

All in all these limitations in existing methods makes it desirable to develop a new post-processing method that would work for high b-value sequences.

4 Method

This chapter outlines the eddy current compensation method that was developed for this thesis. It explains the different parts of this new method that together make up a software package that can readily be used to correct an eddy current distorted diffusion image set. It also explains the three different steps that were used to verify the functionality of this new method.

4.1 Modeling and Comparing Images

The correction method explained here is a post-processing method that is based on the different distortions that has been shown to be possible in the previous chapter. This is the same type of compensation that has commonly been used ever since the first post-processing methods. However, here the compensation is further linked to the actual physics of Diffusion MRI by linking the correction parameters to the diffusion gradients that are applied to generate the image. This method also differs from earlier suggestions by co-registering images in a way that has not previously been used. This makes it possible to use this method for b-values larger than 1000. The basic idea is to transform parts of the image set and co-register the images. The total correlation in this set is used as a fitness function for a numeric optimization method explained below, resulting in a set of best fitting transformation parameters. The whole method is implemented as a MATLAB program where the heavily used comparison and transformation functions are written in C for speed.

4.1.1 A Model for Eddy Current Distortion

The basics of the correction method is to find parameters that link the applied diffusion gradients to the shearing, scaling and translation that corrects each image in the set. From the previous theory chapter it is known that the distortion in an image is given by the three parameters ε_x, ε_y and ε_0 (ε_z is always zero) that cause the three different distortion effects. The correction transformation can be applied as an affine image transformation according to

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \alpha & 1+\beta & \gamma \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4.1}$$

where α is shearing, β is scaling, γ is translation and the image is assumed to be located with the phase encoding direction along y. Note that this γ is just a naming of a variable, it is not related to the previous use of γ as the gyromagnetic ratio. All mentions of γ from here on refers to the translation of images. To find these parameters, we create a model whereby all correction parameters for the entire image set can be calculated by determining a smaller set of parameters. This model is dependent on the assumption that the eddy current effects are linearly dependent on the applied diffusion gradients, which is according to finds in previous research [42, 49]. The model calculates the correction parameters in the following way:

$$\begin{aligned} \varepsilon_x: &\quad \alpha = p_1 G_r + p_2 G_p + p_3 G_s \\ \varepsilon_y: &\quad \beta = p_4 G_r + p_5 G_p + p_6 G_s \\ \varepsilon_0: &\quad \gamma = p_7 G_x + p_8 G_y + p_9 G_z \end{aligned} \tag{4.2}$$

where G_r, G_p and G_s are the diffusion gradients in the read, phase and slice directions of the images. Note that γ is instead dependent on the diffusion gradients in x, y and z directions in the scanner's frame of reference. The reason that this parameter differs from the two previous ones is that ε_0 comes from eddy current effects on the $\mathbf{B}_0$ field along z, while ε_x and ε_y are dependent on the eddy currents in the image's plane which differs if the slice stack is tilted due to the patient's position. $p_1 \dots p_9$ are the nine parameters that need to be found to enable the full calculation of correction parameters. Using all nine parameters allows both direct and cross terms from all gradients to come into effect on all three distortions.

The approach so far is similar to some previous correction methods. One more element is however added that needs to be considered for the kind of high b-value diffusion experiments used in this thesis. It has been found, for example by Foerster et al. [51], that the main magnetic field is not stable during long sequences using high power magnetic field gradients. The induced heat and physical vibrations cause the $\mathbf{B}_0$ field to drift, causing a separate translation of the images in the phase direction. The most intense sequence used in the study found a time dependent translation that reached several millimeters during a 30 min sequence, which is

very similar to the sequences used in Lund. To make it possible to adjust for this offset, a time dependent function is added to the translation, which thus becomes

$$\gamma = p_7 G_x + p_8 G_y + p_9 G_z + f(t) \tag{4.3}$$

where $f(t)$ is set to be a second-order polynomial on the form

$$f(t) = p_{10} t^2 + p_{11} t \tag{4.4}$$

The parabola given by this function can bend in either direction, and can also degenerate into just a line. The lack of a non-time dependent parameter in the function is due to the boundary condition that the time dependent translation should always be zero at $t = 0$.

This completes the model for the image distortion effects in a full sequence, giving a total of eleven parameters to find in order to correct the image set.

4.1.2 Comparing Images Using Local Correlation

An important step in any post-processing strategy is the co-registration of images. This involves calculating some sort of value that shows how alike two images are. As discussed in the theory chapter, two methods have been commonly used in different post-processing methods: cross correlation and mutual information. The latter method is currently seen as the gold standard for co-registration [52, 53]. It is based on calculating the joint entropy of the image data, which has the drawback of being a very computationally intensive calculation. Since comparing all the images in a data set would require thousands of these calculations, mutual information was deemed impractical for this project. The other common method, cross correlation, is impossible to use for images with very different signal strength, which is the case with diffusion images. Since both these methods have distinct drawbacks, a third method called local correlation was studied for use in this thesis. This method has been shown to be equal to mutual information for co-registering, while being both faster and avoiding the problems with cross correlation [54, 55].

Cross correlation compares pixels in two images by calculating their difference compared to the averages over the whole images. With the image data designated $\mathbf{r}$ and $\mathbf{t}$ respectively, the correlation becomes

$$CC(\mathbf{r}, \mathbf{t}) = \frac{\sum_i (r_i - \bar{\mathbf{r}})(t_i - \bar{\mathbf{t}})}{\sqrt{\sum_i (r_i - \bar{\mathbf{r}})^2 \sum_i (t_i - \bar{\mathbf{t}})^2}} \tag{4.5}$$

where $\bar{\mathbf{r}}$ and $\bar{\mathbf{t}}$ designate the average signal value over respective image and i is an index summing over all the pixels in the images. It is immediately apparent that this calculation is very dependent on the signal strengths being similar in order for the two images to weigh in at the same level. Local correlation solves this problem by considering the average of the Moore neighborhood (the pixel and its eight surrounding pixels) for each pixel instead of the global average. Two neighborhoods are compared using the following calculation:

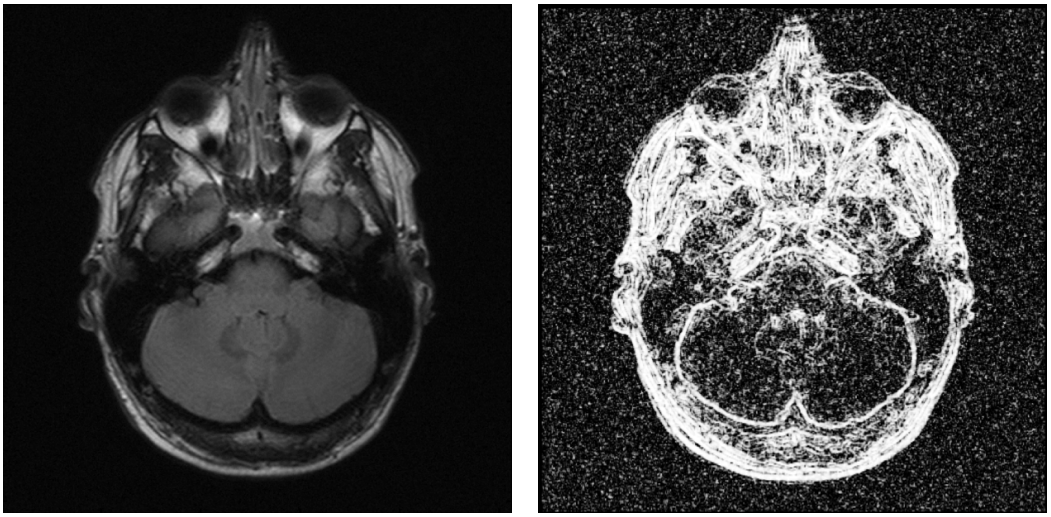
$$[\mathbf{uv}]_i = \sum_{j \in n(i)} (u_j - \bar{\mathbf{u}}_i)(v_j - \bar{\mathbf{v}}_i) \tag{4.6}$$

where we sum over the neighborhood $n(i)$ of pixel i and $\bar{\mathbf{u}}_i$ is the average value for the neighborhood. The complete local correlation is then calculated as a sum over such partial sums as

$$LC(\mathbf{r}, \mathbf{t}) = \frac{1}{N} \sum_i \frac{[\mathbf{rt}]_i^2}{[\mathbf{rr}]_i [\mathbf{tt}]_i} \tag{4.7}$$

This value will be much more robust than cross correlation when it comes to differences in signal intensity due to its focus on local differences. The definition above makes local correlation an efficient comparison that indicates how well the edges in two images line up with each other. The way each pixel contributes to the total local correlation is illustrated in figure 4.1.

During the testing and evaluation of local correlation, it was discovered that the optimization process could be improved by discarding parts of the data. The calculation as defined in (4.7) is dependent on the total number of used neighborhoods N. This has the effect of reducing the local correlation if more pixels are included, even if they are completely black and do not add any further information. This was avoided by adding a filtering function whereby a neighborhood $[\mathbf{uv}]$ is only included in the calculation when its value is



(a) (b)

Figure 4.1: *(a) A normal MRI image and (b) the distribution of values for pixel neighborhoods when the local correlation is calculated for the image compared to itself. It clearly shows that local correlation works by focusing on the lineup of the detected edges in the compared images. The right image also demonstrates that the neighborhood values are rather noisy even when the signal to noise ratio in the original image is very high.*

larger than a certain limit, meaning that the neighborhood must contain a level of contrast to be considered. The local correlation is thus only divided by the number of neighborhoods N_t that pass this threshold. This filtered version of local correlation is used in this thesis and is defined as

$$LC(\mathbf{r}, \mathbf{t}) = \frac{1}{N_t} \sum_i \frac{[\mathbf{rt}]_i^2}{[\mathbf{rr}]_i [\mathbf{tt}]_i} \tag{4.8}$$

where at least one of $[\mathbf{rr}]$ and $[\mathbf{tt}]$ is larger that the designated threshold t. Using this, the local correlation has a much larger difference when images are correctly aligned. Figure 4.2 shows this effect. The vast improvement in local correlation value when the images are aligned make it easier to detect this situation. When many such comparisons are combined, filtering becomes an effective way to make sure that the correct global maximum is more likely to be found by numeric optimization.

4.1.3 Co-Registration Strategy

The filtered local correlation described above is used to co-register different images in the image set. Images are compared according to a scheme designed to include the images where alignment is most likely to be detected. Most post-processing methods developed so far correct the images by finding the highest similarity between the non-diffusion weighted image of a slice and the different diffusion weighted images taken of the same slice. While this works well for b-values up to about 1000, it does not function properly for larger b-values since these differ too much from the non-diffusion weighted image in both image content and signal strength. The scheme is designed to eliminate this problem by using a chain of comparisons between images captured using different b-values. This scheme is illustrated in figure 4.3.

The basic idea of this new strategy is to utilize the fact that two diffusion images captured using the same diffusion gradient direction but with different b-values should be very similar to each other. This means that while it is not possible to make a decent comparison between images with $b = 0$ and $b = 4000$ for example, it is feasible to compare images with $b = 1000$ and $b = 4000$. Using this, we create a chain of comparisons. The image with $b = 0$ is compared to the image in direction A with the lowest non-zero b-value. That image is then compared to the image with the second lowest b-value and so on. Each comparison results in a local correlation value between the two images, and the global fitness function is then defined as the sum of all such local correlations. The perfect transformation should then result in the correction of all distortions, making all the images as similar as possible.

As can be seen in figure 4.3, the scheme does not only compare images captured with the same direction for the diffusion gradients, but also compare images captured with different directions (the arrows connecting

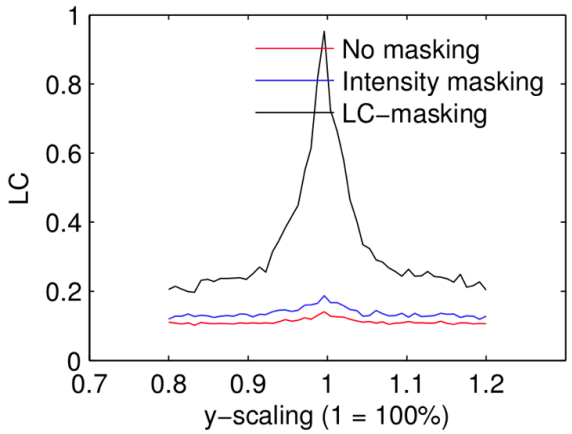


Figure 4.2: *The effect on enabling filtering for local correlation. The three lines show the local correlation values for an image compared to a scaled version of itself, where a y-scaling of 1 is the correlation between the image an itself. The red line shows the local correlation as originally defined in* (4.7)*. The blue line is a test where only pixels above a certain signal threshold was included in the calculation. Finally, the black line shows the effect of using filtered local correlation as defined in* (4.8)*. As can be seen, the function peak is vastly improved, making it much easier to detect for an optimization function.*

images labeled Dir A with images labeled Dir B). The idea here is to provide further information by using the correlation between images with similar gradient directions but the same b-value. For each direction, the closest neighboring direction is found, and each image is then also compared to its neighboring image when the fitness function is calculated for each transformation. This neighboring direction can be either the direction closest to the image's direction, or the direction closest to the exact opposite of the image's direction. This is due to the fact that two diffusion images captured with exactly opposite gradient directions should be identical, as discussed in the theory chapter.

Hence, for an image set with 5 b-values, 10 gradient directions and 15 slices, there is a total of 51 images per slice (there is one non-diffusion weighted image per slice). These 51 images result in 75 separate local correlations for the slice: one for each image compared to the image with lower b-value and one for each pair of diffusion images between neighboring images (same pattern as in figure 4.3 where 7 images result in 9 comparisons). With fifteen slices, testing a set of transformation parameters would require the calculation of $75 \cdot 15 = 1125$ local correlations that are then summed to form the fitness value for the parameter set. Due to the large cost of calculating thousands of local correlations, the calculation is often limited to a subset of the slices, such as five of the fifteen slices in the example. Also note that there are no comparisons made between images from different slices, which is due to different physical regions being present in different slices, making them hard to compare to each other.

4.1.4 Numeric Optimization

The hard part of this post-processing method is to find the transformation parameters that maximize the fitness function that is the sum of all the local correlation calculations made. This requires some kind of numeric optimization method. Test calculations where images were scaled and sheared and then compared with local correlation showed a difficult topography filled with small local maxima. This means that methods such as the Gauss-Newton method, which would otherwise fit very well with local correlation, would easily get stuck in such maxima. Instead, particle swarm optimization (PSO) was chosen as a suitable method for the numeric optimization [56]. PSO is an optimization method inspired by the behavior of bird flocks. The basic idea is to allow a number of particles to explore the multi-dimensional space spanned by the possible range of solutions. In this case, this is the eleven dimensional space consisting of all possible sets of transformation parameters $p_1, \ldots, p_{11}$ within the bounded search-space. Each particle is given a random starting position in this space, which is a candidate solution for the problem. It is also given a random velocity in this space. For each iteration, the fitness function is calculated for the position of each particle. In each iteration the particles will move according to their current velocity. This velocity is updated in each iteration, where each particle is drawn to both the best position it has found itself, and towards the globally best position. This means that even when a

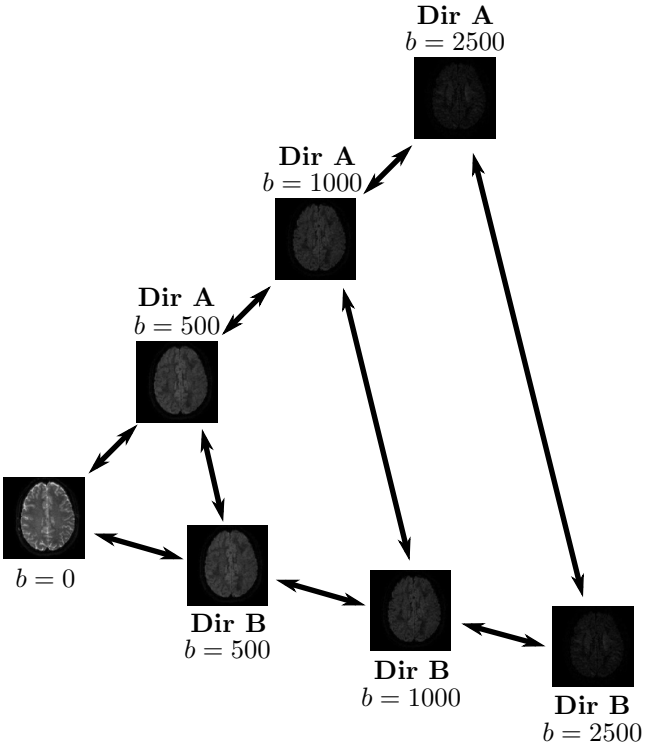


Figure 4.3: *Illustration of the co-registration scheme as a sum of local correlations between different images. This is a partial view of the images that are taken of one slice of the subject, showing only the non-diffusion weighted image ($b = 0$) and the diffusion images for two different gradient directions (A and B). Each arrow indicates the calculation of one local correlation. Hence, each image is compared to the image taken with a lower b-value and to the image with a larger b-value, both using the same gradient direction. It is also compared to the image taken with the same b-value, with the gradient direction that is closest to its own. Thus, the images in the illustration would contribute with nine different local correlation values that are added to the global summation. Note that the length of the arrows is a consequence of the illustration and lacks any importance.*

particle finds a local maximum, its velocity will make it pass over the position and thus avoid being stuck there. In the end, all particles will gravitate towards the best position found by any particle and will further explore the space surrounding that point. The optimization process will be done when the particles have converged so that the average speed is below some predetermined breakpoint. Hence, the method starts with exploration of the total search-space, which slowly changes into a behavior when it exploits the best positions found so far. This makes the method function very well for numeric optimizations in spaces filled with local maxima, such as the one encountered in this thesis. The optimization was run with the default parameters as given in [56], since they were found to work well and result in repeated convergence towards the global maximum.

4.1.5 Application on Diffusion Images

With these tools in place, the full post-processing chain can be applied to diffusion image sets. The method runs on images stored using the DICOM standard (Digital Imaging and Communications in Medicine), which is a common format for storage of medical images. These images currently must contain some tags that are specific to images captured using Philips MRI scanners, but the tag reading could easily be adapted for other manufacturers. When the image sets are read, the DICOM files contain info regarding b-values, gradient directions and timings involved in the sequence. This allows for the calculation of the total imaging time for the set and the individual capture time of each image, which is required to correctly apply the translation calculated from the time dependent function. The only information that needs to be supplied by the user is a target directory with DICOM images, the number of slices in the stack (required to simplify the loading of the images) and the values for Δ and δ used in the diffusion sequence. The two final values are used to calculate the diffusion gradient strength for each image according to (3.6), where the b-value and the gradient direction

are read from the image files.

The method currently assumes that the image set is created using simple slice interleaving where all the odd numbered slices are captured first, followed by the even numbered slices using the same setting. This is commonly done to allow a slice as long relaxation time as possible before it is excited for a new image to be taken. The method can be configured to use all slices in the set for the correction, or to only use a subset. The experiments conducted during this thesis work commonly used 6-8 slices for each run, spaced evenly through the image set. This reduction of slices is done in order to decrease the total optimization time since it was found to have no negative effect on the final parameters.

When the optimization process is started, about 150 particles (depending on settings) are initialized to random positions and given starting velocities as described in the section above. During each iteration, the following steps are performed:

1. For each particle, the parameter values are put into the method to calculate the transformation parameters α, β and γ for each image in the slices involved in the optimization.
2. Each image is transformed according to the these parameters using bicubic interpolation [57] to make them as accurate as possible.
3. The transformed images are compared to each other using local correlation, with the scheme outlined above. The total fitness function for each particle is found as the sum of all these comparisons.
4. The global best and particle best positions are updated, and all the particle velocities are changed.
5. If the average speed of the particles is below the threshold, convergence has occurred and the final solution is found.

This process is preferably run in parallel, to increase the calculation speed. Since MATLAB has no built-in support for parallelization, a multicore enabling plugin written by Marcus Buehren was used [58]. This made it possible to distribute the calculation of all the fitness functions over several cores.

4.1.6 Parameter Reduction

Unfortunately, the numeric optimization of all eleven parameters $p_1 \dots p_{11}$ using PSO turned out to be very hard. Testing showed that while the method converged each time, convergence was not stable enough to be considered a robust correction method. To eliminate this problem, we turned to parameter reduction. The parameters p_3, p_6, p_7 and p_8 as defined in (4.2) were set to zero, leaving seven parameters to find for the numeric optimization. The effect of leaving out p_3 and p_6 is that the gradient in slice direction is no longer allowed to influence the shearing and scaling. Similarly, with p_7 and p_8 removed, the x and y gradients cannot affect the translation of the images. While this removes several dimensions of freedom for the final solution, this reduction is motivated by the findings of previous research in the field, such as the article by Chen et al. [49]. This research shows that the effect of gradients is much larger in the direction of the gradient than the effect on other directions caused by cross terms. Hence, it makes sense to remove the parameters that indicate cross terms if parameter reduction is required. This action ensured that the end product was a correction method with stable conversion towards a credible solution. The final calculation of α, β and γ with parameter reduction in place becomes

$$\begin{aligned} \alpha &= p_1 G_r + p_2 G_p \\ \beta &= p_4 G_r + p_5 G_p \\ \gamma &= p_9 G_z + p_{10} t^2 + p_{11} t \end{aligned} \tag{4.9}$$

4.2 Simulations

With a finished post-processing method in place, a method to verify its functionality was required. This was provided by simulated data with known eddy current distortions. Thus, if the method works as intended, it should be able to find these known parameters when applied to the simulated data. The simulated data was generated by an advanced simulation framework created by Markus Nilsson at the Department of Medical Radiation Physics at Lund University. This framework has previously been used in several studies, such as the one by Mannfolk et al. [59].

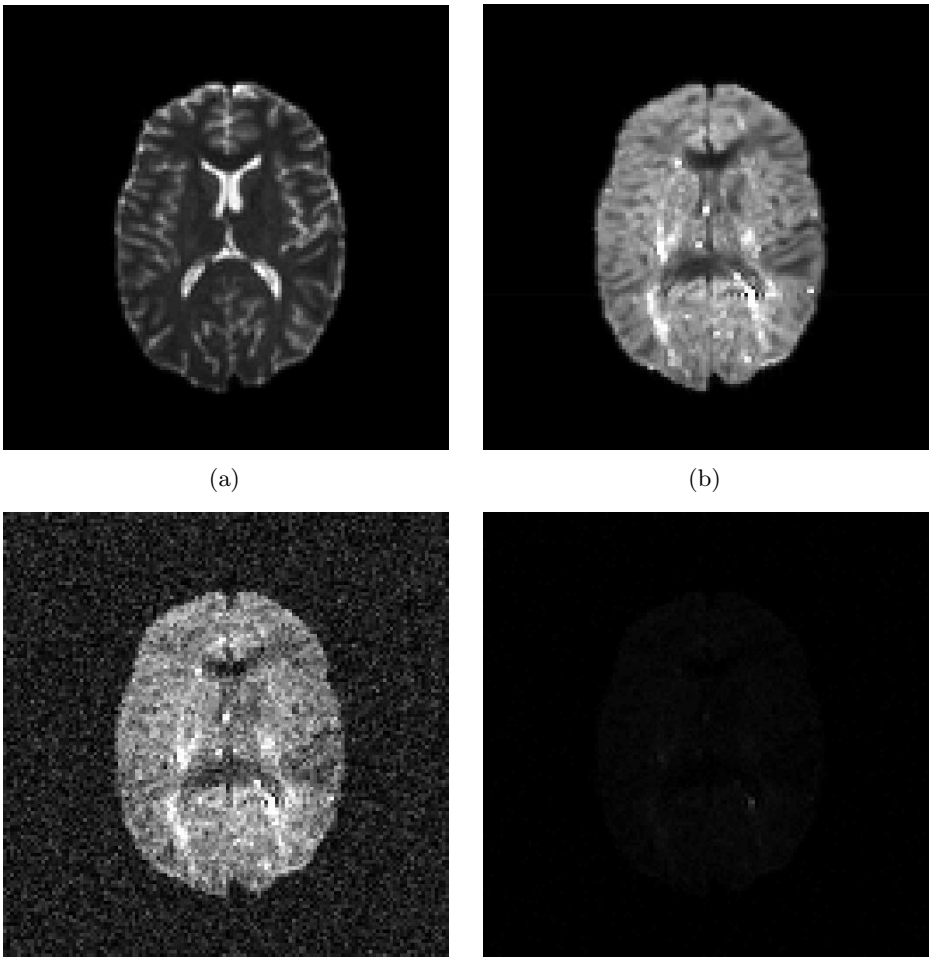


Figure 4.4: *Simulated MRI images showing the same slice. (a) Image without diffusion weighting ($b = 0$). (b) Image simulated with $b = 4000$. (c) Same image as in (b), but with added noise. Note that the images are shown in different scale. In reality, the maximum signal strength in the first image is ten times larger than in the other images. The image in (d) is the same image as in (b), but shown using the same scale as (a).*

The simulations are based on a virtual brain created from a high resolution DTI scan. This brain is divided into cells where the tissue is given realistic properties. Simulated sequences are then played out on this subject, with each cell reacting to the RF pulses and diffusion gradients according to established MRI theory. Each voxel in the final image consists of about 10^3 cells, with each simulation cell having the shape of a cube with sides that are approximately 200 µm. The final result from these simulations is a realistic noise free image with known distortions. This simulated image set has the characteristics found in table 4.1, where TR is the repetition time of the sequence, which is the time difference between two images of the same slice.

The simulations are however somewhat limited. They only include direct eddy currents from each gradient direction, which means that cross-terms aren't taken into account. There is also no time dependent translation effect in the images. Both translation and noise can be added at load time before the correction algorithm is started, which makes it possible to test how these effects are handled as well. Example images from the simulated data set are shown in figure 4.4.

When noise was added to the images, it was randomly added to all pixels using a Rician distribution, which is the noise distribution in MRI images due to the way that the image reconstruction is done. The noise level was adjusted to generate a signal to noise ratio (SNR) of about 30 in the non diffusion-weighted image, which is the normal level for this kind of image. The SNR is calculated according to the method described in [1], using the following equation:

$$\mathrm{SNR} = 0.66 \cdot \frac{S}{\sigma} \tag{4.10}$$

Table 4.1: *Sequence parameters for the simulated diffusion protocol.*

Parameter	**Value**
Nr of slices	60
Nr of gradient directions	15
b-values	$0, 500, 1000, 1500, 2000, 2500, 4000$
Δ	40.0 ms
δ	30.0 ms
TR	8000.0 ms

where S is the mean signal value in the area of the image containing the subject, and σ is the standard deviation of the signal found in the area that is outside of the subject. 0.66 is the Rayleigh distribution correction factor, accounting for the fact that the noise in the images is not normally distributed.

The first results from simulated images with added noise showed that the method failed to find the true parameters with the noise present, instead resulting in transform parameters that were roughly one factor to small. This was found to be due to the background noise outside the subject, which lowered the local correlation values until there was no difference in local correlation for small changes in shearing and scaling. To remove this negative effect, all signal in the input data set below a certain threshold was set to zero for the optimization process, illustrated in figure 4.5. To identify a suitable cutoff, the signal distribution for the images captured with the largest b-value was studied, since these are the images that have the lowest SNR and thus show the smallest difference between the background noise and the real signal from the subject. The threshold was chosen so that it removed most of the background noise while leaving almost all of the real signal untouched. This thresholding of the signal was successful, and was since used in the correction of all data sets from the MRI scanner. The threshold level was set individually for each data set.

4.3 Testing on a Phantom

After the verifications on simulated data, the next step was to test the correction method on real MRI data. An imaging phantom is an object that simulates the behavior of a real subject. It was used in this thesis to eliminate any impact that might come from a moving subject. The phantom needed to have a directionally varying sensitivity to diffusion gradients to make it similar to a human body. It was also preferable that the phantom had an inner structure to give contrast in the image in order to have edges that the eddy current algorithm can detect and use to find the correction parameters. These requirements ruled out the normal silicon oil phantoms that are often used in MRI. Instead, asparagus stems placed in water was chosen as the phantom for this study. Asparagus has different diffusion properties in different directions since water will diffuse more easily along the fibers, and it also gives the phantom the required inner structure. An image of the phantom is found in figure 4.6.

The images of the phantom were acquired on a Philips 3T Achieva camera at the Skåne University Hospital in Lund. Two different diffusion protocols were used, with five series captured using each sequence. The idea behind using different protocols was that it would make it possible to determine by how much the distortion parameters vary between different sequences. By capturing several series for each protocol, it is also possible to test whether the distortion parameters are stable when the same sequence is used several times on a static subject. If the parameters are stable, it might be possible to find the distortion for a certain sequence using a phantom, and then apply these precomputed parameters on the data series from live patients. Thus, only motion correction would have to be calculated in a post-processing step for each patient. The data for the two different sequences are found in tables 4.2 and 4.3. Two example images from the phantom data sets are shown in figure 4.7, where the loss of data for large b-values is very apparent.

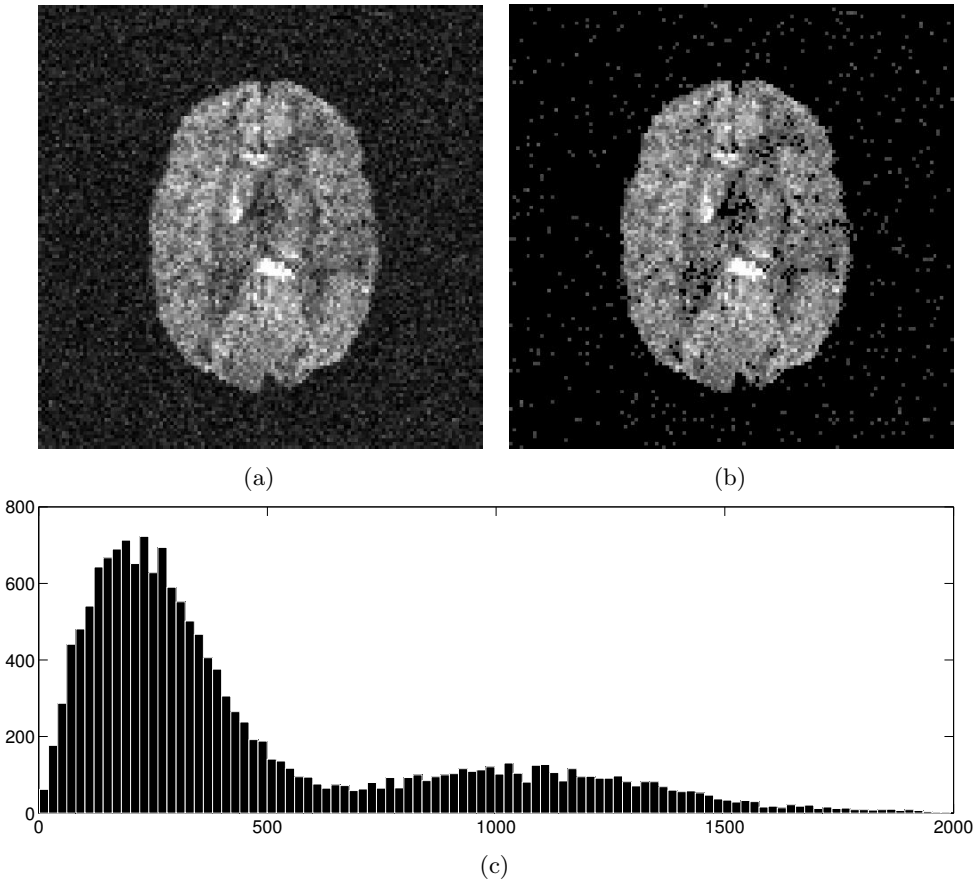


Figure 4.5: *Filtering of noise was found to be necessary since the local correlation was lowered by the background noise. A simulated image is shown in (a), with a realistic noise level added. The image is captured using the highest b-value, making the real signal only slightly stronger than the background noise. The distribution of signal values in this image is found in (c), where the large amount of low signal values to the left shows the noise. The smaller rise in the middle of the image is the real signal. To make the method work on noisy images, all signal below a threshold was removed to eliminate the background noise. (b) shows the same image as (a), using the same scale, but with all signal below $S = 500$ removed. As can be seen in both the image and the histogram, this removes almost all the background noise, while allowing true signal to remain.*

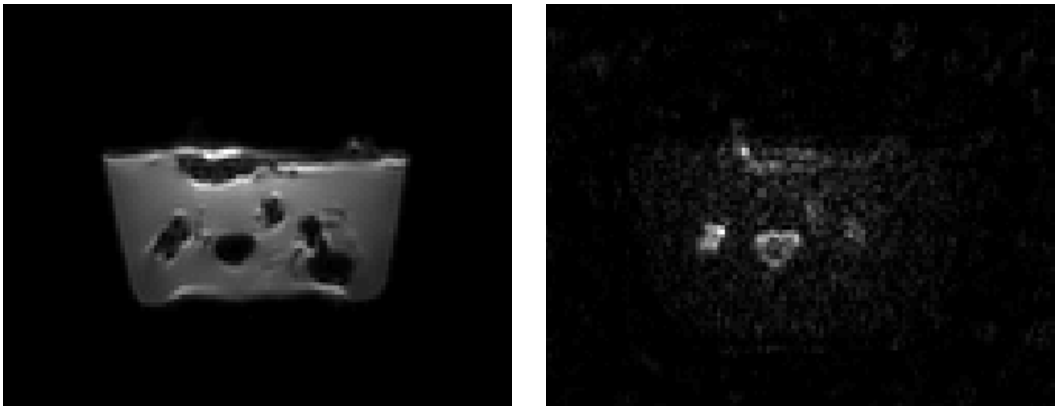
Table 4.2: *Parameters for the first sequence used on the asparagus phantom.*

Parameter	**Value**
Nr of slices	35
Nr of gradient directions	15
b-values	$0, 500, 1000, 1500, 2000, 3000$
Δ	43.7 ms
δ	25.5 ms
TR	5624.7 ms

Figure 4.6: *The phantom used to generate real image data to test the correction method. It consists of asparagus stems placed in water, which gives it both a distinct structural contrast and a directionally varying diffusion sensitivity. This makes it a perfect phantom for testing eddy current distortion. A plastic lid was placed on the container before it was put in the MRI scanner.*

Table 4.3: *Parameters for the second sequence used on the asparagus phantom.*

Parameter	**Value**
Nr of slices	15
Nr of gradient directions	15
b-values	$0, 1000, 1500, 2000, 3000$
Δ	37.9 ms
δ	9.7 ms
TR	6000.0 ms



(a) (b)

Figure 4.7: *Images from the diffusion sequence on the asparagus phantom. (a) No diffusion weighting,* $b = 0$ *and (b) maximum diffusion gradient,* $b = 3000$*. It is apparent that the water surrounding the asparagus stems allows free diffusion, which is why the signal from the water is almost zero in the right hand image. The structure of the stems hinder the diffusion and gives the comparatively high signal that remains inside them.*

4.4 Measuring in Vivo

The final test was to apply the method on in vivo data from diffusion measurements. To accomplish this, three healthy volunteers were examined after informed consent. A kurtosis sequence was used for the measurements. The same sequence was repeated five times in a row on each volunteer, to make it possible to test by how much the movements of a patient would affect the final transformation parameters. These measurements would also be used to find out whether the transformation parameters are the same for different brains imaged using the same kurtosis sequence. The outcome of this would tell if the method can be calibrated using a phantom for later immediate application on in vivo data, or if the optimization process has to be run separately for each subject. Example images from these measurements are found in figure 1.2, while the settings for the imaging sequence are located in table 4.4.

Table 4.4: *Parameters for the in vivo kurtosis measurements.*

Parameter	**Value**
Nr of slices	35
Nr of gradient directions	15
b-values	$0, 500, 1000, 1500, 2000, 3000$
Δ	43.5 ms
δ	25.5 ms
TR	5625.0 ms

5 Results

This chapter summarizes the results from running the post-processing method on the different data sets outlined in the previous chapter. For each data set, the method was run five times with identical settings in order to determine the stability of the resulting transformation parameters. All the results were generated on an Apple Mac Mini Intel Core 2 Duo 2.4 GHz, utilizing both cores. This performed the iterations it generally took for the 150 particles to converge in about four hours. The converging criterion was that the average movement of the particles in each dimension was less than 2% of the allowed parameter space. This generally required about 100 iterations, requiring a total of 15000 evaluations.

5.1 Results from Simulations

5.1.1 Unadjusted Data Set

The first run on the simulated data was made without any noise or time dependent translation being added. Thus, there were only five parameters left to be found, p_1, p_2, p_4, p_5 and p_9 as defined in (4.9). Since the parameters that generated the data set are known, this experiment becomes a test of the ability of the method to find the correct transformation. The method was run on a subset of the full 60 slices, consisting of slices 15, 23, 31, 39, 47 and 55. The results for the five runs are found in table 5.1, along with the mean values for the parameters and their standard deviation. Here are also the true values that should theoretically be found based on the known generating parameters used to affect the simulated image readout.

The results show that the optimization algorithm is capable of consistently finding the parameter set that maximizes the local correlation. The parameters found actually results in a higher local correlation than the expected parameters, as seen in the table, showing that the optimization algorithm works perfectly. When the parameters are multiplied with the gradients, the largest transform parameters found are $\alpha = 0.0140$ (shearing), $\beta = 0.0144$ (scaling) and $\gamma = 0.0023$ (translation), although these are for different images in the set. When these transformations are translated into image pixels the largest pixel movement from shearing inside the imaged subject is 0.48 pixels, the largest scaling is 0.56 pixels and the largest translation is 0.18 pixels (the transformation from translation is calculated as percentage of the image size). This shows that method is capable of finding transformations down to a subpixel level. The largest combined translation of the image set becomes 1.1 pixels, which is a substantial difference when calculating kurtosis for a diffusion set. Each pixel is about 2 mm in the imaged subject.

Table 5.1: *Results for five different parameter optimization runs on the simulated data set without any added noise or time dependent translation. The best result found for each run is presented, along with the number of iterations required for convergence and the local correlation (LC) for the best parameter set discovered. I is the number of iterations required for convergence. The bottom half of the table shows the mean and standard deviation (SD) for the five parameters. The bottom row holds the true values that should theoretically be the final result. The LC presented for the true values is the correlation found if these values are applied to the image set.*

	p_1	p_2	p_4	p_5	p_9	I	LC
Run 1	0.2976	0.0829	-0.3069	0.2906	0.0502	49	0.94080
Run 2	0.2980	0.0830	-0.3079	0.2927	0.0501	45	0.94076
Run 3	0.2974	0.0823	-0.3072	0.2954	0.0500	50	0.94076
Run 4	0.2973	0.0831	-0.3080	0.2903	0.0503	51	0.94076
Run 5	0.2974	0.0781	-0.3072	0.2881	0.0502	49	0.94076
Mean	0.2976	0.0819	-0.3074	0.2914	0.0502	48.8	0.94077
SD (10^{-3})	0.2746	2.1461	0.4826	2.7356	0.1419	2.28	0.018
True	0.3272	0.1091	-0.3272	0.3272	0.0511	n/a	0.94058

As is clearly seen, the results from the optimization do not exactly match the true values that are expected to be the correct transformation parameters. The values found are about 10% lower than expected, except for p_2 where the difference is almost 25% and p_9 where the difference is less than 2%. When this is translated into

real image pixel differences, the largest difference found between the true parameter set and the one found is 0.11 pixels, which is 10% of than the largest transformation applied by the correction.

One way to analyze the impact of the different transformations is to measure the amount of signal found outside the position of the imaged subject. The image taken without any diffusion weighting shows the subject without any eddy current distortions, and the goal of the correction method is to properly realign the other images with this position. Thus, a segmented mask could be created from the image with $b = 0$ in each slice. When this mask is applied to the diffusion weighted images in the same slice, the total image signal outside the segmented position of the subject gives an indication of the effect of the transformations. An effective transform should show that the signal outside the segmented subject should be substantially decreased compared to the situation before any distortion correction is applied. This process is illustrated in figure 5.1, with a mask created from a non-diffusion weighted image shown in 5.1b. When this method was applied slice by slice to one of the transformed data sets, the method was shown to have reduced the total signal outside of the segmendted subject by 11.75%. When the known true transformation parameters were applied, the reduction of signal was 12.32%. This shows that the true parametes are in fact a better correction (as expected), even if they showed a slightly lower local correlation value for the image set. The found value does however show a performance that is very close to the true parameters.

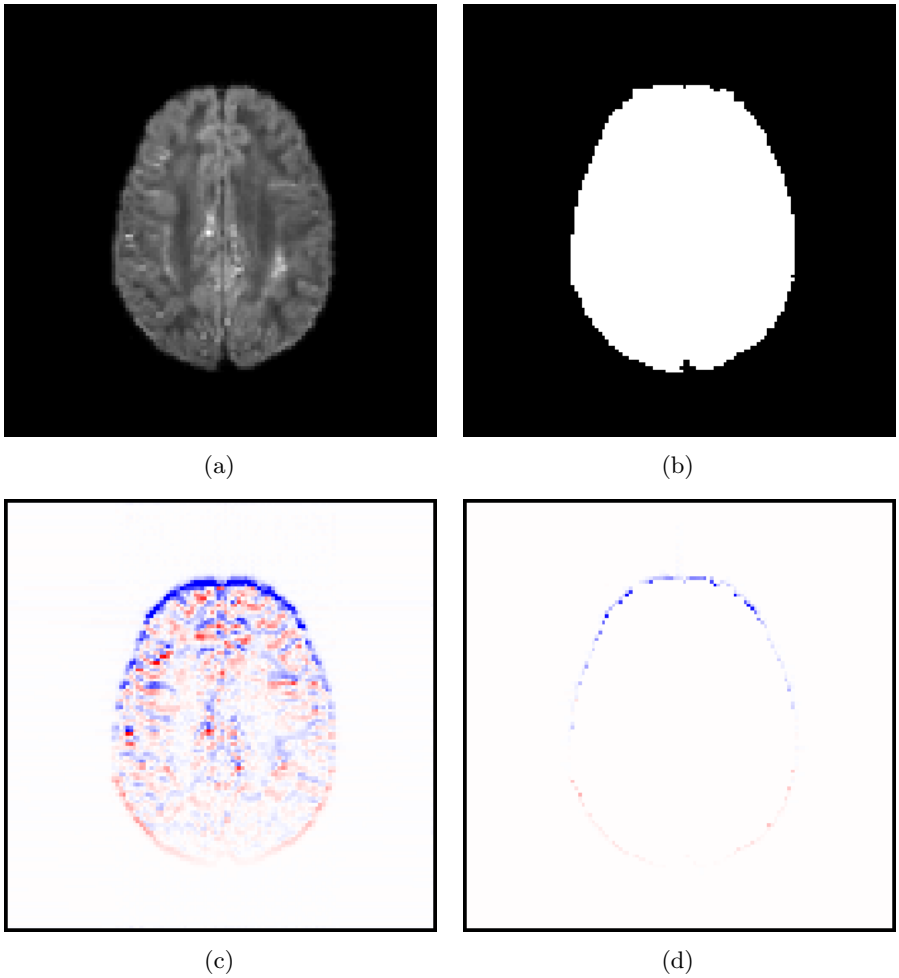


Figure 5.1: *Illustration of the principle of applying a mask to calculate the effectiveness of the correction. (a) uncorrected image from the simulated data set. A mask was created by thresholding the image with* $b = 0$ *in the same slice, using a visually determined suitable level (b). (c) the effect on the image by applying the determined transformation having* $\alpha = 0.001$, $\beta = -0.006$ *and* $\gamma = -0.002$, *which is mainly scaling and translation. Blue color shows area where signal was removed by the transform, with red color the opposite. (d) shows the change for the area outside of the image mask from (b). This shows that signal is mainly removed from the area outside of the brain, exactly what would be expected by a working correction method. In this case, 20.2% of the signal outside of the segmented brain was removed by the correction, with the remaining signal being mainly noise.*

5.1.2 Time Dependent Translation

For the second test using the simulated data, a time dependent translation was added to the image set, shown in figure 5.2. This translation grows with time, until it reaches its maximum of 2%, or 2.56 pixels, when the sequence ends after a simulated acquisition time of 728 seconds. The translation function is quadratic, which is the same shape and size of the translation that is described in the article [51] by Foerster, et al. This adds two new parameters, p_{10} and p_{11}, to the set that needs to be found, making it necessary to find 7 parameters instead of 5.

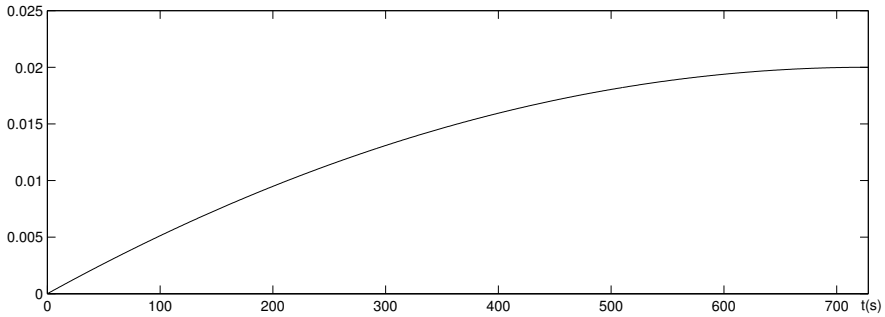


Figure 5.2: *The time dependent translation added to the simulated data set, in percent. The translation function is a quadratic function with its maximum after 728 seconds, when the translation is 2%. Since the images are* 128×128 *pixels, this equals 2.56 pixels.*

The detailed results for five different runs on the data are found in table 5.2, while mean values and standard deviations, along with the true values, reside in table 5.3. The results show that the method is still very consistent, while the added translation is clearly not a problem. The values found for p_{10} and p_{11} are close to the true parameter values, off by 10% and 4%, respectively. While this may seem like a large difference, when the real difference in pixel movement is calculated, it becomes only 2.9%, or 0.07 pixels. Note that these parameters give a negative translation offset, which is required to counter the positive translation that was originally added to the images. The lineup of the found parameters and the true change is seen in figure 5.3, showing a very good lineup.

The values for the five remaining parameters are slightly further off from the true values than they were when no time dependent translation was included. The mean value of p_2 is now 40% lower than the true value, and the mean value of p_5 is 20% off. This is however compensated by a better match for p_4. The largest difference in image pixels becomes 0.22, about 8% of 2.68 pixels, which is the largest pixel movement found. How the aggregated correction works is demonstrated for a single image in 5.4, showing that the algorithm performs very well. Since the extra translation applied to the images move many of them quite far from their correct position, the analysis method using the image mask of the non-diffusion weighted images show completely different values than before. The correction method now manages to remove 64.2% of the signal outside of the image mask. It takes almost twice as many iterations for the PSO to converge, compared to the case without any time dependent translation.

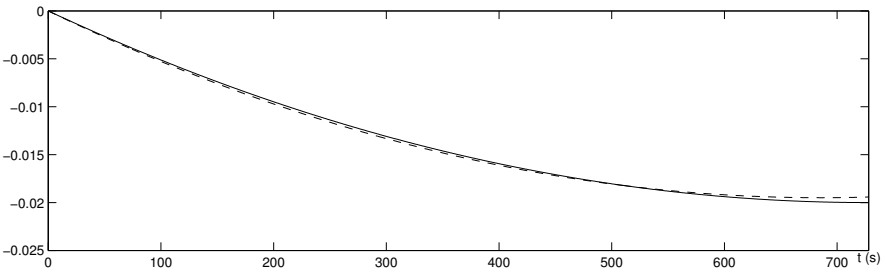


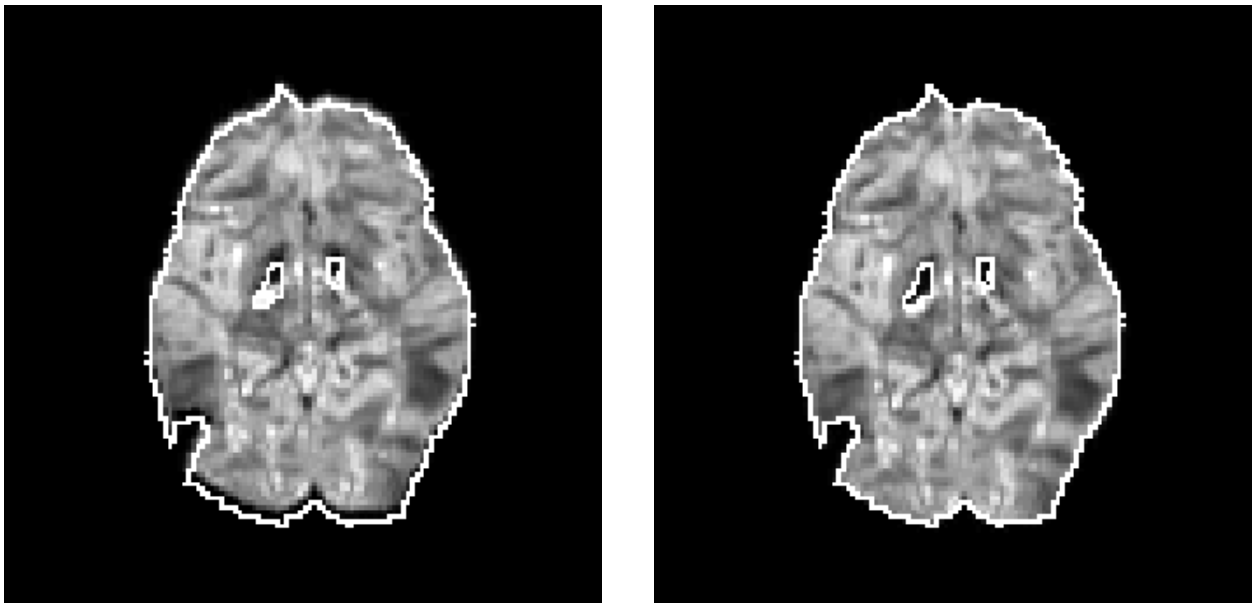
Figure 5.3: *The time dependent translation that should be found is shown as a solid line, while the dashed line shows the translation found by the method during the first run on the data set.*

Table 5.2: *Detailed results for the five runs of the correction method on the simulated data with added time dependent translation. The algorithm is still very consistent when expanded to 7 parameters. It does however take almost twice the number of iterations (I) to reach convergence. The capability of the optimization algorithm is demonstrated by it being able to correctly find the translation parameter p_{10} despite it being in the region of 10^{-8}.*

	p_1	p_2	p_4	p_5	p_9	$p_{10}(10^{-8})$	$p_{11}(10^{-5})$	I	LC
Run 1	0.2802	0.0633	-0.3158	0.2672	0.0496	4.145	-5.685	92	0.92559
Run 2	0.2828	0.0615	-0.3125	0.2661	0.0495	4.153	-5.691	93	0.92559
Run 3	0.2788	0.0678	-0.3134	0.2591	0.0496	4.187	-5.736	88	0.92559
Run 4	0.2786	0.0659	-0.3142	0.2592	0.0491	4.166	-5.733	93	0.92559
Run 5	0.2790	0.0654	-0.3102	0.2586	0.0493	4.164	-5.724	92	0.92559

Table 5.3: *Mean and standard deviation for five runs completed on the same data set as in table 5.2. The bottom row holds the known true values for the simulated data set. The standard deviations are at least two magnitudes smaller than the mean values, showing great consistency between runs. Once more, the true values turn out to result in a slightly lower local correlation value than the values found by the correction method.*

	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	LC
Mean	0.2799	0.0648	-0.3132	0.2621	0.0494	4.163×10^{-8}	-5.714×10^{-5}	0.92559
SD (10^{-3})	1.748	2.414	2.099	4.213	0.228	1.586×10^{-7}	2.401×10^{-4}	1.2×10^{-3}
True	0.3272	0.1091	-0.3272	0.3272	0.0511	3.775×10^{-8}	-5.496×10^{-5}	0.92489



(a) (b)

Figure 5.4: *The effect of correction on a simulated image with added time dependent translation. To the left is the uncorrected image. The bright border shows the outline of the image with $b = 0$ in the same slice, and thus shows the ideal position for the image. It is apparent that the image is far off from where it should be. To the right is the image after correction, showing that the image is now well positioned. The translation applied to this image during the correction was -0.0179, or -2.3 pixels.*

5.1.3 Adding Noise

The noise that was added to the images gave a realistic SNR of 35 in the non-diffusion weighted images, and SNR = 4.6 in the images where b was largest ($b = 4000$). Different random noise was added to the data set before each run of the optimization method. As discussed in the method chapter, the first attempt at correction using the method failed to find the known true correction parameters. The summarized results for these runs are shown in table 5.4. The mean values are very close to zero and far off from the true parameters. It should be noted that the true parameters actually gives a lower local correlation value than the ones found by the method, showing that the optimization method finds the maximum for the fitness function. The background noise apparently causes any gain from applying the correct transformation parameters to be lost.

Table 5.4: *Results when the method was run five times on simulated data with added noise, without any thresholding of the data. The parameters found are close to zero, except for the time dependent translation. This is probably due to the true parameter settings being hidden by the large background noise in the images. The only parameters that are found as expected by the method are the time dependent translation parameters.*

	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	**LC**
Mean	0.0196	0.0031	-0.0365	0.0162	0.0147	5.052×10^{-8}	-5.293×10^{-5}	0.28975
SD (10^{-3})	1.478	7.357	3.251	2.740	1.341	5.725×10^{-6}	9.343×10^{-3}	0.284
True	0.3272	0.1091	-0.3272	0.3272	0.0511	3.775×10^{-8}	-5.496×10^{-5}	0.28664

The problem with noise defeating the method was solved by applying a lower threshold on the input data used by the optimization algorithm. The mean value of the signal S in the area of the imaged subject was roughly 1000 in the images captured with $b = 4000$. The threshold was set to 500, which removed almost all the noise outside of the subject but left the noise that was applied to the area of the imaged subject. The results from five runs on this data are found in tables 5.5 and 5.6. The results here show that with the low level signal removed, the method is once more capable of finding the true parameters. The results are however further from the true parameters than without the noise, and also show larger standard deviation than before. The latter is probably due to separate noise being added to the data for each run. This is done to avoid a situation where a single lucky or unlucky noise distribution might severely affect the results.

The added noise causes a larger deviation from the true parameters than the case with only time dependent translation added. The largest difference between the true parameters and the found parameters varies between 0.33 and 0.49 pixels. This is roughly twice as much as the difference found when no noise was present. Given that the largest pixel transformation applied in each set is slightly more than 2.5 pixels, this is still less than 20% of the total change.

The masking method used before also shows a significant improvement in the amount of signal that is located inside the mask created from the images with $b = 0$. The reduction in signal outside the mask is $63.5 - 68.7\%$ for the thresholded data set. When the removed noise is included in the calculation, this value drops to $3.9 - 4.1\%$, indicating that a very large part of the signal outside of the mask is only added noise. While the true parameters have a lower local correlation value than the found parameters, they once more result in a better reduction of signal outside the mask. The real parameters give $65.6 - 70.2\%$ on the thresholded data, slightly better than the found parameters.

Plots showing the size and variation of the correction parameters for each image are found in figure 5.5.

Table 5.5: *Results for five runs on the noisy simulated data set where the input data to the optimization algorithm was thresholded at $S = 500$. The results are less consistent than previous results, which is probably due to the fact that the noise added to the data set is different for each run. It takes more iterations for the algorithm to converge with noise present, as can be expected.*

	p_1	p_2	p_4	p_5	p_9	$p_{10}(10^{-8})$	$p_{11}(10^{-5})$	**I**	**LC**
Run 1	0.3122	0.0489	-0.3285	0.1612	0.0715	4.310	-5.755	98	0.27696
Run 2	0.2225	0.0536	-0.2935	0.2437	0.0391	5.453	-6.542	121	0.27756
Run 3	0.2792	0.0542	-0.2392	0.2389	0.0342	5.213	-6.416	100	0.27751
Run 4	0.4040	0.0698	-0.4133	0.1797	0.0661	3.825	-5.429	95	0.27718
Run 5	0.2105	0.0329	-0.2641	0.2616	0.0427	5.542	-6.611	107	0.27764

Table 5.6: *Aggregated results for the five runs from table 5.5. The mean values are close to the true values, showing that the thresholding method enables the algorithm to find the true parameters regardless of the presence of noise. As before, it turns out that the local correlation value is lower for the true parameters than for the parameters found for each run. The local correlation value shown here for the true parameters is an average value since the separately applied noise made it different for each run.*

	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	LC
Mean	0.2857	0.0519	-0.3077	0.2170	0.0507	4.869×10^{-8}	-6.151×10^{-5}	0.27737
SD (10^{-3})	78.08	13.20	67.79	43.83	16.89	7.607×10^{-6}	5.278×10^{-3}	0.287
True	0.3272	0.1091	-0.3272	0.3272	0.0511	3.775×10^{-8}	-5.496×10^{-5}	0.27631

5.2 Phantom Results

For the asparagus phantom, two series captured using each of the two sequence protocols were arbitrarily selected for processing using the correction algorithm. As before, five runs were completed on each selected data set, for a total of twenty runs on the four data sets. The input data was thresholded in the same way as with the noisy simulated images. The maximum signal level reaches 47 in the images with the largest b-value (compared to 1400 when $b = 0$), which occurs in the asparagus stems. The surrounding water has a very small signal due to its homogeneous diffusion characteristics. The thresholding level was set to 1 to only remove the background noise surrounding the phantom, and not the small signal from the water. The signal contribution from the water is large when $b = 0$, as shown in figure 4.7, resulting in a large SNR of 188. For $b = 3000$, SNR drops to 5.76, which is comparable to the value for the noisy simulated data set.

Table 5.7: *Results from running the correction method on two image series captured using the first sequence for the asparagus phantom, defined in table 4.2. Values shown are mean parameter values and standard deviation for the best parameters found for the five runs completed on each of the series. The phantom was not moved between acquisitions.*

Sequence 1	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	LC
Series 1								
Mean	0.0089	-0.1710	-0.0122	-0.1360	0.0334	-4.732×10^{-8}	0.3119×10^{-5}	0.93058
SD (10^{-3})	14.41	41.69	14.77	14.30	3.302	1.771×10^{-6}	6.435×10^{-4}	0.0151
Series 2								
Mean	-0.0078	-0.1756	-0.0049	-0.0763	0.0319	-2.625×10^{-8}	-0.2720×10^{-5}	0.93216
SD (10^{-3})	18.95	16.61	7.069	20.00	1.698	1.344×10^{-6}	6.726×10^{-4}	0.0125

The results for the first sequence from table 4.2 are found in table 5.7. The standard deviation shows that the method again converges consistently for all parameters. The values actually show lower deviations than the ones found for the noisy simulated data. They are also very similar between the two series, the only major difference is that the SD for p_2 show a much larger value for the first series due to a minor outlier in one of the runs. The mean values show that the method generates similar correction results for the two series. While the mean values for p_{10} and p_{11} might seem quite different with the latter actually showing different sign, the combined effect of them is very close. The transformation parameters that are generated for the mean values show that the shearing and translation are almost identical. Larger difference is present for scaling, where p_4 is almost twice as large for series 1, causing about twice as much scaling of the images. The transformation parameters for the separate sequences can be studied in figure 5.7. The plots show the large correlation between the results for the two different series, but also that the first series result in a large correction change. There is a significant time dependent translation present, reaching a maximum change of 0.94 pixels for series 1. The gradient dependent translation is much smaller with a maximum change of 0.08 pixels. Shearing and scaling effects are not as large as the total translation, peaking at 0.22 pixels and 0.21 pixels, respectively. All these values are the maximum change found in the central 80×80 part of the images. The largest correction difference for a single image between the two series is 0.19. Since the largest combined pixel movement for an

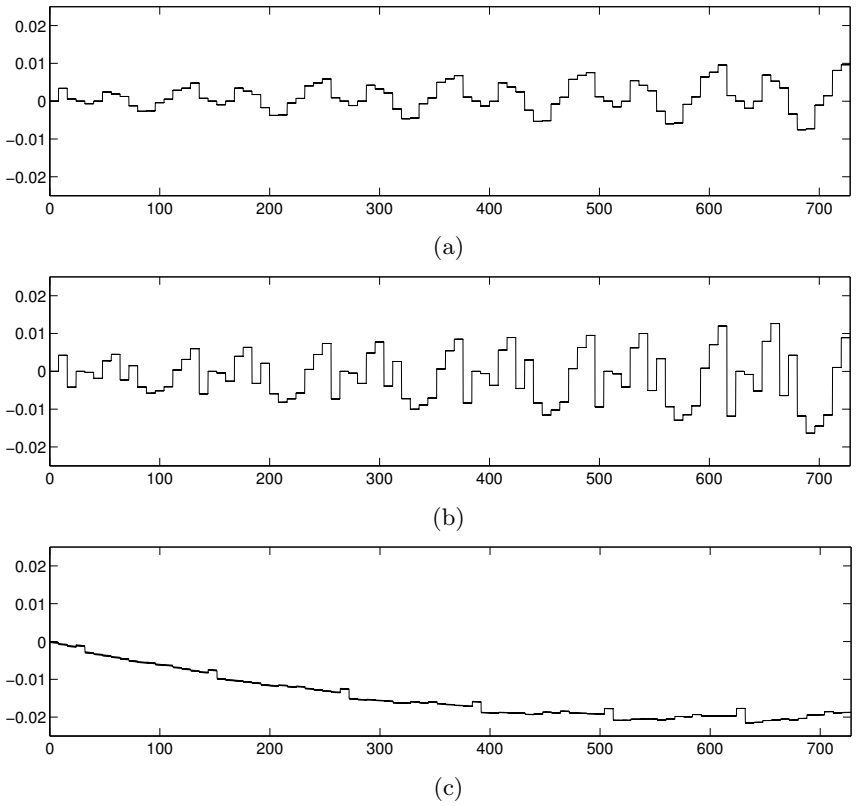
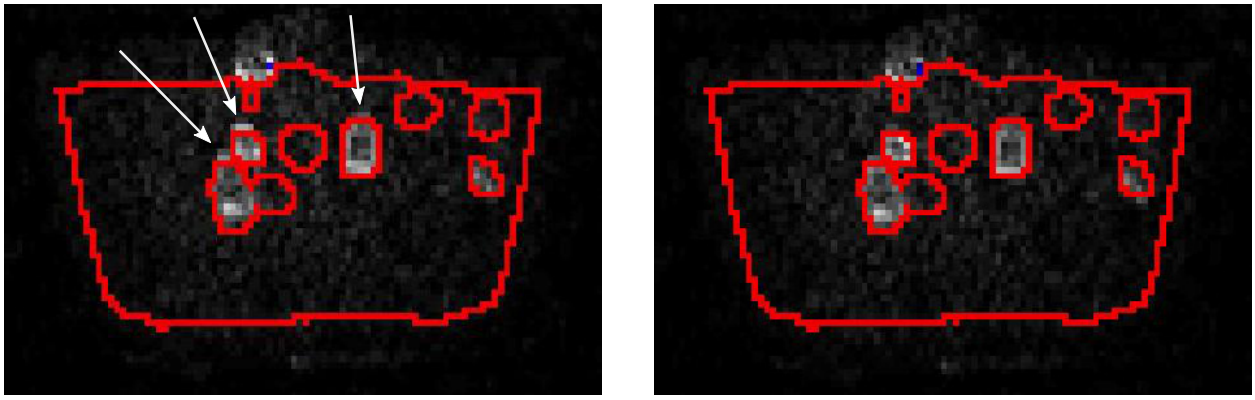


Figure 5.5: *Shearing (a), scaling (b) and translation parameters (c) generated from the best values found in run 5 on the noisy simulated data set. The values are shown as a function of time in seconds. The translation parameters are dominated by the time dependent translation that grows with time. The two other parameters also grow with time, but in this case because the b-value and hence the gradients become larger. The plot clearly shows how the sequence cycles over the fifteen parameters. The horizontal blocks where the parameter stays the same occurs when each slice in sequence is imaged using the same gradient setting. The maximum shearing value of about 0.01 gives a largest pixel movement of 0.40 in the central* 80×80 *pixels area of the image where the subject is located. The largest scaling at -0.015 creates a maximum movement of 0.58 pixels in the same area. The translation is in percent of the image size (128 pixels) and gives a largest movement of 2.56 pixels.*

image is 0.90 pixels, the difference in result for the two series is about 20%.

To determine if these transformations actually result in better alignment of the image set, visual inspection and signal reduction was used as for the simulated data. Figure 5.6 shows one image from the first series, before and after correction has been added to the image. When the edge found from the non-diffusion weighted image from the same slice is overlaid, the correction shows better alignment with the $b = 0$ image for most of the spots were signal is present. The method using image masks and change in signal level outside the mask was used with the threshold set to $S = 25$. The first series gave a reduction of signal outside of the mask by 2.27% - 2.47%. While this may seem like a very small change, this is due to the largest part of the signal existing in the asparagus stems that are all inside of the mask before and after correction is applied. The reduction of signal outside the mask still indicates an improvement of the image alignment. For the second series the reduction was 1.63% - 1.90%. That this reduction is smaller than for the first series show that the larger corrections found for series 1 does in fact result in a better image alignment than for series 2.

While the two series from sequence 1 generated very similar correction parameters, the situation is completely different for sequence 2. The aggregated results from five runs on each series are found in table 5.8. Both series converge consistently with low standard deviation, except for p_5 in series 1 where run 4 is an outlier. They do however converge to very different results. The shearing parameters are similar for both series, but series 2 results in three times as much scaling as series 1. The huge difference occurs for the time dependent translation, where the two series show completely different trajectories, illustrated in figure 5.8. Series 2 result in the same kind of translation as for the previous sequence, but correction on series 1 has resulted in a positive translation in a parabola, reaching its maximum in the middle of the sequence. To determine if either correction result in



(a) (b)

Figure 5.6: *Image from the first phantom sequence, captured at $b = 3000$, before (a) and after (b) correction from the first run is applied. The red border shows the outline of the non-diffusion weighted image of the same slice, thresholded at $S = 200$. The white arrows point to areas where visual inspection indicates that the image after correction is better lined up with the $b = 0$ image than before. The correction effect on this particular image was $\alpha = 0.0035$, $\beta = -0.0026$ and $\gamma = -0.0053$, giving a pixel movement that varies between -0.19 and -0.67 pixels depending on the position in the image. While the overall correction does seem favorable with less signal outside of the true location of the subject, it should also be noted that the outline of the asparagus stem on the far right does seem to line up slightly worse after correction.*

any improvement, the most reasonable approach is to once again look at the reduction of signal outside of the true position of the phantom. Series 1 does not show any change at all, the runs actually increase the signal outside the image mask slightly, with 0.00% - 0.14%. Series 2 instead shows a decent change of the signal, reducing it by 3.81% - 4.03%. It should however be noticed that while series 1 did not reduce the signal outside the phantom, there is also no visual indication that the data is misaligned before correction. The non-diffusion weighted images from the two series overlap as expected, showing that the phantom was not moved between the capture of the two data series.

The correction method converged consistently for all the four different image series that were used for the evaluation of the phantom measurements. It did so using slightly fewer iterations that for the noisy simulated data, using an average of 97.7 iterations. Another parameter that indicates easier convergence is the large value for local correlation for all the data sets. The average is more than 0.93 for the phantom data, which can be compared to a mean value of about 0.27 for the simulated data with added noise. This shows that the real images from the phantom do in fact have much less random noise than the amount that was added to the simulated images.

Table 5.8: *Results for running the correction method five times on two separate data series for the second sequence on the asparagus phantom. The two series result in very different correction parameters, with the second series being consistent with previous results while the first series show completely different values, especially for the time dependent translation.*

Sequence 2	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	**LC**
Series 1								
Mean	-0.0229	0.0488	0.0347	0.0372	-0.0010	-9.520×10^{-8}	4.084×10^{-5}	0.95184
SD (10^{-3})	4.157	5.175	10.59	3.856	0.8708	3.052×10^{-6}	9.761×10^{-4}	0.0200
Series 2								
Mean	-0.0351	0.0464	0.0106	0.1094	-0.0011	-0.501×10^{-8}	-1.758×10^{-5}	0.94969
SD (10^{-3})	6.266	7.855	6.899	7.177	0.1207	1.249×10^{-6}	4.470×10^{-4}	0.0055

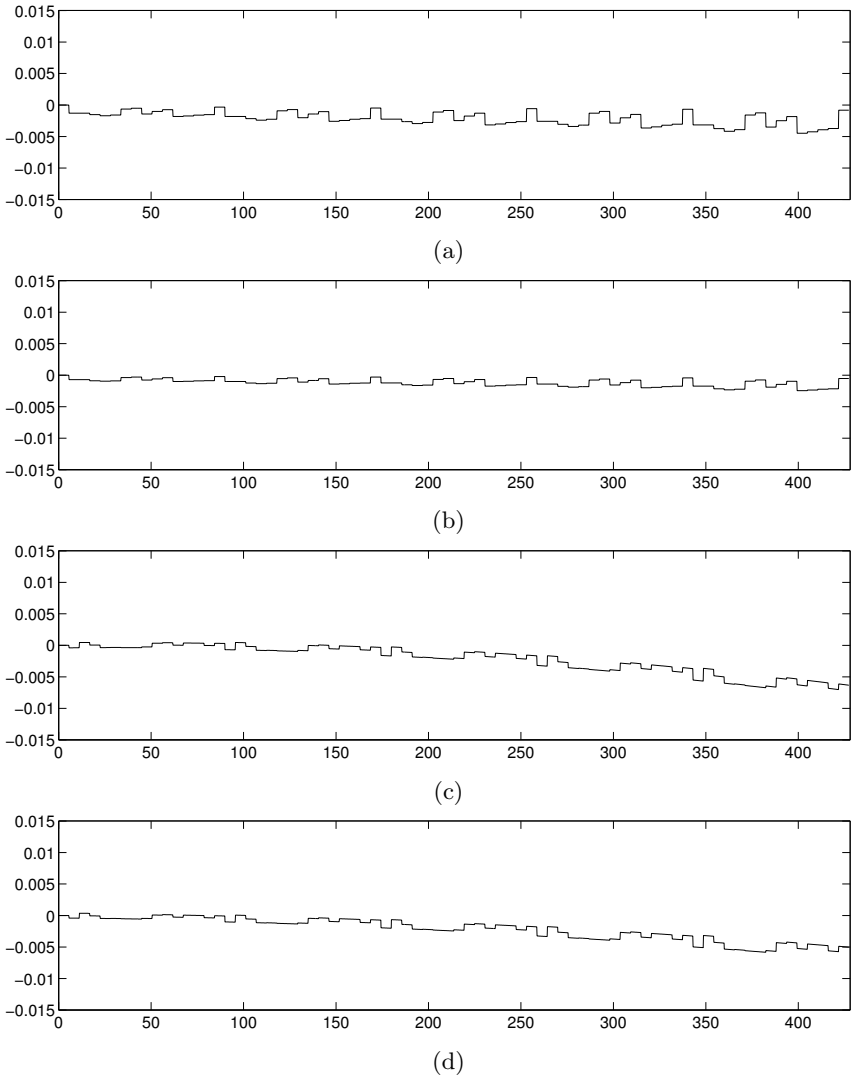


Figure 5.7: *Transformation parameters as a function of time (s) resulting from the mean parameters from the two series of sequence 1 for the asparagus phantom. (a) and (b) show scaling from the first and second series, while (c) and (d) show the translation. The plots show that the correction method on the two different series result in the same shape for the plotted parameters. The first series does however show almost twice as large scaling effect, and also a slightly larger translation than the second series. The shearing parameters are not shown here, but they show large correlation between the two series. The greatest difference arising when these transformations are applied to the images is 0.19 pixels. This should be compared to the largest pixel movement for any of the images in the two series, which is 0.90 pixels.*

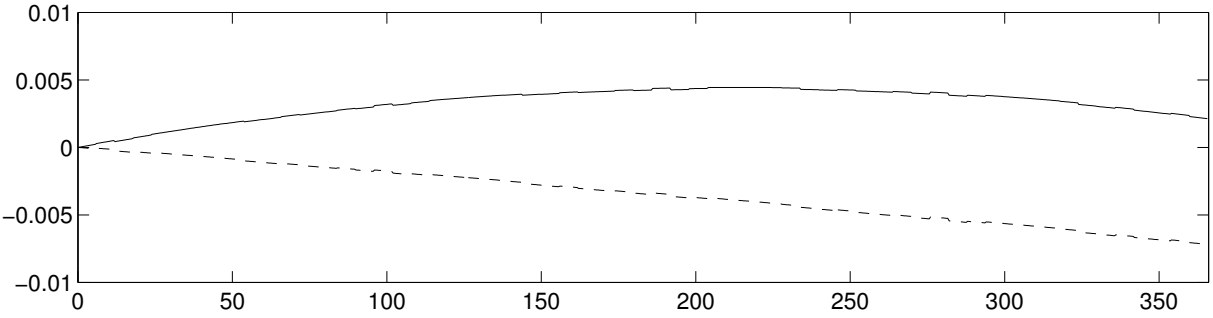


Figure 5.8: *Translation as a function of time in seconds for the second sequence protocol on the asparagus phantom. The solid line is series 1 and the dashed line is series 2. Both are captured using the exact same protocol and settings, and would not be exptected to differ. Series 2 shows the kind of translation that was the result from correction on sequence 1, while series 1 has resulted in a positive translation that reaches its maximum in the middle of the sequence. This goes against the theory for expected distortions. The largest translation for series 2 is -0.92 pixels and the same for the first series is 0.57 pixels.*

5.3 Results in Vivo

Results on real humans are provided from the three healthy volunteers. The correction algorithm was run five times each on two different series for each person. All the series were captured using the same sequence protocol. The results from the correction algorithm are collected in tables 5.9, 5.10 and 5.11. The algorithm converges in about 100 iterations for all the series. As in all the previous cases, it also converges consistently to the same solution for all runs made on the same series. For volunteer 1, both series result in parameters that cause quite large corrections to be applied to the images. Figure 5.10 show that this results in a good correction where the images align correctly with the true position of the subject. This is reflected in the percentage of signal removed outside of the brain: the reduction is between 1.70% and 3.56% for the ten runs. Note how the uncorrected image in figure 5.10 is clearly misaligned with the real position of the brain. The time dependent translation for both series on volunteer 1 behaves in the same way as for the phantom, resulting in a translation that grows with time until it finally reaches almost -2 pixels at the end of the sequence at $t = 427$ s.

Table 5.9: *Results for five correction runs on data from two different series on volunteer 1 using the same sequence protocol. The correction algorithm converge consistently for both series, and the mean parameters are very similar. Both series show a decent amount of shearing from p_2, and a quite large scaling effect in p_5. While the values for p_{10} and p_{11} might seem very different for the two series, their combined effect is in fact very similar, both resulting in a negative translation that reaches a minimum value of about -0.015 (-2 pixels in the image) at the end of the series.*

Volunteer 1	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	**LC**
Series 1								
Mean	-0.0295	0.1255	-0.0841	0.2174	0.0017	-5.925×10^{-8}	-0.825×10^{-5}	0.6574
SD (10^{-3})	5.239	5.397	4.714	2.763	0.868	0.527×10^{-6}	2.706×10^{-4}	0.0056
Series 2								
Mean	-0.0173	0.1099	-0.0804	0.2437	0.0025	0.663×10^{-8}	-3.595×10^{-5}	0.6635
SD (10^{-3})	8.579	1.669	2.024	3.635	1.341	2.252×10^{-6}	9.154×10^{-4}	0.0048

The situation is quite different for the two other volunteers. Figure 5.11 shows the correction for volunteer 3. The uncorrected image does not show any visually obvious misalignment. The correction applied by the solution only make minor adjustments, making it impossible to tell whether the correction results in any improvement to the image using visual inspection. Figure 5.12 illustrates how different the correction applied to volunteer 3 is to the one found for volunteer 1. The situation for volunteer 2 is the same as for volunteer 3. The biggest difference between the four data series for volunteer 2 and 3 and the two series for volunteer 1 is that the time

dependent translation behaves very differently. Figure 5.9 shows the translation for the six different series. Three different types of translation are seen. The expected translation that results in a negative adjustment is found, but there are also positive and negative parabolas that reach their minimum and maximum in the middle of the sequence. The extrema for the functions add a translation of about 0.6 pixels in either positive or negative direction. The same shape was seen for one of the phantom series as well, indicating that it is not due to any movement by the volunteer. While this type of correction was an unexpected result, it does not seem to negatively affect the image series. All the series still show good alignment with the true position of the brain, and result in a reduction of the signal outside the brain. This reduction varies between 0.72% and 1.61%. It is natural that the change is less than for volunteer 1, since none of the series for the other volunteers show as much distortion effect to correct. It should be noted that all of the series that lack the large negative translation also lack any large initial deviation from the true position, and thus seem to be much less affected by distortions overall.

Table 5.10: *Results for volunteer 2, for two different series using the same protocol as for volunteer 1. Both series converge consistently, and the results for the two series are very alike for shearing and scaling. The combined effect of p_1 and p_2 is a shearing that is about half of that found for volunteer 1. The case for scaling is the same, p_5 is dominant giving slightly less than half the scaling for volunteer 1. The time dependent translation varies between the two series. Neither gives a large effect, but they show different signs, one maximizing at plus 0.05, the other reaches a minimum of -0.05 instead.*

Volunteer 2	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	**LC**
Series 1								
Mean	-0.0875	0.0225	-0.0495	0.1046	-0.0006	-7.102×10^{-8}	3.563×10^{-5}	0.6627
SD (10^{-3})	5.321	3.068	2.984	3.628	0.5870	1.204×10^{-6}	3.740×10^{-4}	0.0093
Series 2								
Mean	-0.0696	0.0319	-0.0524	0.1050	-0.0043	3.892×10^{-8}	-2.617×10^{-5}	0.6672
SD (10^{-3})	4.211	3.167	1.517	3.975	0.6137	1.646×10^{-6}	5.894×10^{-4}	0.0071

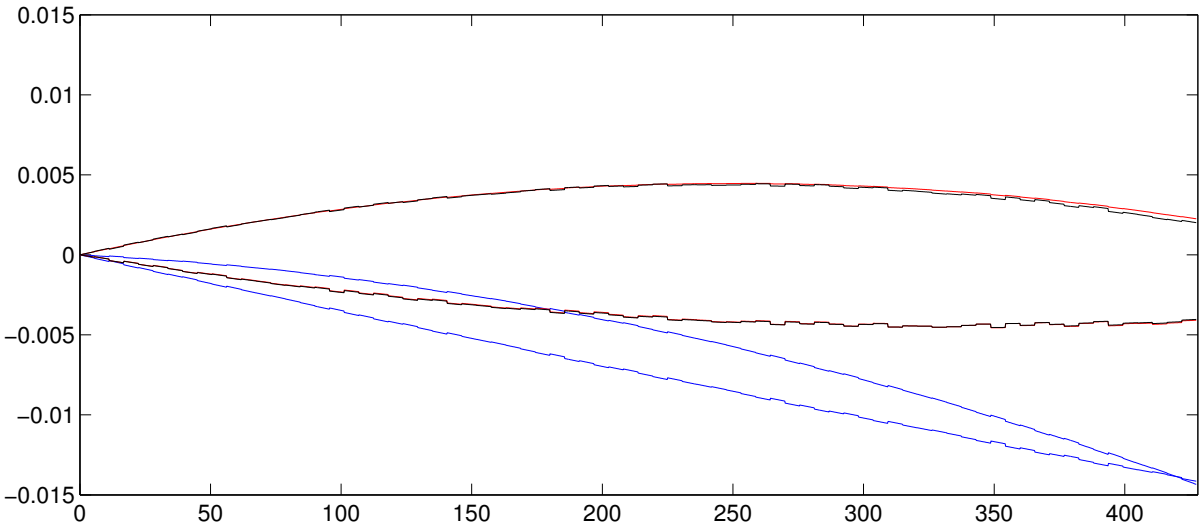
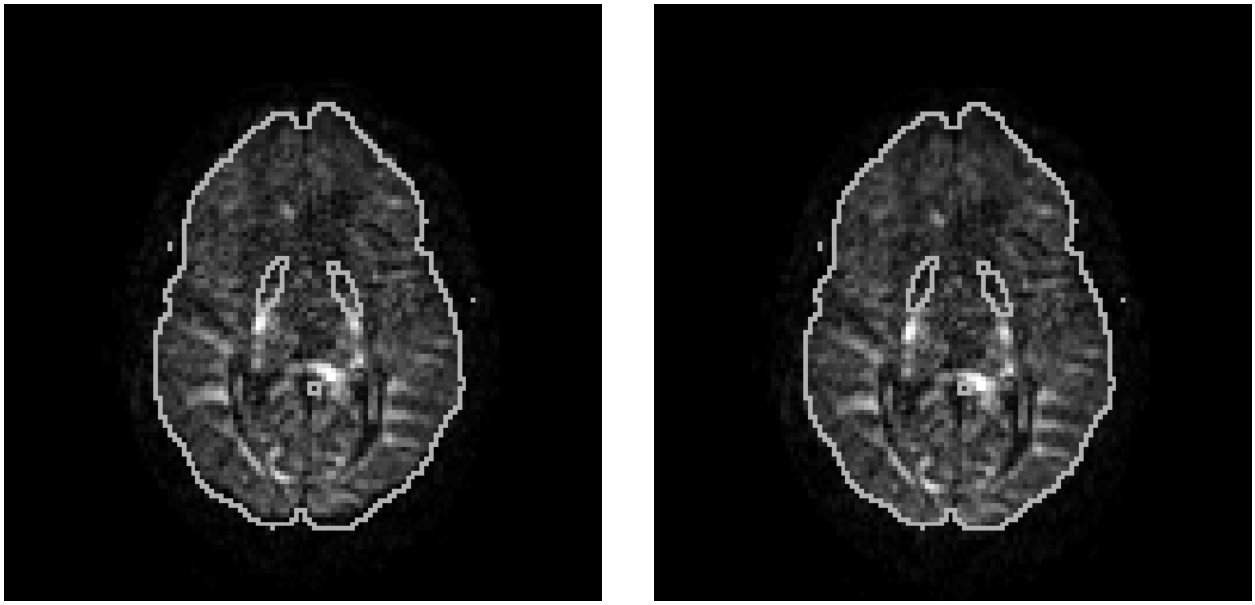


Figure 5.9: *Translation as a function of time in seconds for the six different data series on the volunteers. Blue is volunteer 1, red is volunteer 2 and black is volunteer 3. The two series on volunteer 1 are similar and behave as most series did for the phantom, ending in a negative translation of about -2 pixels for the final image. For the other two volunteers, the two series are distinctively different. Each volunteer has one series that results in a parabola having positive translation and one series with a trend that is almost the opposite. The shape of this translation is very similar to the one seen for the phantom in figure 5.8.*

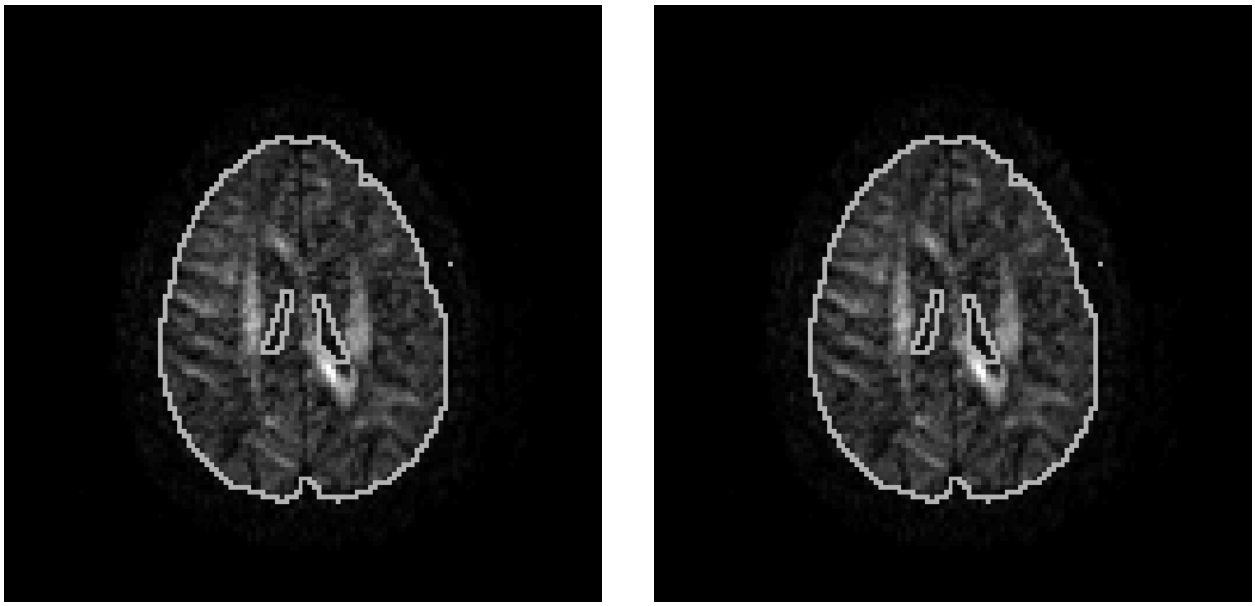
Table 5.11: *The results for the final volunteer show the same kind of consistency as the others, for both series. There is virtually no shearing at all present for either series, while the scaling is on about the same level as for volunteer 2, p_5 once more dominating. The time dependent translation once again differs between the two series, one showing a slight positive translation, the other a slight negative translation.*

Volunteer 3	p_1	p_2	p_4	p_5	p_9	p_{10}	p_{11}	**LC**
Series 1								
Mean	-0.0382	-0.0123	-0.0558	0.0799	-0.0028	-7.302×10^{-8}	3.593×10^{-5}	0.6624
SD (10^{-3})	1.119	2.726	1.979	2.751	0.7263	0.355×10^{-6}	1.212×10^{-4}	0.0029
Series 2								
Mean	-0.0229	0.0039	-0.0505	0.1008	-0.0039	4.023×10^{-8}	-2.661×10^{-5}	0.6618
SD (10^{-3})	2.4238	2.516	3.323	4.782	0.6904	0.289×10^{-6}	1.437×10^{-4}	0.0056



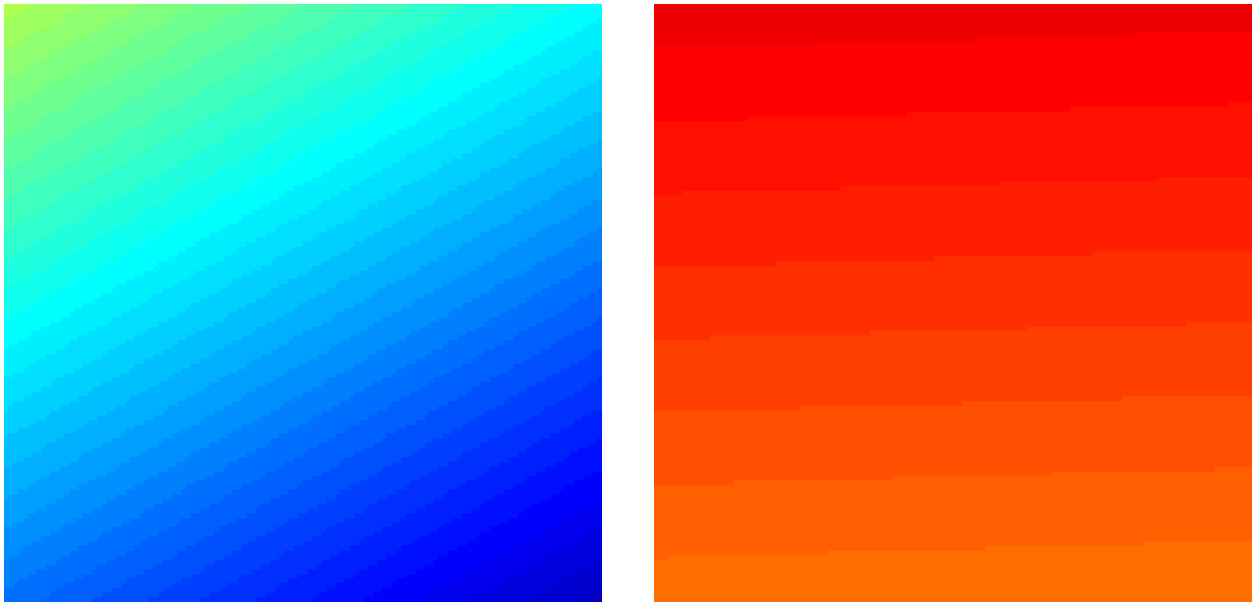
(a) (b)

Figure 5.10: *Image of volunteer 1, before (a) and after (b) correction is applied, with the real position of the brain overlaid. The correction applied is the average parameters found on series 1. The image is of the upper part of the brain and is taken using the largest b-value, $b = 3000$. The gray border is the detected position of the non-diffusion weighted image from the same slice. The borders that are seen inside the brain is the position of the ventricles. Before correction there is substantial misalignment of the image, especially visible at the bottom of the image where there is several pixels between the brain in the image and the real position. This is adjusted when the correction has been applied, the brain now aligning with the desired position. It is also apparent that the ventricles now align correctly. In the uncorrected image the areas of large signal in the brain partly overlaps with the ventricles, which is clearly wrong. The transformation parameters applied to the image are $\alpha = -0.0546$, $\beta = 0.0098$ and $\gamma = -0.0143$, resulting in an adjustment of the subject between -1.21 pixels and -2.41 pixels, depending on the position in the image. The small zones of border seen right in the center of the image and outside of the brain are due to a few pixels falling just below or above the threshold used to segment the true position.*



(a) (b)

Figure 5.11: *Uncorrected (a) and corrected (b) image, as in figure 5.10, but for series 1 on volunteer 3 instead. The situation here is completely different than for volunteer 1. There is no large misalignment in the uncorrected image, and hence no visually detectable improvement to the image after correction. The resulting transformation parameters are* $\alpha = -0.0001$, $\beta = 0.0040$ *and* $\gamma = 0.0021$. *The movement from this varies between 0.01 pixels and 0.33 pixels within the subject.*



(a) (b)

Figure 5.12: *The change applied to each pixel in the images corrected in figures 5.10 and 5.11, with the change for volunteer 1 in (a), and volunteer 3 in (b). The scale is from blue at -3 to red at 1. The image to the left shows much correction, ranging from -2.77 pixels in the bottom right corner to -0.85 pixels in the top left corner. Shearing is clearly present, shown by the way the stripes with the same pixel change are at an angle. The different colored stripes show that scaling is also present in the image. On the right, there is almost no shearing, and much less scaling. At the bottom the scaling almost exactly counters the positive translation, giving a total change of 0.01 pixels, while the change reaches 0.54 pixels at the top of the image.*

6 Discussion and Conclusions

6.1 Feasibility of the Method

This thesis work has resulted in a robust method that consistently improve the data quality of the image sets. The method converges to the same solution every time it is run on a data set, making sure that only one run is required to find the translation parameters that best improve the images.

The application of the method on the simulated data shows that the method is capable of finding a solution that is very close to the known distortions even in the presence of large amounts of noise. The corrections that were found resulted in parameters that are 10% - 20% from the perfect values. While not being identical to the true solution, this is still a huge improvement when values such as kurtosis are to be calculated using a full data set since it adjusts the images by more than 2.5 pixels. This kind of distortion would otherwise make the calculated values almost useless.

The asparagus phantom made it possible to estimate the real effects on a subject that is unaffected by movement. Two different sequence protocols were tested, with two different series tested for each protocol. All but one of the series showed obvious visual improvement of the alignment of the images. Those three series also showed that while shearing and scaling is present, the largest distortions were due to a time dependent translation caused by the heating of the scanner. The algorithm found that this translation slowly grew with time, which is as expected from literature. The fourth series did not find this type of translation, but did not show any initial misalignment either.

In vivo data from three volunteers show a similar behavior to the phantom data. The two series on one of the volunteers show large distortions, and very good alignment of the images after correction. The other four series do behave more like the fourth series on the phantom, having neither any big misalignment before correction, nor any large movement due to the correction. Running the correction algorithm on these six series result in a reduction of the amount of signal outside the true position of the brain for all the series. This is a good indication of the result of correction, since correction of the images should result in signal being moved from outside of the location of the brain in the non-diffusion weighted images to inside this area. It is however impossible to judge if the method has corrected for all distortions that are present. While there are possibly some misalignment left in the images, they are reduced from several pixels to sub pixel levels, which is a vast improvement.

While the biggest actual pixel movement arises from the time dependent translation that is due to $\mathbf{B}_0$ field drift, there is also quite large shearing and scaling effects detected in some data series. This shows that it is absolutely necessary to include the field drift translation, but that there are also eddy current effects present in the images. It does however seem to be almost zero effect from the gradient dependent translation (p_9). This strengthens the case to disregard the cross-terms between the x and y gradients and the translation effect, since this is as small as was expected.

Since the method is capable of detecting and correcting the large distortions that are present in some of the image series, the overall recommendation should be to run it on all data sets that are to be used for calculating kurtosis, DTI, or other values that are dependent on the images aligning correctly. The method has never showed any tendency to result in worse alignment of the data set, which means that it is safe to use even when no distortions are present.

6.2 Distortion Characteristics

When the work on this thesis started, the perception was that eddy currents were the source of the large distortions present in MRI diffusion images, while $\mathbf{B}_0$ field distortions might be a minor factor. The results in this thesis instead supports the opposite situation: the translation due to field drift is much larger than the eddy current effects. This translation is easily detected by visual inspection of the images, and is clearly a factor that needs to be corrected for. The results do also support that eddy currents effects are present, but these do contribute with less than a pixel of movement in the images, compared to a distortion of several pixels for the time dependent translation. One reason that the eddy current effect is not larger might be built-in correction software in the MRI scanner. This could explain why earlier diffusion data seemed to be much more affected by eddy currents. The large distortion caused by the $\mathbf{B}_0$ field drift is probably caused by the fact that the kurtosis sequences push the scanner to its limits, causing more drift than has been noticed in more

moderate diffusion sequences, such as the ones used for DTI.

Another idea that can be dismissed by the results is that the distortions would be completely dependent on the sequence protocol, regardless of whether a phantom or a patient was the subject. If this was the case, it would be possible to use correction results from a single reference measurement to correct future image sets captured using the same protocol. The idea with capturing several series on two different sequences for the phantom was to determine whether this was the case or not. The results for both phantom and volunteers clearly show that the distortions vary between series captured with the same protocol. Some series show a large distortion of several pixels, as in figure 5.9, while others show no apparent distortions at all (figure 5.11). The largest variation is the time dependent translation that is due to the $\mathbf{B}_0$ field drift, which is easily detected by ocular inspection. The results also show large variation in scaling and shearing, but whether this is due to actual differences or to other unknown factors that affect the convergence is hard to tell. The notion that reference corrections could be created by running the algorithm on data from a phantom can be completely dismissed. Instead, the algorithm will have to be performed separately for each data series. While this takes about four hours using the hardware that was available for this thesis, it could easily become much faster since the evaluations are parallelizable. A current eight core computer should be able to run the process until convergence in about one hour.

That both the eddy current effect and the time dependent translation varies as much as it seems to do is an effect that is not accounted for in the available literature on the subject. Since there is not much data available on the behavior of high b-value diffusion gradients, real field measurements should ideally be performed to find out how they actually behave. The model used here assume linear eddy current effects in accordance with existing theory. The large deviations found using the correction algorithm creates some doubt regarding whether they are linear. Measurements of the actual effects would add valuable knowledge regarding the requirements of correction. Measurements of the $\mathbf{B}_0$ field drift should most definitely be made since this is the single largest source of distortions in the images, and also seems to differ a lot between data series. The method currently assumes that the time dependent translation returns to zero quickly when each data series acquisition is finished. This has later been shown in Lund to be false, and the method should be expanded to take this into account.

A large factor in the puzzle when it comes to distortions is the MRI scanner itself. Each manufacturer has their own software and hardware, and such complex machines can also be expected to vary somewhat in between scanners. The amount of correction that is built into the software is an unknown factor and might turn out to be completely different for another manufacturer. Since the results here are based on data from a single Philips 3T MRI scanner, this might be a factor. The data acquisition for this thesis was made over a period of six months, which means that differences in calibration over time is another possible source of distortions. Image series from different scanners should be used as input for the algorithm, and preferable also data from scanners from other manufacturers. This would add valuable information regarding the scope that the method can handle.

6.3 Further Analysis of Results

While the correction method does improve the data quality of the image series after correction, we don't actually know how much distortion remains after the correction has been applied. Visual inspection can only determine that most of the distortions have been accounted for. Sub pixel shearing and scaling distortions are clearly beyond the human eye to measure. The kurtosis data sets consist of between 1000 and 2500 images each, which creates a requirement for computational methods to determine the distortion effects that are not accounted for. The results generated here are based on the study of a handful of images per data set, and while the signal reduction used here is a good indication that the method makes a difference, it tells us nothing about any remaining distortions. The model is based on existing theory, but could there be other factors present? The measurements mentioned above could go a long way to answering this.

Further analysis of the results could also be performed by comparing the method to some of the existing methods discussed in section 3.2.4. While these are not capable of correcting distortions in the high b-value series used here, they could be compared against results for our method on image series that use b-values up to about 1000. If our method works well it should return comparable results to the existing methods, although it is not known how well it works with fewer b-values available.

Finally, the analysis of the results could benefit from creating some sort of computational measurement of the amount of registration error that remains in the image series after correction has been performed. The

article by Zitová and Flusser [60] discussed various ways that such a measurement could be created. This kind of analysis was however considered to be outside the scope of this thesis.

6.4 Improving the Method

The results show that the method is far from perfect when it comes to finding the true optimization solution. The method was 10% - 20% off from the true solution when noise was added to the simulated images. The optimization algorithm found the solution that maximized the fitness function, but this was not the true solution. The real solution did show a lower fitness function, proving that it was not the optimization that went wrong. The problem was that the fitness function reaches its maximum at a point that is close to, but not at the true solution. The calculation of the reduction of signal outside the brain showed that the real solution resulted in larger reduction than the found solution, clearly proving that it was in fact the superior correction. Thus, the fitness function needs some tweaking to further improve the results. The required change might be to introduce a different pattern of comparisons than the one shown in figure 4.3, or to introduce different weighting for the different comparisons that are made. Another possibility is that further measurements must be added to the fitness function. For example, the reduction of signal outside the subject might be a suitable value to include.

Another aspect of the method is that it can be adjusted by many different parameters. There are parameters controlling how the values are distributed at the start, and several different parameters control how the optimization process performs. Finding the optimum for each of these parameters has not been possible during the time that was available for this work. Since five runs on a data series took about 20 hours to run until convergence, only a small set of possible parameters have been tested. Finding the perfect balance for this process might be the key to being able to drop the reduction of parameters used here. The full method using twelve parameters would be much preferable since this would allow cross-terms between any of the gradient directions. While the contributions from and to the z gradient are generally smaller, including these could improve the possibility of finding small corrections that might remain now.

The method currently assumes that the eddy current decay completely in a slice before the next image acquisition takes place (TR in the sequence specifications). This might not be one hundred percent true, and the method could benefit by adding a parameter that adjusts the decay rate. This could either be set based on measurements, or could be allowed to be found by the method itself. The latter would however add at least one more parameter to be found, increasing the demands on the method.

6.5 Handling Movement Correction

The method as it currently stands does not include any correction of the involuntary motion of a patient during the image acquisition. The typical TR between two images of the same slice is in the range 5 - 10 s. The time between two images with the same gradient direction but different b-values that are compared is even greater, often 30 seconds or more. Thus it is quite likely that a patient would move the head a few millimeters between the capture of two images that are compared to each other, causing the method to fail since such motion is completely different from the distortions that the model allows.

Given those circumstances it is apparent that the method must take patient movement into account, especially since patients probably move much more than healthy volunteers. There are several established ways that motion could be detected and corrected for, since this is a problem that is common to many other fields of MRI as well. While several eddy current correction methods try to correct for both the eddy current effect and the motion and the same time, this is mainly done by allowing any kind of spatial transform that improves the image registration, opening up for the possibility that completely non-physical corrections are performed. For the method presented here it makes more sense to try to measure the patient motion and correct for it before the eddy current correction is applied. Measurement of movements could be performed either by an external camera or by inserting so called navigational echoes into the image sequence. Navigational echoes capture a low resolution volume of the head that is enough to determine if any movement has occurred. The results from either method would then be used to correct for the movements. It is very important to rotate the applied gradients when this is done, since the method rely on the exact direction of the gradients to find the correction parameters. How and why this should be performed is discussed in detail in the article by Leemans and Jones [61]. Applying a movement correction in this way unfortunately has the effect of changing the correction method from a pure post-processing method by requiring either changes to the sequence protocol

or special sensor equipment to the used during the acquisition. While this somewhat defeats the purpose outlined in the introduction, it should be noted that the method is still capable of correction without motion correction, as shown in this thesis. A post-processing method has recently been proposed by Ben-Amitay et al. that could solve the movement problem, discussed in the following section.

Whether the results on the volunteers are affected by involuntarily motion or not is very hard to tell. It does however seem unlikely that the large translation seen for volunteer 1 and the lack of the same for the other two are the result of movement since both effects have been noticed for the motionless phantom as well. It could however have an influence on the results, perhaps explaining why the results differ as much as they do for the different volunteers. The translations that are positive for one series and negative for the next in figure 5.9 could possibly be an effect of such motion. It would be very interesting to perform studies were volunteers perform large deliberate motions to find out how this would affect the method.

6.6 Recent Studies on High b-Value Diffusion

When the work on this thesis started back in 2010, there was no literature on the subject of eddy current correction in high b-value diffusion MRI. Since then, a few different articles have been published that propose methods that could be used to counter these effects.

Ben-Amitay et al. published an exciting method in 2011 that could correct for motion in a post-processing step [62]. The basic idea is to start by running a DTI scan and calculate the diffusion tensor for the subject. This is a step that is well known and have established correction methods for both eddy currents and patient movement. They next create a simulated data set from this that shows how the high b-value data might look. Finally they take the real high b-value images and register them to the simulation, thus canceling any patient motion that might be present. This seems to work very well, and could potentially be used as a pre-processing step to the method outlined here, or to completely replace it.

Another article, by Zhuang et al. from 2012 proposes a method for eddy current distortion correction that is somewhat similar to the method in this thesis [63]. Instead of comparing diffusion weighted images with different b-values, they co-register diffusion images with the same b-value and similar gradient directions, and do not use any comparison to the non-diffusion image of the same slice. The reduction of distortions is indicated by longer fiber tracking (as in figure 1.3) being possible after correction is applied. While this method seems to reduce distortions, it also shares some of the shortages in older studies, especially allowing for non-physical transforms of the images.

6.7 Conclusions

The method that has been developed during this thesis work is, while not perfect, capable of finding the visually apparent distortions in high b-value kurtosis series. It is robust and converges to the same solution every time on a given image series. The distortions have been shown to differ a lot between series using the same sequence protocol, which requires the method to be run separately on each image series. The method can find distortions and does not show any tendency to negatively affect data with wrongful transformations. It should thus be run on all high b-value image series, even if this takes a few hours per data set with ordinary computer capability, since the benefit clearly outweighs the negative aspects. The biggest distortion effect in the kurtosis series come from B_0 field drift, meaning that this effect should most definitely not be left uncorrected.

The level of distortions left in the images after correction is unknown and further work is required to determine this and improve the method. Some kind of movement correction strategy should also be implemented to make this a fully fledged method to correct for all distortions that are present in high b-value diffusion images.

References

[1] D. W. McRobbie et al. *MRI from Proton to Picture, 2nd edition.* Cambridge University Press, 2006.

[2] R. Brown. "A Brief Account of Microscopical Observations Made in the Months Of June, July And August, 1827, on the Particles Contained in the Pollen of Plants; And on the General Existence of Active Molecules in Organic and Inorganic Bodies." In: *Philosophical Magazine* 4 (1828), pp. 161–173.

[3] A. Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen." In: *Annalen der Physik* 17 (1905), pp. 549–560.

[4] K. Tanaka. "Self-diffusion Coefficients of Water in Pure Water and in Aqueous Solutions of Several Electrolytes With ^{18}o and ^{2}h as Tracers". In: *Journal of the Chemical Society, Faraday Transactions 1* 74 (1978), pp. 1879–1881.

[5] M. E. Moseley et al. "Early Detection of Regional Cerebral Ischemia in Cats: Comparison of Diffusion and T2-weighted MRI and Spectroscopy." In: *Magnetic Resonance Imaging* 14 (1990), pp. 330–346.

[6] P. J. Basser et al. "MR Diffusion Tensor Spectroscopy and Imaging". In: *Biophysical Journal* 66 (1994), pp. 259–267.

[7] D. Topgaard and O. Söderman. "Experimental Determination of Pore Shape and Size Using Q-space NMR Microscopy in the Long Diffusion-time Limit". In: *Magnetic Resonance Imaging* 21 (2003), pp. 69–76.

[8] J. H. Jensen et al. "Diffusional Kurtosis Imaging: The Quantification of Non-Gaussian Water Diffusion by Means of Magnetic Resonance Imaging". In: *Magnetic Resonance in Medicine* 53 (2005), pp. 1432–1440.

[9] J. Lätt et al. "In Vivo Visualization of Displacement-Distribution-Derived Parameters in Q-space Imaging". In: *Magnetic Resonance Imaging* 26 (2008), pp. 77–87.

[10] M. Nilsson et al. "On the Effects of a Varied Diffusion Time in Vivo: Is the Diffusion in White Matter Restricted?" In: *Magnetic Resonance Imaging* 27 (2009), pp. 176–187.

[11] L. G. Hanson. "Is Quantum Mechanics Necessary for Understanding Magnetic Resonance?" In: *Concepts in Magnetic Resonance Part A, Vol. 32A* 5 (2008), pp. 329–340.

[12] E. M. Haacke et al. *Magnetic Resonance Imaging - Physical Principles and Sequence Design.* Wiley-Liss, 1999.

[13] F. Bloch. "Nuclear Induction". In: *Physical Review* 70 (1946), pp. 460–473.

[14] D. K. Cheng. *Fundamentals of Engineering Electromagnetics.* Addison-Wesley, 1993.

[15] H. Y. Carr and E. M. Purcell. "Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments". In: *Physical Review* 94 (1954), pp. 630–638.

[16] G. B. Folland. *Fourier Analysis and its Applications.* Brooks/Cole, 1992.

[17] S. Ljunggren. "A Simple Graphical Representation of Fourier-Based Imaging Methods". In: *Journal of Magnetic Resonance* 54 (1983), pp. 338–343.

[18] D. B. Twieg. "The k-trajectory Formulation of the NMR Imaging Process With Applications in Analysis and Synthesis of Imaging Methods". In: *Medical Physics* 10(5) (1983), pp. 610–621.

[19] C. B. Ahn, J. H. Kim, and Z. H. Cho. "High-Speed Spiral-Scan Echo Planar NMR Imaging-I". In: *IEEE Transactions on Medical Imaging* 5(1) (1986), pp. 2–7.

[20] J. G. Pipe. "Motion Correction With PROPELLER MRI: Application to Head Motion and Free-Breathing Cardiac Imaging". In: *Magnetic Resonance in Medicine* 42 (1999), pp. 963–969.

[21] J. Philibert. "One and a Half Century of Diffusion: Fick, Einstein, Before and Beyond". In: *Diffusion Fundamentals* 2 (2005), pp. 1.1–1.10.

[22] A. Fick. "On Liquid Diffusion". In: *Philosophical Magazine* 10 (1855), pp. 30–39.

[23] O. E. Stejskal and J. E. Tanner. "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient". In: *Jornal of Chemical Physics* 42 (1965), pp. 288–292.

[24] S. Mori. *Introduction to Diffusion Tensor Imaging.* Elsevier, 2007.

[25] R. V. Mulkern et al. "Biexponential Apparent Diffusion Coefficient Parametrization in Adult vs Newborn Brain". In: *Magnetic Resonance Imaging* 19 (2001), pp. 659–668.

[26] M. M. Cheung et al. "Does Diffusion Kurtosis Imaging Lead to Better Neural Tissue Characterization? A Rodent Brain Maturation Study". In: *Neuroimage* 45 (2009), pp. 386–392.

[27] J. Lätt et al. "Accuracy of Q-space Related Parameters in MRI: Simulations and Phantom Measurements". In: *IEEE Transaction on Medical Imaging* 26 (2007).

[28] M. A. Bernstein, K. F. King, and X. J. Zhou. *Handbook of MRI Pulse Sequences.* Academic Press, 2004.

[29] P. Jehenson, M. Westphal, and N. Schuff. “Analytical Method for the Compensation of Eddy-Current Effects Induced by Pulsed Magnetic Field Gradients in NMR Systems”. In: *Journal of Magnetic Resonance* 90 (1990), pp. 264–278.

[30] P. Jezzard and R. S. Balaban. “Correction for Geometric Distortion in Echo Planar Images from B_0 Field Variations”. In: *Magnetic Resonance in Medicine* 34 (1995), pp. 65–73.

[31] P. Jezzard, A. S. Barnett, and C. Pierpaoli. “Characterization of and Correction for Eddy Current Artifacts in Echo Planar Diffusion Imaging”. In: *Magnetic Resonance in Medicine* 39 (1998), pp. 801–812.

[32] D. L. Bihan et al. “Artifacts and Pitfalls in Diffusion MRI”. In: *Journal of Magnetic Resonance Imaging* 24 (2006), pp. 478–488.

[33] P Mansfield and B Chapman. “Active Magnetic Screening of Coils for Static and Time-Dependent Magnetic Field Generation in NMR Imaging”. In: *Journal of Physics E: Scientific Instruments* 19 (1986), pp. 540–545.

[34] C. B. Ahn and Z. H. Cho. “Analysis of the Eddy-Current Induced Artifacts and the Temporal Compensation in Nuclear Magnetic Resonance Imaging”. In: *IEEE Transaction on Medical Imaging* 10 (1991), pp. 47–52.

[35] C Boesch, R Gruetter, and E Martin. “Temporal and Spatial Analysis of Fields Generated by Eddy Currents in Superconducting Magnets: Optimization of Corrections and Quantitative Characterization of Magnet/Gradient Systems”. In: *Magnetic Resonance in Medicine* 20 (1991), pp. 268–284.

[36] J. C. Haselgrove and J. R. Moore. “Correction for Distortion of Echo-Planar Images Used to Calculate the Apparent Diffusion Coefficient”. In: *Magnetic Resonance in Medicine* 36 (1996), pp. 960–964.

[37] M. E. Bastin. “Correction of Eddy Current-Induced Artefacts In Diffusion Tensor Imaging Using Iterative Cross-Correlation”. In: *Magnetic Resonance Imaging* 17 (1999), pp. 1011–1024.

[38] M. E. Bastin. “On the Use of the Flair Technique to Improve the Correction of Eddy Current Induced Artefacts in MR Diffusion Tensor Imaging”. In: *Magnetic Resonance Imaging* 19 (1999), pp. 937–950.

[39] J.-F. Mangin et al. “Distortion Correction and Robust Tensor Estimation for MR Diffusion Imaging”. In: *Medical Image Analysis* 6 (2002), pp. 191–198.

[40] N. Bodammer et al. “Eddy Current Correction in Diffusion-Weighted Imaging Using Pairs of Images Acquired With Opposite Diffusion Gradient Polarity”. In: *Magnetic Resonance in Medicine* 51 (2004), pp. 188–193.

[41] Y. Shen et al. “Correction of High-Order Eddy Current Induced Geometric Distortion in Diffusion-Weighted Echo-Planar Images”. In: *Magnetic Resonance in Medicine* 52 (2004), pp. 1184–1189.

[42] M. A. Horsfield. “Mapping Eddy Current Induced Fields for the Correction of Diffusion-Weighted Echo Planar Images”. In: *Magnetic Resonance Imaging* 17 (1999), pp. 1335–1345.

[43] M. Jenkinson and S. Smith. “A Global Optimisation Method for Robust Affine Registration of Brain Images”. In: *Medical Image Analysis* 5 (2001), pp. 143–156.

[44] M. Jenkinson et al. “Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images”. In: *Neuroimage* 17 (2002), pp. 825–841.

[45] S. M. Maniega, M. E. Bastin, and P. A. Armitage. “A Quantitative Comparison of Two Methods to Correct Eddy Current-Induced Distortions in DT-MRI”. In: *Magnetic Resonance Imaging* 25 (2007), pp. 341–349.

[46] J. L. R. Andersson and S. Skare. “A Model-Based Method for Retrospective Correction of Geometric Distortions in Diffusion-Weighted EPI”. In: *Neuroimage* 16 (2002), pp. 177–199.

[47] G. K. Rohde et al. “Comprehensive Approach for Correction of Motion and Distortion in Diffusion-Weighted MRI”. In: *Magnetic Resonance in Medicine* 51 (2004), pp. 103–114.

[48] A. L. Alexander, J. S. Tsuruda, and D. L. Parker. “Elimination of Eddy Current Artifacts in Diffusion-Weighted Echo-Planar Images: The Use of Bipolar Gradients”. In: *Magnetic Resonance in Medicine* 38 (1997), pp. 1061–1021.

[49] B. Chen, H. Guo, and A. W. Song. “Correction for Direction-Dependent Distortions in Diffusion Tensor Imaging Using Matched Magnetic Field Maps”. In: *Neuroimage* 30 (2006), pp. 121–129.

[50] T. G. Reese et al. “Reduction of Eddy-Current-Induced Distortion in Diffusion MRI Using a Twice-Refocused Spin Echo”. In: *Magnetic Resonance in Medicine* 49 (2003), pp. 177–182.

[51] B. U. Foerster, D. Tomasi, and E. C. Caparelli. “Magnetic Field Shift due to Mechanical Vibration in Functional Magnetic Resonance Imaging”. In: *Magnetic Resonance in Medicine* 54 (2005), pp. 1261–1267.

[52] W. M. W. III et al. “Multi-Modal Volume Registration by Maximization of Mutual Information”. In: *Medical Image Analysis* 1 (1996), pp. 35–51.

[53] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. “Mutual Information Based Registration of Medical Images: A Survey”. In: *IEEE Transactions on Medical Imaging* 22 (2003), pp. 986–1004.
[54] T. Netsch et al. “Towards Real-Time Multi-Modality 3-D Medical Image Registration”. In: *Proceedings Eighth IEEE International Conference on Computer Vision ICCV 2001* 1 (2001), pp. 718–725.
[55] T. Netsch and A. van Muiswinkel. “Quantitative Evaluation of Image-Based Distortion Correction in Diffusion Tensor Imaging”. In: *IEEE Transactions on Medical Imaging* 23 (2004), pp. 789–798.
[56] M. Wahde. *Biologically Inspired Optimization Methods*. WIT Press, 2008.
[57] R. G. Keys. “Cubic Convolution Interpolation for Digital Image Processing”. In: *IEEE Transactions on Acoustics, Speech and Signal Processing* 29 (1981), pp. 1153–1160.
[58] M. Buehren. *Multicore - Parallel Processing on Multiple Cores (external plugin for MATLAB)*. 2013. URL: `http://www.mathworks.com/matlabcentral/fileexchange/13775-multicore-parallel-processing-on-multiple-cores` (visited on 2013-09-03).
[59] P. Mannfolk et al. “Assessment of Spatial BOLD Sensitivity Variations in fMRI Using Gradient-Echo Field Maps”. In: *Magnetic Resonance Imaging* 28 (2010), pp. 947–956.
[60] B. Zitová and J. Flusser. “Image Registration Methods: a Survey”. In: *Image and Vision Computing* 21 (2003), pp. 977–1000.
[61] A. Leemans and D. K. Jones. “The *B*-Matrix Must Be Rotated When Correcting for Subject Motion in DTI Data”. In: *Magnetic Resonance in Medicine* 61 (2009), pp. 1336–1349.
[62] S. Ben-Amitay, D. K. Jones, and Y. Assaf. “Motion Correction and Registration of High *b*-Value Diffusion Weighted Images”. In: *Magnetic Resonance in Medicine* 67 (2011), pp. 1694–1702.
[63] J. Zhuang et al. “Correction of Eddy Current Distortions in High Angular Resolution Diffusion Imaging”. In: *Journal of Magnetic Resonance Imaging* 37 (2012), pp. 1460–1467.