Separation of Water and Fat Signal in Magnetic Resonance Imaging

Advances in Methods Based on Chemical Shift

JOHAN BERGLUND
Dissertation presented at Uppsala University to be publicly examined in Hedstrandsalen, Entrance 70, Uppsala University Hospital, Uppsala. Friday, October 21, 2011 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

Magnetic resonance imaging (MRI) is one of the most important diagnostic tools of modern healthcare. The signal in medical MRI predominantly originates from water and fat molecules. Separation of the two components into water-only and fat-only images can improve diagnosis, and is the premier non-invasive method for measuring the amount and distribution of fatty tissue.

Fat-water imaging (FWI) enables fast fat/water separation by model-based estimation from chemical shift encoded data, such as multi-echo acquisitions. Qualitative FWI is sufficient for visual separation of the components, while quantitative FWI also offers reliable estimates of the fat percentage in each pixel. The major problems of current FWI methods are long acquisition times, long reconstruction times, and reconstruction errors that degrade image quality.

In this thesis, existing FWI methods were reviewed, and novel fully automatic methods were developed and evaluated, with a focus on fast 3D image reconstruction. All MRI data was acquired on standard clinical scanners.

A triple-echo qualitative FWI method was developed for the specific application of 3D whole-body imaging. The method was compared with two reference methods, and demonstrated superior image quality when evaluated in 39 volunteers.

The problem of qualitative FWI by dual-echo data with unconstrained echo times was solved, allowing faster and more flexible image acquisition than conventional FWI. Feasibility of the method was demonstrated in three volunteers and the noise performance was evaluated.

Further, a quantitative multi-echo FWI method was developed. The signal separation was based on discrete whole-image optimization. Fast 3D image reconstruction with few reconstruction errors was demonstrated by abdominal imaging of ten volunteers.

Lastly, a method was proposed for quantitative mapping of average fatty acid chain length and degree of saturation. The method was validated by imaging different oils, using gas-liquid chromatography (GLC) as the reference. The degree of saturation agreed well with GLC, and feasibility of the method was demonstrated in the thigh of a volunteer.

The developed methods have applications in clinical settings, and are already being used in several research projects, including studies of obesity, dietary intervention, and the metabolic syndrome.

Keywords: Magnetic resonance imaging, digital image reconstruction, chemical shift imaging, water and fat separation, Dixon method, fat suppression, quantitative MRI, whole-body MRI, fatty acid composition, fat unsaturation, triglycerides, adipose tissue, liver fat, T2* mapping

Johan Berglund, Uppsala University, Department of Radiology, Oncology and Radiation Science, Radiology, Akademiska sjukhuset, SE-751 85 Uppsala, Sweden.

© Johan Berglund 2011

ISSN 1651-6206 0346-5462
ISBN 978-91-554-8154-4
urn:nbn:se:uu:diva-158111 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-158111)
Nature is more subtle, more deeply intertwined and more strangely integrated than any of our pictures of her. It is not merely that our pictures are not full enough; each of our pictures in the end turns out to be so basically mistaken that the marvel is that it worked at all.

— Jacob Bronowski
This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


III **Berglund J** and Kullberg J. Three-dimensional water/fat separation and T₂* estimation based on whole-image optimization – application in breathhold liver imaging at 1.5 T. Accepted for publication in *Magnetic Resonance in Medicine* (2011).


This material is reproduced with permission of John Wiley & Sons, Inc.
Related Work

The author has also contributed to the following work:


Den teknologi, som möjliggör avbildning av människokroppens inre med hjälp av magnetkamera, kallas magnetisk resonanstomografi (MRT). En magnetkamera visas i fig. 3.3 på sidan 38. MRT har funnits sedan slutet på 1970-talet och har kommit att bli ett av de viktigaste verktygen för diagnostik inom modern sjukvård. En orsak till denna utveckling är att MRT inte är förknippat med samma risker som andra bildgivande metoder, såsom röntgen. En annan förklaring är metodens flexibilitet; en undersökning kan ge bilder med många olika typer av kontrast (fig. 3.1, sid. 36).

Bildsignalen i MRT alstras genom magnetisk kärnresonans, ett fenomen som kan observeras hos vissa atomkärnor då de placeras i ett magnetfält. I medicinska sammanhang nyttjas oftast magnetresonansen hos väte. En MRT-bild skapas i två steg; *datainsamling* i magnetkamerans starka magnetfält, samt *bildrekonstruktion* i en dator (fig. 2.12, sid. 31). Rekonstruktionen sker oftast direkt i magnetkamerans dator.

Vid avbildning av människokroppen härrör bildsignalen i huvudsak från väteatomer i vatten- och fettmolekyler. Signaler från vatten respektive fett har olika resonansfrekvens och kan därför separeras. Bildsignalen kan sedan delas upp i en vatten-bild och en fett-bild. Sådana bilder ger information om var fettet finns respektive inte finns, vilket är kliniskt intressant och kan underlätta diagnostik. Detta är dessutom den bästa metoden för att icke-invasivt
mäta mängd och fördelning av fettvävnad och kan även användas för att mäta fettinnehåll i inre organ.


Denna avhandling ger en sammanställning av tidigare FWI-metoder, samt beskriver nya helautomatiska metoder i fyra olika delarbeten. Alla de nya metoderna utvärderades genom att samlar in bilddata med vanliga kliniska magnetkameror. Ett genomgående fokus var snabb tredimensionell bildrekonstruktion.

**Delarbete I**

Denna artikel beskriver en metod där data samlas in med tre olika ekotider. Metoden utvecklades för den specifika tillämpningen snabb tredimensionell helkroppsavbildning. Sådana bilder är av stort värde vid mätning av kroppssammansättning, i synnerhet mängden fettvävnad och hur denna är fördelad i olika depåer (underhudsfett, bukfett o.s.v.).

Experiment utfördes genom att avbilda 39 frivilliga försökspersoner. Den beskrivna metoden gav snabb bildrekonstruktion och uppvisade bäst bildkvalitet i en jämförelse med två referensmetoder.

**Delarbete II**


**Delarbete III**

De två första delarbetena beskrev *kvalitativa* metoder för att separera vatten- och fettsignal. Med *kvantitativ* FWI kan man erhålla ett tillförlitligt värde på
fettinnehållet i varje pixel (% av total signal). Detta förutsätter att data samlas in med fler ekotider och att mer avancerade signalmodeller används.

Delarbete III beskriver en generell metod som kan användas för både kvalitativ och kvantitativ FWI. Magnetfältet beräknades med en befintlig metod som utökades till tre dimensioner. Dessutom föreslogs en snabbare lösningsalgoritm för att kunna hantera de större datamängder som är förknippade med tredimensionella bilder. Snabb bildrekonstruktion med få rekonstruktionsfel demonstrerades i experiment med tredimensionell avbildning av buken hos tio frivilliga försökspersoner.

Delarbete IV
I denna studie beskrivs en metod som utökar FWI för att kvantitativt mäta vissa egenskaper hos fett. Fettmolekylerna i en pixel kan karakteriseras i term av fettsyrororna genomsnittliga kolkedjelängd och mättnadsgrad (antal dubbelbindningar). Genom att samla in data med ett större antal ekotider (32 användes i denna studie), kan dessa egenskaper beräknas matematiskt i varje pixel.


Utsikt

Ett förnyat intresse för FWI har märkts de senaste åren. Inte minst gäller detta kvantitativ FWI för att diagnosticera leverförfettning. De stora tillverkarna av magnetkameror har nu börjat implementera färdiga lösningar för FWI.

Förhoppningsvis bidrar detta avhandlingsarbete till förädlingen av FWI, och att dessa metoder får ökad spridning. De metoder som utvecklats i avhandlingen är i princip kliniskt tillämpbara och används redan i flera forskningsprojekt, inklusive studier av diet, fetma och det metabola syndromet.
1. Introduction

Modern tomographic techniques are able to produce images of the internal structures of the human body. This amazing fact is rendered possible through sophisticated use of physics, electrical engineering, mathematics, and signal processing. This thesis is concerned with the most intriguing tomographic technology – magnetic resonance imaging (MRI).

1.1 Background

The signals that constitute magnetic resonance images emerge from the imaged object itself. This is not the case for other tomographic techniques, such as positron emission tomography (PET), where the signal is emitted by an externally administered radioactive isotope, or computer tomography (CT), where the signal is formed by x-rays penetrating the object. The absence of ionizing radiation makes MRI attractive both in clinical settings and for research purposes. MRI is a versatile technique offering true volumetric (three-dimensional) imaging. In fact, the images may extend in further dimensions such as time or resonance frequency. The resolution domain ranges from microscopic imaging to 3D imaging of the whole body. The versatility and safety of MRI has led it to develop in the last decades into one of the most important diagnostic tools of modern healthcare.

Most of the signal in MR images of the human body originates from $^1$H nuclei in water and fat molecules. The fat signal appears bright in most types of images, while the water signal is usually of greater medical importance. Therefore, suppression of the fat signal is often desired. Conveniently, water and fat have distinct resonance frequencies – there is a chemical shift between water and fat. Beyond the spatial dimensions, chemical shift imaging (CSI) enables encoding of the chemical shift dimension. With proper signal modeling, the contributions from water and fat can be estimated and separated with only a few samples in the chemical shift dimension. In this thesis, such methodology is referred to as fat-water imaging (FWI).

Being able to produce separate water and fat images, FWI provides unique information on the location of fat. The water image is free from fat signal, which is desirable in clinical settings. Advantages of FWI over conventional fat suppression techniques include molecular specificity, noise efficacy, and the possibility to estimate and compensate for magnetic field inhomogeneity.
The main disadvantages are prolonged acquisition times, long image reconstruction times, and vulnerability to reconstruction errors that degrade image quality. FWI is also important in body composition research, comprising a supreme basis for segmentation of adipose (fatty) tissue. Moreover, FWI includes the possibility to quantitatively measure the fat percentage in each pixel.

1.2 Overall aim

The aim of this thesis is to provide an overview of the FWI literature, to increase the understanding of FWI, and to refine FWI with respect to acquisition time, reconstruction time, and image quality. This includes developing algorithms robust to reconstruction errors, and to remove acquisition parameter constraints in order to allow faster imaging. There is a special focus on rapid algorithm performance in order to make the methods practical in research settings, and to allow future implementation directly on the MRI scanners. An additional aim is to extend conventional FWI to enable mapping of fat quantities such as fatty acid chain length and degree of saturation.

1.3 Structure of the thesis

This is a compilation thesis based on four papers, which are summarized in chapter 5. Paper I describes an FWI reconstruction algorithm tailored for the special case of whole-body imaging with three samples in the chemical shift dimension. Paper II introduces an FWI method with only two flexible samples in the chemical shift dimension, allowing for minimal acquisition times. Paper III describes a general FWI method with three or more chemical shift samples, including a sophisticated problem formulation and solution. Paper IV introduces a method for quantitative mapping of fat chain length and degree of saturation.

The framing text is written to provide a context of these papers. Chapter 2 describes the basic theory of magnetic resonance imaging. Chapter 3 focuses on the medical setting, including hardware, body composition, and different fat suppression techniques. A review of FWI techniques and theory of the novel contributions is given in chapter 4. Chapter 6 offers a discussion of the present, previous, and future work.
2. Magnetic resonance imaging

Resonance is the tendency of a system to change between states with certain frequencies. An example is a pendulum in a gravity field, which will swing with a natural frequency that increases with gravity. Similar is the movement of a magnet, such as a compass needle, in an external magnetic field. The needle will swing about the direction of the north pole with a frequency proportional to the strength of the magnetic field. This is known as magnetic resonance. In principle, some atomic nuclei behave like tiny magnets. When placed in a strong magnetic field, the phenomenon of nuclear magnetic resonance can be observed [1, 2]. Different types of resonance are illustrated in fig. 2.1.

Nuclear magnetic resonance (NMR) is the foundation of MRI, although ‘nuclear’ has been left out due to the negative connotations of the word. In order to receive a signal, the nuclei are manipulated with rotating electromagnetic fields in the radio frequency range. To obtain an image, the signal is spatially encoded by adding linearly varying gradient fields to the static field. Thus, three types of magnetic fields are required for MRI; a strong static field (measured in units of Tesla, T), rotating radio frequency fields, and linear gradient fields.

Before NMR was used for imaging, NMR spectroscopy (or magnetic resonance spectroscopy, MRS) was developed for chemical analysis. MRS is an important tool for structural and functional analysis of molecules, not least in the study of proteins, and has several applications in medicine. MRS is further described in section 2.4.

Figure 2.1: Different types of resonance. (a): Mechanical resonance. (b): Magnetic resonance. (c): Nuclear magnetic resonance.
2.1 Nuclear magnetic resonance

The magnetic properties of atomic nuclei are due to the quantum mechanical
property of spin, which is a type of angular momentum. Protons and neutrons
have spin 1/2. Depending on the number of protons and neutrons, a specific
isotope may or may not have residual spin. Isotopes with residual spin include
\(^1\text{H}\), \(^{13}\text{C}\), \(^{14}\text{N}\), \(^{17}\text{O}\), \(^{31}\text{P}\), and \(^{129}\text{Xe}\), of which \(^1\text{H}\) is the most important for
medical MRI. A table with the properties of some isotopes can be found on
page 41.

Spinning nuclei have a magnetic moment that will tend to align with an
externally applied magnetic field. Due to the angular momentum, the magnetic
moment of a nucleus \textit{precesses} in a conical motion about the direction of the
external field rather than oscillates (see fig. 2.1c). This is however a simpli-
fied picture, since individual nuclei obey the non-intuitive laws of quantum
mechanics. Yet, MR experiments are always conducted on large ensembles
rather than single nuclei. Therefore, quantum mechanical explanations are su-
perfluous in most aspects of MRI [3].

The total angular momentum of an ensemble of spins can be described by
a vector \(\mathbf{A}\). The angular momentum is related to the magnetic moment

\[
\mathbf{M} = \gamma \mathbf{A}
\]

where the \textit{gyromagnetic ratio} \(\gamma\) is specific to the isotope (for \(^1\text{H}\), \(\gamma/2\pi = \gamma = 43\) MHz/T). In a magnetic field \(\mathbf{B}\), the magnetic moment will experience a
torque

\[
\tau = \mathbf{M} \times \mathbf{B}
\]

forcing the angular momentum to change as

\[
\frac{d\mathbf{A}}{dt} = \tau
\]

According to equations 2.1 – 2.3, the rate of change of the magnetic moment
is perpendicular to both itself and the magnetic field [2]:

\[
\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{B}
\]

Since the magnetic moment only changes perpendicular to itself, its magni-
tude is constant.

By convention, the \(z\)-axis, also referred to as the \textit{longitudinal axis}, is
aligned along the static magnetic field. The perpendicular \(xy\)-plane is called
the transversal plane. In the case of a static field

\[ \mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix} \]  

(2.5)

the differential equation 2.4 has the solution

\[ \mathbf{M}(t) = M_0 \begin{pmatrix} \cos(-\gamma B_0 t) \sin(\theta) \\ \sin(-\gamma B_0 t) \sin(\theta) \\ \cos(\theta) \end{pmatrix} \]  

(2.6)

i.e. \(|\mathbf{M}| = M_0\) is constant, and the magnetization precesses about the static field at a constant angle \(\theta\), with a characteristic resonance frequency

\[ \omega = \gamma B_0 \]  

(2.7)

also known as the Larmor frequency. Precession of the magnetic moment is illustrated in fig. 2.2 a.

In order for an additional field \(B_1\) to be perpendicular to both the magnetic moment and the static field, it must also rotate with the resonance frequency about the direction of the static field so that the total field becomes

\[ \mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix} + \begin{pmatrix} -B_1 \sin(-\gamma B_0 t) \\ B_1 \cos(-\gamma B_0 t) \\ 0 \end{pmatrix} \]  

(2.8)

\[ \begin{array}{c}
(a) \\
(b) \\
(c)
\end{array} \]

Figure 2.2: (a): Precession of magnetic moment about the direction of a static \(B_0\) field. (b): Excitation of magnetic moment with a rotating \(B_1\) field. (c): Mechanical “excitation”.

19
For such a combination of a static field and a rotating field, eq. 2.4 has the solution

$$M(t) = M_0 \begin{pmatrix} \cos(-\gamma B_0 t) \sin(-\gamma B_1 t + \theta) \\ \sin(-\gamma B_0 t) \sin(-\gamma B_1 t + \theta) \\ \cos(-\gamma B_1 t + \theta) \end{pmatrix}$$ (2.9)

In contrast to eq. 2.6, the angle of precession is no longer constant. The presence of a $B_1$ field rotating with the resonance frequency is able to increase the precession angle of the total magnetic moment. This is referred to as excitation, and is illustrated in fig. 2.2 b. The manipulation of the precession angle with a rotating magnetic field may seem abstract, but is similar to a mechanical pendulum, such as a swing, where applying synchronous pushes (with the right frequency) will gradually increase the elevation angle (fig. 2.2 c).

### 2.2 Relaxation

The dynamic behavior given by eq. 2.4 accounts for the interaction of the spinning nuclei with external magnetic fields. In practice, the spinning nuclei also interact with each other (spin-spin interaction) and with other nuclei, atoms and molecules (spin-lattice interaction). These interactions cause relaxation of the magnetization, forcing it to approach an equilibrium magnetization. The spin-spin relaxation gradually puts the spins out of phase, and the spin-lattice relaxation causes the spin distribution to become slightly oriented along the direction of the static field. Eventually, the equilibrium magnetization becomes parallel to the static field. The magnitude is given by a first order approximation of the Boltzmann distribution:

$$M_0 = \rho \frac{\gamma^2 \hbar^2}{4k_B T} B_0$$ (2.10)

where $\hbar$ is the reduced Planck constant, $k_B$ is the Boltzmann constant, $T$ is the temperature, and $\rho$ is the spin density. This magnetization is small compared to the total magnetic moment of the nuclei. Relaxation towards equilibrium magnetization is illustrated in fig. 2.3. Relaxation is similar to friction and air resistance in the case of the mechanical pendulum, which causes the pendulum to 'relax' and eventually stay pointing in the direction of the gravity field.

The *Bloch equation* extends eq. 2.4 to account for relaxation [2]:

$$\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{B} + \begin{pmatrix} -M_x/T_2 \\ -M_y/T_2 \\ (M_0 - M_z)/T_1 \end{pmatrix}$$ (2.11)

where $T_1$ and $T_2$ are relaxation time constants.
A common situation in MRI is that a rotating $B_1$ field is applied to the magnetization until it precesses perpendicular to the static field. Then, the $B_1$ field is switched off, leaving only the static field (eq. 2.5). It is informative to study the behavior of the magnetization in this special case. If the magnetization is oriented along the x-axis at time $t = 0$, the solution to eq. 2.11 is

$$M(t) = M_0 \begin{pmatrix} \cos(-\gamma B_0 t) e^{-t/T_2} \\ \sin(-\gamma B_0 t) e^{-t/T_2} \\ 1 - e^{-t/T_1} \end{pmatrix}$$

(2.12)

This is the situation illustrated in fig. 2.3 a. The magnetization along the z-axis increases according to the time constant $T_1$. $T_1$ is also called spin-lattice relaxation time or longitudinal relaxation time, and the relaxation process is called $T_1$ recovery. Simultaneously, the magnitude of the rotating magnetization in the xy-plane decreases exponentially according to the time constant $T_2$. This process is called $T_2$ decay. $T_2$ is also known as spin-spin relaxation time or transversal relaxation time. The relaxation times are characteristic for different materials and tissues, but vary between magnetic field strengths. Typically, $T_2$ decay is faster than $T_1$ recovery.

Relaxation is assumed to be the result of stochastic processes such as thermal motion. However, deterministic conditions, such as inhomogeneity of the static field amplitude $B_0$, causes different spins to precess with different resonance frequencies (eq. 2.7). As a result, the spins will dephase, causing decay of the total transversal magnetization. Therefore, the effective transversal relaxation time will be $T_2^*$ defined as

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

(2.13)

where $T_2$ is the stochastic contribution and $T_2'$ the deterministic. By definition, $T_2^*$ is always shorter than $T_2$. Due to its deterministic nature, the $T_2'$ signal loss
can be recovered, such as in a spin echo experiment which will described in section 2.3.1.

Longitudinal and transversal relaxation are complicated processes involving many kinds of interactions between particles. As such, they were introduced as phenomenological processes. Although, there have been attempts to derive relaxation times from first principles, such as Bloembergen-Purcell-Pound theory (BPP theory) [4].

2.3 Signal generation

The generation of a nuclear magnetic resonance signal is described in the following. A sample (or a patient) is placed in a strong static magnetic field $B_0$, typically created by a superconducting magnet. $B_0$ is on the order of a few Tesla (T), causing the $^1$H nuclei to precess on the order of 100 MHz. The net magnetic moment of the nuclear spins will relax to equilibrium, as described above. Then, for a short time, the sample is exposed to a weaker field $B_1$, rotating with the resonance frequency in the transverse plane. Since the resonance frequency is in the radio frequency range, this is called a radio frequency pulse (RF-pulse). According to eq. 2.9, the RF-pulse increases the precession angle of the net magnetic moment from zero to an angle $\alpha$, which depends on the duration and amplitude of the pulse. Such an RF-pulse is called an $\alpha^\circ$-pulse, and $\alpha$ is called the flip angle. The largest transversal magnetic moment is obtained by a 90$^\circ$-pulse. After the RF-pulse, the magnetic moment precesses about the direction of the static field. Due to Faraday’s law of induction, this changing magnetic moment induces electric currents in nearby conducting materials. Thus, the magnetic moment can be measured. If two coils are aligned with their symmetry axes in the transversal plane and at right angles to each other (in quadrature), the precession of the magnetic moment will induce electric signals in each of the coils. Together, these signals can be viewed as a complex signal, corresponding to the evolution of the transversal magnetic moment over time. This signal will decay due to $T_2^*$ relaxation, and is therefore called free induction decay. As pulse sequences tend to get complicated, they can be represented using timing diagrams. A timing diagram for

![Figure 2.4: Free induction decay. The received signal decays according to $T_2^*$ after excitation.](image)
the simple pulse sequence described above is shown in fig. 2.4. In MRI, the free induction decay is typically not measured. Instead, so-called echoes are created, the most important being spin echoes and gradient echoes.

### 2.3.1 Spin echoes

After applying a $90^\circ$-pulse, the transversal magnetization will decay exponentially according to $T_2^*$, defined by eq. 2.13. As discussed previously, part of this decay is an effect of different spins not precessing equally fast. If a second RF-pulse with a flip angle of $180^\circ$ is applied, all spins will rotate $180^\circ$ about the direction of the $B_1$-field. This means that the longitudinal magnetization is unchanged, but the relative phases of the spins in the transversal plane are inverted. Faster spins that were leading in phase before the $180^\circ$-pulse, will now be lagging in phase. Following the $180^\circ$-pulse, these spins will catch up with the slower spins. The rephasing of the spins results in a signal echo [5]. A timing diagram of the spin echo pulse sequence is shown in fig. 2.5. The time between the $90^\circ$-pulse and the centre of the echo is called **echo time**, $TE$. The signal amplitude at $TE$ depends on $T_2$ rather than $T_2^*$. Several spin echoes, a so-called Carr-Purcell train [6], may be created by expanding the pulse sequence with additional $180^\circ$-pulses.

### 2.3.2 Stimulated echoes

Just like two RF pulses create a spin echo, three RF pulses result in a *stimulated echo* [5]. In the simplest case, the $180^\circ$-pulse of a spin echo sequence is split into two $90^\circ$-pulses separated by a **mixing time**, $TM$. The timing diagram is shown in fig. 2.6. After the first pulse, the transversal magnetization relaxes according to $T_2^*$. It is then flipped onto the negative longitudinal axis, and subject to $T_1$ relaxation during $TM$. After the third pulse, the signal rephases. The signal amplitude is now a function of both $T_1$ and $T_2$. Sequences where stimulated echoes are recorded are called **stimulated echo acquisition mode** (STEAM) [7]. When used for imaging, the last $90^\circ$-pulse can be split up into
Figure 2.6: Timing diagram of a stimulated echo sequence. The first two pulses create a spin echo (SE), and the combination of the three pulses yields a stimulated echo (STE).

several smaller pulses, each yielding a (smaller) stimulated echo. If each stimulated echo is differently spatially encoded, rapid imaging is possible [8]. In addition to the stimulated echo (STE), spin echoes (SE) will also appear, as indicated in fig. 2.6.

2.3.3 Inversion recovery

An inversion recovery sequence starts with an inversion pulse of 180°. Then, $T_1$ relaxation occurs during the inversion time $TI$, before the sequence proceeds. Inversion recovery is a ‘magnetization preparation pulse’, the rest of the sequence typically being a spin echo or a gradient echo. The timing diagram of a spin echo sequence with inversion recovery is shown in fig. 2.7. Inversion recovery introduces $T_1$ dependence of the signal. After the inversion pulse, the magnetization is aligned along the negative $z$-axis and relaxes through zero towards equilibrium on the positive $z$-axis. Therefore, the inversion time $TI$ can be set to null signal components with a specific $T_1$. This is utilized in fluid attenuation inversion recovery (FLAIR) [9] and short time inversion recovery (STIR) [10], the latter being used for fat suppression and described in greater detail in section 3.5.1.

Figure 2.7: Timing diagram of an inversion recovery prepared spin echo sequence.
2.3.4 Gradient echoes

Gradient echoes (GRE) are created with a magnetic field gradient instead of an RF-pulse [11]. Gradient fields are described in section 2.5. After the excitation pulse, a gradient is switched on, causing the spins to dephase. Then, the polarity of the gradient is reversed, so that the spins rephase and form a signal echo. The timing diagram of a gradient echo is shown in fig. 2.8. Note that the spins rephase only with respect to the gradient, and not to other sources of dephasing. Therefore, the signal amplitude is subject to $T_2^*$ decay. Multiple gradient echoes can be formed by adding further gradients with alternating polarity [12]. In gradient echo imaging, the excitation pulse is usually smaller than $90^\circ$.

2.3.5 Repeated pulse sequences

In practice, the pulse sequences are always repeated with a fixed repetition interval, $TR$. This is done both for signal averaging to reduce noise and for different space encoding in the different repetitions (explained in more detail in section 2.5). Some signal properties are required to remain constant between repetitions.

Two different situations are of interest; either $TR$ is large enough so that full relaxation is guaranteed in each repetition, or else the pulse sequence must be designed so that a steady state arises [13]. Steady state means that relaxation together with pulse sequence manipulation of the magnetization, results in an identical magnetization at the beginning of each repetition. Often, this does not happen until after one or more repetitions. In these cases, a series of dummy repetitions are played before actually recording the signal.

Since $T_1$ relaxation is slower than $T_2$, it is also important to distinguish if $TR$ is on the order of $T_1$ but longer than $T_2$, or if it is shorter than both types of relaxation. In the first case, a steady state will arise for the longitudinal
magnetization only, while the transverse magnetization is relaxed in each repetition.

In gradient echo imaging, $TR$ is often shorter than both $T_1$ and $T_2$. A common strategy is to *spoil* residual transverse magnetization at the end of each repetition using crusher gradients and/or RF-pulses. Then, steady state needs only to be reached for the longitudinal magnetization. This is called *spoiled gradient echo*. Gradient echo without spoiling requires the same gradient moment of a gradient in each repetition to reach steady state for the transverse magnetization [14]. This is called *steady state free precession* (SSFP). If the gradient moment for each gradient is zero in each repetition, it is said to be *balanced* (bSSFP). Balanced gradient echo sequences have the largest rate of received signal per time, but the signal intensity depends both on $T_1$ and $T_2$, and is sensitive to resonance frequency offsets.

2.4 Magnetic resonance spectroscopy

Until now, it has been assumed that all nuclei of a particular isotope precess with same resonance frequency. However, the magnetic moments of electrons cause local distortions of the static field, called *electron shielding*. In effect, nuclei surrounded by many electrons experience a lower static field, and hence precess more slowly. Therefore, the resonance frequency of a nucleus depends on its chemical context, such as its position within a molecule. Accordingly, the MR signal contains chemical information [15].

Differences in resonance frequency are also called *chemical shifts*. Chemical shifts are measured relative to some reference frequency (tetra-methyl silane for $^{1}H$). Since the shielding effect is small compared to the static field,

\[ \text{chemical shift [ppm]} \]

*Figure 2.9:* Magnetic resonance proton spectrum of the liver. The large resonance at 4.7 ppm is attributed to water protons. The smaller resonances originate from protons in fat.
chemical shifts are given in units of parts per million (ppm). For instance, the $^1$H nuclei in water molecules have a chemical shift $\approx 4.7$ ppm (the exact shift depends on the temperature [16]). To excite all relevant resonance frequencies, the RF-pulses must cover a certain bandwidth.

Complex samples of the signal induced by the transversal magnetization, such as a spin echo, are acquired over time. A complex frequency domain spectrum is then obtained by applying a discrete Fourier transform [17]. An example of a magnetic resonance spectrum is shown in fig. 2.9. By convention, the resonance frequencies decrease along the $x$-axis so that higher frequencies are to the left and lower frequencies to the right. Each specific resonance frequency will correspond to a line (peak) in the spectrum. Exponential $T_2^*$ decay in the time domain causes line broadening in the Fourier domain, giving a so-called Lorentzian lineshape. Faster decay corresponds to greater linewidth. The area under a resonance line is proportional to the number of nuclei processing with the associated resonance frequency.

2.5 Space encoding

In order to produce images, it is necessary to know something about the position of the nuclei emitting the magnetic resonance signal. In the above description, two types of magnetic fields were used; a static field $B_0$ and a rotating field $B_1$. An exception to this was gradient echoes, which used a third type of magnetic field: gradient fields. However, the primary use for magnetic field gradients is spatial encoding [18, 19].

An MRI scanner is equipped with three sets of gradient coils; one for each of the $x$, $y$, and $z$ axes. Each set of coils is able to temporary augment the static field, by an amplitude that varies linearly on its axis. This changes only the amplitude of the total field, not the direction. The gradient amplitudes are much smaller than the amplitude of the static field, and are measured in units of mT/m. If, for example, a patient lies in an MR scanner with a static field of 3.0 T and a gradient field of 40 mT/m is added along the left/right direction, the left side of the patient will experience a total field of 2.99 T, while the total field on the right side will be 3.01 T. Linear combinations of the gradient fields enable linearly varying fields in arbitrary directions. This, in turn, allows arbitrary orientation of the images.

According to eq. 2.7, a spatially varying magnetic field yields spatially varying resonance frequencies. This enables spatial encoding of the MR signal. Chemical shift of the resonance frequencies occur simultaneously with shifts due to field gradients. Although, the field gradient shifts are much larger than the chemical shifts. In the reconstructed images, chemical shift manifests as a minor spatial shift, known as chemical shift artifact of the first kind.
2.5.1 Slice selection

The simplest example of spatial encoding is slice selection, which is used in 2D MRI [20]. The idea is to excite only the spins in a certain thin volume, a slice, so that any recorded signal is known to originate from that slice only (note that even 2D MRI is ‘volumetric’). This is possible by applying a gradient field at the same time as the excitation pulse, such as the 90° pulse in a spin echo sequence. The frequency profile of the RF-pulse together with the field gradient determines which spins will be excited. Spins with resonance frequencies outside the bandwidth of the excitation pulse will not be excited. If the frequency profile of the RF-pulse is exactly rectangular, a rectangular slice of spins will be selected. In principle, rectangular frequency profiles are not achievable in finite time, so trade-offs must be made. The selected slice is perpendicular to the direction of the gradient, which is called the slice encoding direction in this context. Slice selection is illustrated in fig. 2.10. In 3D MRI, slice selection is also called slab selection. Due to larger RF bandwidth, a volume rather than a thin slice is excited.

2.5.2 Localized spectroscopy

To perform MR spectroscopy in vivo (in live animals or humans), it is necessary to localize the acquisition so that the spectrum can be associated with a particular volume of interest (a voxel = volumetric pixel). One kind of localized spectroscopy is chemical shift imaging (CSI), which is described in section 3.5.5. More common, however, is single voxel spectroscopy. The technique used is similar to slice selection, but the selection must be performed along three spatial directions. This is easily achieved for STEAM sequences,
where field gradients can be applied in each of the three directions simultaneously with the three 90°-pulses [21]. Alternatively, a spin echo sequence with two 180° pulses can be used. By applying field gradients along each direction during the 90° and the two 180° pulses, the second spin echo will originate from spins in the selected volume only. This technique is called point-resolved spectroscopy (PRESS) [22]. A newer technique that uses adiabatic (frequency modulated) RF pulses is called localization by adiabatic selective refocusing (LASER) [23].

The most important in vivo application of MRS is 1H spectroscopy of the brain. In such applications, suppression of the signal from water protons is often required (similar to fat saturation, described in section 3.5.2).

2.5.3 Fourier encoding

As explained above, a gradient field introduces position-dependence of the resonance frequency. If, for example, a gradient $G_x$ is applied along the $x$-axis, the resonance frequency at position $x$ will be

$$\omega(x) = \gamma(B_0 + G_x \cdot x)$$

If a gradient is applied during sampling of the signal echo (i.e. during readout), the sampled data will have contributions from all frequencies corresponding to positions occupied by the sample/object/patient. Just as in spectroscopy, a discrete Fourier transform can be applied to the sampled signal to obtain the frequency distribution [24]. Since the frequency distribution is linear in space, the Fourier transform of the signal sampled over time will correspond to the spatial signal distribution. Note that the spatial distribution is obtained along the direction of the applied gradient only.

The spatial distribution along the direction of the readout gradient can be regarded as a one-dimensional image. It is desirable to acquire three-dimensional images or at least two-dimensional image slices. However, only one direction can be spatially encoded during each readout; the spatial dimensions are not separable using linear gradient fields. This is solved by repeating the pulse sequences with different space encoding in each repetition [24]. For 2D imaging, one direction is designated as the fast encoding direction. A gradient is applied along this direction during each signal readout. Perpendicular to both the slice encoding direction and the fast encoding direction is the slow encoding direction. Between the excitation pulse and signal readout, a gradient is applied along the slow encoding direction. The amplitude of this gradient is uniformly varied between repetitions as indicated by the multiple overlaid gradients in fig. 2.11. In this way, the fast and the slow direction are independently spatially encoded.

The procedure described above is the most commonly used imaging technique in MRI. It is known as spin warp imaging [25] or Cartesian imaging.
The fast encoding direction is also called the frequency encoding direction and the slow encoding direction is also known as the phase encoding direction.

If $M$ data points are sampled during each signal readout, and the pulse sequence is repeated $N$ times, an $M \times N$ data matrix $S(k_x, k_y)$ is obtained. The coordinates $k_x$ and $k_y$ of a point in the data matrix are proportional to the gradient moment along the $x$ and $y$ directions. That is, the time integral over the gradient between excitation and sampling of that data point [26]. Since the readout gradient is preceded by a dephasing gradient with negative amplitude, and the slow gradient ranges from negative to positive amplitudes, the central point of the data matrix corresponds to zero gradient moments, $k_x = k_y = 0$. A change of the gradient moment can be thought of as a translation in the data space, more commonly referred to as the $k$-space. A k-space data matrix is shown in fig. 2.12. The 2D image $y(x, y)$ is given by a discrete 2D Fourier transform of the data matrix [24]:

$$y(x, y) = \sum_{k_x, k_y} S(k_x, k_y) e^{i(xk_x + yk_y)}$$  \hspace{1cm} (2.15)

Typically, only the magnitude of the complex image is displayed and the phase information is discarded. The magnitude of a Fourier transformed data matrix, i.e., an image, is shown in fig. 2.12.

The extension to 3D imaging is straightforward; one fast and two slow encoding directions are used. This gives a 3D k-space, and the 3D image is obtained through a 3D Fourier transform. 3D imaging requires many repetitions, so $TR$ must be short. For this reason, 3D imaging is mostly used in gradient echo sequences [27]. For spin echo sequences, multiple slices can be obtained.
in an interleaved fashion. This is also a kind of 3D imaging, although there may be gaps between the slices.

The concept of \textit{k-space} \cite{26} is rather abstract, but is mathematically related to the image space through the Fourier transform. Points in image space correspond to spatial positions. Conversely, points in k-space correspond to spatial frequencies and specific gradient moments. The central k-space corresponds to low spatial frequencies, and the k-space periphery corresponds to high spatial frequencies. Note that each point in image space depends on each point of k-space (eq. 2.15). Conversely, each k-space sample has contributions from the entire imaging volume.

It should be emphasized that k-space can be interpreted both as a frequency domain (in terms of spatial frequency, not temporal) and as a time domain (since the k-space points are samples of the signal evolving over time). This false dichotomy may be subject to some confusion.

If the measured data is assumed to have limited spatial extension (called \textit{field of view}, FOV) along the fast encoding direction, an applied readout gradient of a certain strength results in a bandwidth limited signal. For instance, a FOV of 0.5 m and a readout gradient strength of 40 mT/m gives a bandwidth of \(\pm 430\) kHz for \(^1\)H imaging. The signal then needs to be sampled no faster than 860 kHz according to the Nyquist sampling criterion. A certain number of samples, results in the same number of pixels within the FOV after the Fourier transform. Thus, faster sampling extends the FOV while more samples (longer readout) increases the resolution.

### 2.6 Image reconstruction

The process of magnetic resonance imaging is summarized in fig. 2.12. In particular, obtaining an image from the data matrix is referred to as \textit{image reconstruction}. The simplest type of image reconstruction is a discrete Fourier transform of a spin-warp data matrix. Image reconstruction may also include

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{image12.png}
\caption{The process of magnetic resonance imaging consists of two distinct steps; Data acquisition and image reconstruction.}
\end{figure}
procedures such as post-processing or model fitting of the data (which is covered in this thesis, with respect to separation of water and fat signal).

In spin-warp imaging, k-space is sampled on a rectangular grid. This makes sense, since images are defined on rectangular grids, and the Fourier transform of a rectangular grid is also a rectangular grid. However, this is not the only way to sample k-space. The concept of moving through k-space by applying gradient moments [28] allows arbitrary sampling patterns. Some of the most important k-space trajectories are depicted in fig. 2.13. After data acquisition, the samples are often interpolated onto a rectangular grid. This is called gridding. For a rectangular grid, the multi-dimensional discrete Fourier transform is separable, so that consecutive one-dimensional discrete Fourier transforms can be applied along each dimension. Additionally, the equidistant spacing of the samples allows the use of fast Fourier transforms (FFT) [29]. FFT algorithms are fundamental for performing MRI with practical reconstruction times.

One of the most common post-processing operations is zero filling (or zero padding), which expands the data matrix by adding zeroes to the k-space periphery [24]. This increases the nominal resolution and corresponds to an interpolation of the image by convolution with a sinc-kernel. Obviously, no actual data is added, so the effect is merely cosmetic. Zero filling can be used to increase the number of data points along each dimension to become a power of two, so that more efficient FFT algorithms can be used. Although sometimes claimed otherwise, FFT algorithms exist for arbitrary numbers of data points [30].

Figure 2.13: Different k-space trajectories. (a): Cartesian. (b): Spiral. (c): Polar. In the Cartesian case, no signal is sampled along the sloped sections of the trajectory, which correspond to gradient dephasing. The spiral trajectory shown here consists of six interleaved spirals, and the polar trajectory consists of 16 radial spokes.
2.7 Imaging speed

MRI is a relatively slow process; a typical medical examination with MRI takes 30–60 minutes. Thus, attempts of increasing imaging speed are as old as MRI itself. The advantages of faster imaging are obvious and include reduced patient discomfort, higher patient throughput, reduced costs and decreased motion artifacts. Motion artifacts are the result of some physical motion, such as breathing, being faster than the image acquisition. Therefore, motion artifacts are more pronounced along the slow encoding direction.

One of the earliest techniques for fast imaging is echo planar imaging, EPI [12]. EPI employs a fast encoding gradient, with alternating polarity to form multiple refocused echoes, together with a slow encoding gradient. This enables recovering data from an entire plane in k-space in one repetition, rather than only a single line. This is also achievable by other k-space trajectories, such as spiral imaging (fig. 2.13 b).

Another way to encode more k-space per repetition is through the use of a Carr-Purcell train [6] with different phase encoding for each spin echo [31]. This is known as fast spin echo, turbo spin echo or RARE. Fast spin echo can even be combined with EPI to encode several k-space lines for each spin echo, known as gradient and spin echo (GRASE). A problem with acquiring much data in a single excitation is that relaxation occurs during the course of signal sampling, so that different parts of k-space become subject to different amounts of relaxation.

Relaxation times were once believed to be a limiting factor for MRI acquisition speed. Therefore, using repetition times shorter than the relaxation times have been a crucial step towards faster imaging. Short TR combined with 90°-excitation pulses results in a steady state with small longitudinal magnetization and, consequently, low signal strength. To maintain signal strength between repetitions, smaller flip angles can be used. This is the key idea of fast low-angle shot (FLASH) [11]. If the $T_1$ relaxation rate is known, calculation of the flip angle that maximizes the steady state signal (the so-called Ernst angle) for a given TR is straight-forward [17].

As previously suggested, greater amplitudes of the gradients allow faster signal sampling and more compact pulse sequences. However, modern gradients have reached safety limits associated with peripheral nerve stimulation of patients due to fast gradient switching [32]. Stronger gradients may also experience problems related to self-induced eddy currents. Using stronger gradients results in lower signal-to-noise ratio (SNR), because of the reduced signal readout time (not because of the increased bandwidth [33]).

Recent development of accelerated MRI has been focused on parallel imaging, referring to that multiple receiver coils are used in parallel to sample the data. The use of multiple coils enables sampling below the Nyquist limit. Spins closer to the coil induce a greater signal, giving the coils different sensitivity profiles. Thus, the coils constitute a coarse space encoding,
fundamentally different and complementary to that of the gradient fields. This allows parallel imaging to reconstruct images from an under-sampled k-space. Approaches to parallel imaging include SMASH [34], SENSE [35], and GRAPPA [36]. A new and promising approach to parallel imaging is called regularized nonlinear inversion [37]. This technique has been combined with radially encoded FLASH [38] for 2D real-time imaging of the heart at 50 frames per second [39].

The theory of compressed sensing [40] is another new area of research interest, motivated by the fact that images can be compressed with small or no loss of image quality. Inversely, it should be possible to faithfully reconstruct an image from under-sampled data. This is done by incorporating sparsity into the image reconstruction. As opposed to the fully sampled case, multiple solutions (images) may be consistent with the under-sampled data. The solution is chosen that maximizes the sparsity in some transform domain [41].
3. Magnetic resonance in medicine and biology

Medical applications of NMR were introduced in the 1950s by Erik Odeblad [42]. However, the interest in medical NMR grew tremendously since 1971 when Raymond Damadian suggested that relaxation times can differentiate malignant tumors from healthy tissue [43]. Damadian demonstrated the first MR image of the human body in 1977 [44], albeit using a method that saw limited use in practice. Instead, spatial encoding using gradient fields turned out to be the crucial component for MRI as we know it, resulting in MR images of a human finger in 1974 [45] and of the human body in 1980 [25]. Subsequently, the decision to award the 2003 Nobel price in physiology or medicine to Sir Peter Mansfield and Paul Lauterbur, but not to Damadian, was subject to some controversy [46]. Today, medical applications of MRI dominate the development of new techniques.

A great advantage of MRI in medical settings is its safeness [47], not least in light of the ionizing radiation of computer tomography (CT). Another advantage, again in relation to CT, is the attainable excellent contrast between different soft tissues. However, perhaps the most attractive feature of MRI is its flexibility, offering great versatility and continuous development of new applications. A common usage is to acquire images weighted by $T_1$ or $T_2$ relaxation, offering different types of contrast. Examples of different contrast mechanisms are shown in fig. 3.1.

Beyond relaxation times, MRI is also sensitive to flow. This enables depiction of flow in the blood vessels, which is called MR angiography [48]. Angiography can be further enhanced by contrast agents [49], a kind of relaxation catalysts (see fig. 3.1 c). Contrast agents were proposed already in 1946 by Bloch [2], although not with medical applications in mind. Intravenous administration of a contrast agent shortens the relaxation times of the blood, resulting in hyper-intense signal with some pulse sequences. This is often helpful in differentiating tumors from healthy tissue. Contrast agents are also useful for monitoring physiological function. Following the contrast agent over time can reveal blood supply to different organs, heart function etc. Fig. 3.2 illustrates the influence of a contrast agent on the image appearance.

Another interesting technique is functional MRI (fMRI). It is based on the fact that deoxygenated blood has a shorter $T_2^*$ than oxygenated, known as the BOLD effect [50]. This enables indirect measurements of brain activity [51]. Typically, the patient is alternately exposed and not exposed to some stimulus
or given some cognitive task, while $T^*_2$-weighted images are acquired. Signal change statistically correlated with the stimulus is provided as a map, which is overlaid on anatomical images of the brain. In psychology, fMRI is met with great interest. It may also provide important information prior to brain surgery.

Yet another important use of MRI is diffusion weighted imaging (DWI), where the signal intensity is related to microscopic movement of the spins [52, 53]. A diffusion weighted image is shown in fig. 3.1 d. Tissues with restricted diffusion, such as some tumors, appear intense. It is also possible to reveal the major axis of diffusion, which can be used to track the orientation of fibers in white brain matter [54].
3.1 Hardware

A complete MRI system has a large number of components. The most important are the main magnet, the gradient coils, RF coils for signal transmit and receive, and the computer system. Typically, the computer system consists of a host computer providing user input and output, a real-time computer (called the pulse-programmer) controlling the hardware, and a computer performing image reconstruction (called the array processor) [55].

There are several companies providing complete whole-body MRI systems for clinical use, including Philips, GE Healthcare, Siemens, Toshiba, and Hitachi. MRI machines require special environments. The scanners are installed in special rooms which are shielded from electromagnetic radiation (a so-called Faraday cage). The scanner room must be free from magnetic items for safety reasons. An MRI scanner can also be used for MR spectroscopy.

3.1.1 The magnet

The main magnet is responsible for creating the static $B_0$ field. Typically, a static field of 0.2–3.0 T is used. This can be compared to the Earth's magnetic field of approximately 50 μT. A higher field gives a stronger signal, according to eq. 2.10. The static field can be created by a permanent magnet (up to 0.3 T), an iron-cored electromagnet (up to 0.6 T), or a superconducting electromagnet (up to 8 T) [55]. Superconducting magnets are created of materials that are superconducting at very low temperatures, such as niobium-titanium. Such magnets require constant cooling by liquid helium [55].

Ideally, the amplitude of the static $B_0$ field should be constant throughout the field of view. In practice, attaining a perfectly homogenous magnetic field is difficult. $B_0$ field inhomogeneities are caused both by magnet imperfection, and by susceptibility differences in the imaged object. In modern magnets, inhomogeneities in the applied field are relatively small at the magnet isocentre, and typically increases towards the periphery of the magnet bore. Substantial $B_0$ inhomogeneity is provoked by a human body. This is particularly problematic at interfaces between air, bone, and soft tissue. The effect is even greater in the proximity of metallic implants. Static field inhomogeneity can be partly
compensated for by *shimming*, meaning that additional fine-tuned magnetic fields are created by so-called shim coils. By shimming, a homogeneous magnetic field can be produced within a small region of interest. However, using a large field of view, the inhomogeneity problem remains. Local offsets of the $B_0$ field amplitude result in resonance frequency shifts relative to the nominal frequency.

### 3.1.2 Gradient fields

Gradient fields are used for spatial encoding and for creating gradient echoes (see section 2.3.4). Three pairs of gradient coils are used to create gradient fields in three perpendicular directions. Gradient systems are characterized by their maximum gradient amplitude (measured in mT per meter) and their *slew rate* (measured in T per meter per second). Other important qualities are the gradient linearity, and the ability to avoid producing eddy currents [55].

### 3.1.3 Radio frequency coils

The radio frequency coils are used both to generate the rotating $B_1$ field, and to detect the signal. Larger coils detect more noise, so the coil should be as small as possible, while still encompassing the object to be imaged. The magnet
bore has a built-in coil, called the body coil, but different types of separate coils are also available. There are coils designed for specific body parts, such as head coils, as well as general purpose surface coils. Often, the body coil is used for transmitting RF pulses, and smaller coils are used to receive the signal. The amplitude of the detected signal is much smaller than that of the excitation pulses.

A uniform amplitude $B_1$ field must be delivered in order to obtain the same flip angle across the field of view. This is more problematic at higher fields, such as 3.0 T and 7.0 T (since the transmitted wavelength approaches the length of the field of view). One solution is to use multiple element coils and parallel transmit techniques [56]. Multiple element coils are also required for parallel imaging.

3.1.4 Experimental instrumentation

All experiments described in this thesis were performed using clinical whole-body MRI systems at the Uppsala University Hospital. All the equipment used was provided by Philips Healthcare (Best, the Netherlands). Three different standard clinical whole-body systems were used: Achieva 1.5 T, ACS 1.5 T, and Achieva 3.0 T. The 1.5 T Achieva scanner was equipped with a non-standard continuously moving table (COMBI) for whole-body scanning [57]. The built-in body coils, phased-array torso coils, and dual-element surface coils were used for RF-pulse transmitting and signal receiving.

All water and fat separation algorithms were implemented and run on a standard laptop computer. The code was written in c++ within the framework of an in-house image processing platform.
3.2 Body composition

According to the five-level model, body composition can be studied on any of five distinct levels; atomic, molecular, cellular, tissue-system, and whole-body [58].

On the atomic level, the human body is largely composed of hydrogen, carbon, nitrogen, oxygen, phosphorus, and calcium, as shown in table 3.1. Some isotopes of these elements and their properties are listed in table 3.2. Most of the carbon, oxygen, and calcium nuclei have no residual spin and give no MR signal. Compared to $^{14}$N and $^{31}$P, $^1$H has a higher gyromagnetic ratio, which gives a higher sensitivity. It is also more abundant in the body. Therefore, medical MRI is almost exclusively performed with respect to $^1$H. Since the $^1$H nucleus is a single proton, $^1$H imaging and spectroscopy is often referred to as proton imaging and proton spectroscopy.

On the molecular level, about 60% of the body weight can be attributed to water [58]. The three other large groups of molecules are lipids, proteins, and minerals, among which lipids and proteins contribute most to the $^1$H magnetic resonance signal.

The terms lipids and fats are often used interchangeably. Strictly, fat is identical to triglycerides (although liquid form triglycerides are denoted oils). Triglycerides are lipids, while the opposite is not necessarily true. However, most body lipids are triglycerides. Further adding to the confusion, adipose tissue is sometimes called fat, although only 80% of the mass is attributed to fat [60]. Strictly, lipids such as fat belong in the molecular level, while adipose tissue belongs in the tissue-system level [60].

Beyond water and lipids, MR visible resonances originate from proteins, peptides, and small mobile metabolites [61]. Together, such molecules contribute to a wide spectrum of resonances. Relatively large concentrations of any single such species is required to match the great abundance of water and lipids. MR visible lipids other than fat include cholesterol esters and free fatty

| Table 3.1: Typical composition of elements in the human body a) |
|-------------------|-----------------|----------------|
| Element           | Symbol | Percent of body weight |
| Oxygen            | O      | 61%               |
| Carbon            | C      | 23%               |
| Hydrogen          | H      | 10%               |
| Nitrogen          | N      | 2.6%              |
| Calcium           | Ca     | 1.4%              |
| Phosphorus        | P      | 0.8%              |
| Other             |        | 1.2%              |

a) Data from Wang et al. [58]
Table 3.2: Properties of some isotopes $^a$

<table>
<thead>
<tr>
<th>Isotope</th>
<th>Spin</th>
<th>$\gamma$ (MHz/T) $^b$</th>
<th>Natural abundance</th>
<th>Relative sensitivity $^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1$H</td>
<td>1/2</td>
<td>42.6</td>
<td>99.99%</td>
<td>100%</td>
</tr>
<tr>
<td>$^2$H</td>
<td>1</td>
<td>6.5</td>
<td>0.01%</td>
<td>0.96%</td>
</tr>
<tr>
<td>$^{12}$C</td>
<td>0</td>
<td>–</td>
<td>98.89%</td>
<td>–</td>
</tr>
<tr>
<td>$^{13}$C</td>
<td>1/2</td>
<td>10.7</td>
<td>1.11%</td>
<td>1.59%</td>
</tr>
<tr>
<td>$^{14}$N</td>
<td>1</td>
<td>3.1</td>
<td>99.63%</td>
<td>0.1%</td>
</tr>
<tr>
<td>$^{15}$N</td>
<td>1/2</td>
<td>-4.3</td>
<td>0.37%</td>
<td>0.1%</td>
</tr>
<tr>
<td>$^{16}$O</td>
<td>0</td>
<td>–</td>
<td>99.76%</td>
<td>–</td>
</tr>
<tr>
<td>$^{17}$O</td>
<td>5/2</td>
<td>-5.8</td>
<td>0.04%</td>
<td>2.91%</td>
</tr>
<tr>
<td>$^{18}$O</td>
<td>0</td>
<td>–</td>
<td>0.2%</td>
<td>–</td>
</tr>
<tr>
<td>$^{31}$P</td>
<td>1/2</td>
<td>17.2</td>
<td>100%</td>
<td>6.63%</td>
</tr>
<tr>
<td>$^{40}$Ca</td>
<td>0</td>
<td>–</td>
<td>99.94%</td>
<td>–</td>
</tr>
<tr>
<td>$^{43}$Ca</td>
<td>7/2</td>
<td>-2.9</td>
<td>0.14%</td>
<td>0.64%</td>
</tr>
</tbody>
</table>

a) Data from Hausser & Kalbitzer [59]
b) Gyromagnetic ratio
c) Sensitivity relative to $^1$H, assuming an equal number of nuclei

acids, but not lipids in cell membrane bilayers [61]. The concentration of these lipids is negligible compared to that of fat [62, 63).

The assumption, that all $^1$H MR signal of the human body originates from water and fat, is a useful first-order approximation.

3.3 Triglyceride spectrum

Fat molecules, triglycerides, consist of three fatty acids with ester bonds to one glycerol. An example triglyceride is shown in fig. 3.4. Assignment of the triglyceride $^1$H spectrum is well-established [64]. Following Ren et al. [65], the triglycerides may be characterized by ten distinct resonances, denoted $A–J$ (see fig. 3.4 and table 3.3).

Depending on their position within the triglyceride, different protons experience different amounts of electron shielding and hence different chemical shifts. Each triglyceride has five glycerol protons (two $G$, two $H$, and one $I$). Each of the three fatty acids has two protons $\alpha$ ($E$) and two protons $\beta$ ($C$) to the carbonyl group, and three protons at the end of the carbon chain ($A$). A double bond is associated with two olefinic ($J$) and four allylic ($D$) protons, while each additional double bond on the same fatty acid gives two additional
olefinic ($J$) and two diallylic ($F$) protons. The number of remaining methylene protons ($B$) increases with fatty acid chain length.

Knowledge of the mean fatty acid chain length ($CL$), the number of double bonds per triglyceride ($ndb$) and the number of methylene-interrupted double bond pairs ($nmidb$) allows characterization of the triglyceride spectrum [66]. Conversely, knowledge of the triglyceride spectrum enables estimation of $CL$, $ndb$, and $nmidb$, which is described in paper IV.

Since resonance $B$ is much larger than the other resonances, it is commonly assumed that fat has a single resonance at 1.3 ppm.

3.4 Image contrast

Many properties that can be captured with MRI are tissue-specific [43]. This is the foundation for the great success of MRI in medicine. The most important tissue-specific properties remain $\rho$ (proton density), $T_1$, and $T_2$. Depending on the scan parameters, the MR signal amplitude will depend on one or more
The ability to manipulate image contrast through scan parameters can be illustrated by a spin echo sequence with echo time $TE$ repeated with intervals of $TR$. The signal amplitude at the center of the spin echo is given by:

$$y_{SE} \propto \rho \left(1 - e^{-TR/T_1}\right)e^{-TE/T_2}$$  \hspace{1cm} (3.1)$$

The shorter the $TR$, the more $T_1$ dependence of the signal, giving more $T_1$ contrast in the image. Conversely, the longer the $TE$, the more $T_2$ dependence of the signal, giving a more $T_2$-weighted image. If $TR$ is long and $TE$ is short, a proton density weighted image is acquired. If $TR$ is short and $TE$ is long, the image becomes weighted by both $T_1$ and $T_2$, which is not used in practice. Note that tissue with long $T_1$ appears dark in standard $T_1$-weighted images, while tissue with long $T_2$ appears bright in $T_2$-weighted images. Images weighted by proton density, $T_1$, and $T_2$ are shown in fig. 3.5.

At 1.5 T, $T_1$ is typically in the range 0.2–2.0 sec, while $T_2$ is approximately 30–130 msec. Free water such as cerebrospinal fluid has long $T_1$ and $T_2$, while bound water has a shorter $T_2$. Fat has short $T_1$ and long $T_2$.
3.5 Fat suppression

Fat contributes much to the signal intensity in standard $T_1$-weighted images, due to the short $T_1$ of adipose tissue. In $T_2$-weighted sequences, the signal from adipose tissue is also intense, due to the long $T_2$. In addition, adipose tissue has a higher proton density than many other tissues [68], further contributing to the strong signal from fat.

The non-fat signal, which originates predominantly from water protons, often contains clinically valuable information. However, the strong signal from fat may obscure underlying signal of greater medical interest, such as contrast enhanced tissue in $T_1$-weighted images, and edema or inflammation in $T_2$-weighted images. Therefore, some kind of fat suppression is used in most medical exams. Brain exams are an exception of this, since the signal from fat is normally negligible in the brain [69]. The use of fat suppression also avoids chemical shift spatial misregistration artifacts and may reduce motion artifacts. Fat suppression is demonstrated in fig. 3.6.

Two different principles are commonly used for fat suppression; the specific chemical shift of fat and the specific $T_1$ of adipose tissue. There exists a number of techniques to acquire fat suppressed images, which each have their advantages and disadvantages. Reviews of different methods for fat suppression have been made by Kaldoudi and Williams [70], de Kerviler et al. [71], Delfaut et al. [72], and Bley et al. [73].

3.5.1 Short time inversion recovery

Fat suppression by short time inversion recovery (STIR) utilizes the short $T_1$ of adipose tissue [10]. It is a kind of “$T_1$-null method” [70], based on the inversion recovery sequence described in section 2.3.3. After the 180° inversion pulse (at time $t = 0$), the transversal magnetization will be zero as the
net magnetic moment will be rotated to point along the negative z-axis. The longitudinal magnetization (solving eq. 2.11) is given by

$$M_z(t) = M_0(1 - 2e^{-t/T_1})$$

(3.2)

for spins with a common $T_1$. At time $t = T_1 \ln 2$, $M_z$ will be zero. If the inversion pulse is applied to several tissues with different $T_1$, their respective magnetization will be zero at different times. In STIR sequences, the inversion time $TI$ is set so that the magnetization of adipose tissue is zero (about 150 msec at 1.5 T). Then, the pulse sequence proceeds with a spin echo sequence. The adipose tissue will have no magnetization to be excited, while tissues with higher or lower $T_1$ will be excited and start to precess in the transversal plane, eventually forming an echo.

An advantage of STIR over methods based on chemical shift is its insensitivity to static field inhomogeneity. On the other hand, it is not molecule-specific. The $T_1$ relaxation time of fat depends on the molecular environment [74]. Tissue with fatty infiltration or tumors containing fat might not have the same $T_1$ as adipose tissue. On the contrary, tissue with $T_1$ similar to that of adipose tissue will also be suppressed, such as mucoid tissue, hemorrhage, and proteinaceous fluid [72]. In addition, tissue enhanced by contrast agents may, depending on the concentration, obtain a $T_1$ similar to that of adipose tissue. This rules out the use of contrast agents.

After the excitation pulse, the transversal magnetization of long $T_1$ tissue will be $180^\circ$ out of phase compared to the magnetization of tissue with very short $T_1$. Therefore, long and short $T_1$ tissues may appear equally intense in magnitude images. They can be differentiated, however, in the real or imaginary part of the image [71]. Since the longitudinal magnetization of many tissues is low at the time of excitation, STIR images suffer from low signal-to-noise ratio.

Figure 3.6: Axial view of the abdomen. $T_2$-weighted images without (a) and with (b) fat suppression.
The inversion recovery introduces $T_1$ dependence of the signal. In standard $T_1$ weighted images, short $T_1$ tissue appear bright. In STIR images, tissue with long $T_1$ are bright. If $TE$ is long, the image will be weighted by both $T_1$ and $T_2$ (sometimes called $T_2$-STIR). This ‘unconventional’ contrast causes tissue with long $T_1$ and $T_2$ to appear hyper-intense. A $T_2$-weighted image with STIR is shown in fig. 3.1 b.

3.5.2 Fat saturation

Fat saturation is also known as FATSAT or CHESS. In contrast to STIR, fat saturation techniques rely on the molecule-specific chemical shift of fat.

It is assumed that all fat protons are chemically shifted approximately 1.3 ppm, i.e. the chemical shift of the fat bulk methylene protons (resonance B in fig. 3.4). Fat saturation is achieved by applying a selective 90° saturation pulse, tuned to the methylene resonance frequency, before proceeding with a nonselective excitation pulse [75]. The saturation pulse is amplitude modulated to obtain a narrowband frequency response centered around 1.3 ppm, in order not to saturate the water protons, which precess faster and have a chemical shift of 4.7 ppm. After the saturation pulse, a gradient may be applied to spoil the transversal magnetization of the fat protons [76, 77]. In particular, this is necessary if the excitation pulse is smaller than 90°.

Fat saturation can be added to most pulse sequences. In general, narrowband pulses have long duration. Typically, some milliseconds are added to the $TR$. Unlike STIR, fat saturation can be combined with the use of contrast agents. The main disadvantage is the sensitivity to $B_0$ inhomogeneity, which causes resonance frequency shifts of all protons. This may result in local failure of the fat suppression or even inadvertent suppression of water signal. At higher field strengths, inhomogeneity of the rotating $B_1$ field may also be a problem. This can be avoided by using of adiabatic pulses, which are amplitude- and frequency modulated pulses insensitive to $B_1$ inhomogeneity [78].

Even if fat saturation is specific to protons in fat, not all fat signal can be suppressed. For instance, olefinic protons (−CH = CH−; resonance J in fig. 3.4) have a chemical shift of $\approx 5.3$, close to that of water protons.

3.5.3 Hybrid techniques

Some fat suppression techniques combine the principles of fat saturation and STIR, by applying a frequency selective inversion pulse and, after some delay, a nonselective excitation pulse [79]. The inversion pulse is not necessarily 180°, but somewhere between 90° and 180° in order to shorten the time delay before the excitation pulse. These techniques are known as selective partial inversion recovery (SPIR) [79, 80] or spectral inversion at lipids (SPECIAL) [81]. Fat suppression with SPIR is shown in fig. 3.7.
3.5.4 Water excitation

Rather than saturating the fat protons, the idea of *water excitation* is to excite the water protons only. In order to selectively excite water protons simultaneous with slice encoding, *spatial-spectral pulses* can be used [82]. Such pulses combine oscillating slice encoding gradients with specially tailored RF-pulses, allowing single- or multi-slice chemical shift-selective imaging. The duration of these pulses can be substantial [73].

An alternative is to use multiple phase-modulated non-selective excitation pulses with an intermediate delay matching the difference in resonance frequency between water and fat [83]. The simplest example is two 45° excitation pulses separated by a time delay equal to the time required for water and fat protons to get 180° out of phase (in the transversal plane). The phase of the second excitation pulse is modulated to be at 90° angle to the water protons and at $-90°$ angle to the fat protons. As a result, the fat protons will be tilted back to the $z$-axis, while the water protons will have experienced a total excitation pulse of 90°. The pulses can be modified to achieve lower flip angles. Such an excitation scheme can be added to many pulse sequences, adding to the *TR* only the time of 180° dephasing ($\approx 2.3$ msec at 1.5 T). A similar technique is denoted RODEO [84].

Like fat saturation, water excitation is also sensitive to $B_0$ field inhomogeneity, but less sensitive to $B_1$ field inhomogeneity. Water excitation has been reported to be advantageous compared to fat saturation in musculoskeletal imaging [85] and contrast enhanced imaging of the breast [86].

3.5.5 Chemical shift imaging

Fat saturation and water excitation can be said to be chemical shift imaging (CSI) techniques, since the image contrast is based on the chemical shift between water and fat. However, CSI usually refers to techniques producing images with an additional chemical shift dimension. This means that a resonance
frequency spectrum is assigned to each voxel, or that an image is assigned to
each discrete resonance frequency, depending on how the data is displayed.
The image corresponding to the water resonance frequency can be consid-
ered “fat suppressed”. CSI is also known as magnetic resonance spectroscopic
imaging (MRSI).

The first spatially resolved spectroscopy was performed in 1975 by Lauter-
bur et al. [87], who used reconstruction by back-projection. CSI by Fourier
reconstruction was later demonstrated in one [88] and two [89] spatial dimen-
sions, even though the theory allowed for three dimensions [90]. The acquisi-
tion is similar to spin warp imaging, but without the readout gradient. That is,
the signal is sampled with no gradient present, as in spectroscopy. Each spatial
dimension is encoded with slow encoding gradients. Fourier transformation
along the slow spatially encoded dimensions and along the fast chemical shift
dimension gives a four-dimensional spectroscopic image, if three spatial di-
mensions are used. This technique requires the pulse sequence to be repeated
one time per voxel, leading to long acquisition times, especially when encod-
ing three spatial dimensions.

An alternative approach was described by Sepponen et al., who used a read-
out gradient for fast encoding of one of the spatial dimensions, while slowly
encoding the chemical shift dimension [91]. This is possible by keeping a
fixed time between excitation and signal readout (i.e. echo time), while shift-
ing the position of the $180^\circ$-pulse between repetitions, so that the readout
becomes shifted relative to the center of the spin echo. The difference in shift
between repetitions determines the spectral bandwidth, while the number of
shifts determines the spectral resolution. This enables fast acquisition of low
resolution spectra [92].

The main drawback for CSI as a method to obtain images free from fat
signal is the lengthy acquisition time. However, at the same time as Sepponen
et al., Dixon proposed a similar technique allowing fast acquisition of separate
water and fat images [93]. The paper by Dixon can be said to describe the first
model-based water/fat separation by chemical shift imaging, which is the title
of the next chapter.

3.5.6 Other techniques

Except for the fat suppression techniques described above, there exists a few
alternative approaches.

As mentioned in section 2.5, resonance frequency shifts due to the readout
gradient field cannot be distinguished from frequency shifts due to chemical
shift. Therefore, chemically shifted spins become spatially displaced. This en-
ables chemical shift encoding by alternating the direction and amplitude of the
readout gradient [94]. In particular, readout gradients of equal amplitude can
be applied in opposite directions (i.e. with opposite polarity) [95].
Another method utilizes the phenomenon of scalar coupling (J-coupling) to produce separate images of coupled and non-coupled spins, respectively [96]. Since fat spins are coupled, and water spins are not, this method enables separation of water and fat signal.

Further, the phenomenon of magnetization transfer forms a basis for water and fat separation [97]. This technique utilizes that excited protons in membrane phospholipids and proteins will transfer magnetization to water spins, but not to fat spins.

Finally, it has been suggested that contrast between water and fat can be achieved based on differences in diffusivity [98].
4. Model-based water/fat separation by chemical shift imaging

In the simple chemical shift imaging method originally proposed by Dixon [93], the signal is recorded while applying a readout gradient. Just like in the method by Sepponen et al. [91], the spectral dimension is sampled by shifting the position of the 180°-pulse. However, in contrast to the Sepponen method, Fourier transformation is only applied along the spatial dimensions. Additionally, the spectral dimension is only sampled at two well-chosen points; one where the readout gradient echo coincides with the spin echo, and one where the readout is shifted relative to the spin echo with an interval corresponding to the time needed for water and fat protons to get 180° out of phase. The Fourier transformed data can be displayed as two images; one in-phase image (IP) where the voxel brightness reflects the added transversal magnetization of all chemical components, and one opposed-phase image (OP) where each voxel reflects the difference between water and fat magnetization. Separate water and fat images can then be obtained by a simple post-processing step: adding the IP and OP images results in a water image, while subtracting the OP image from the IP image gives a fat image [93].

In- and opposed-phase images are shown in fig. 4.1. In conventional spin echo images, the signal contributions from all chemical species are in phase. In the OP image, water-dominated and fat-dominated tissues are clearly delineated as a result of the partial volume effect. This means that voxels at the interface of such tissues contain substantial amounts of both water and fat pro-

![Figure 4.1: Gradient echo recalled images shifted so that the water and fat protons are in phase (a) and 180° opposed phase (b).](image-url)
tons, which cancel out in the OP image resulting in low intensity. This is also called chemical shift artifact of second type or india ink artifact [99].

Even if Dixon’s method may seem simplistic, it has several qualities. It is basically assumed that the proton spectrum has two discrete spectral lines (which is a reasonable first-order approximation). This sparse spectrum model allows heavy under-sampling of the spectral dimension; only two points are needed if chosen wisely. This enables high spatial resolution to be maintained with short acquisition time. Achieving high spectral resolution only to sample a two-component spectrum can be regarded as inefficient sampling.

The great potential of Dixon’s method is accompanied with a number of challenges, not least to handle phase errors due to static field inhomogeneity. Therefore, a large number of extensions and modifications have been proposed. Important advances cover acquisition strategies, number and timing of the spectral samples, more accurate signal modeling, and algorithms to correct for phase errors.

In this thesis, the term fat-water imaging (FWI) is used for model-based separation of water and fat signal from chemical shift encoded data. Such methods may also be referred to as Dixon techniques [100] or chemical shift based water-fat separation [73]. In addition, other names have been used for specific FWI algorithms. Some examples are DPE [101], CSISM [102], and IDEAL [103].

4.1 Signal model

The spectral dimension is sampled by applying the readout gradient at time shift $t$ relative to a point $t = 0$ where all magnetization is in phase. For spin echo acquisitions, $t = 0$ at the center of the spin echo. If the 180°-pulse is applied at time $\tau$ after excitation, the center of the spin echo occurs $2\tau$ after excitation. For gradient echo acquisitions, $t = 0$ at excitation so that $t = TE$.

\[ \alpha^\circ \]

Transmit $G_{fast}$

Receive

$\underline{t = TE}$

*Figure 4.2:* In gradient echo acquisitions, the signal is in phase directly after excitation, so that the time shift $t$ equals the echo time $TE$. 

52
Therefore, $t$ is sometimes referred to as the echo time, although time shift would be a more correct term in the case of spin echo acquisitions.

The time shifts of spin echo and gradient echo acquisitions is illustrated in figs. 4.2 and 4.3. In spin echo acquisitions, a time shift $t$ can be obtained in two ways: either by shifting $\tau$ and keeping $TE$ constant, or by shifting $TE$ and keeping $\tau$ constant.

In all studies described in this thesis, the data was acquired using gradient recalled sequences.

4.1.1 Naive model

A model of the signal $y$ in a voxel, representing how water and fat interfere as a function of the time shift $t$, is given by:

$$y(t) = W + Fe^{i\omega_B t}$$ (4.1)

where $W$ and $F$ are the water and fat signals (at $t = 0$), and $\omega_B = \gamma B_0(\delta_B - \delta_W)$ is the difference in resonance frequency between water and fat due to chemical shift (see table 3.3 on page 43). Note that $\omega_B$ is known a priori and common to all pixels, while $y(t)$, $W$, and $F$ are spatially dependent.

In this model, the fat signal is assumed to have a single resonance frequency (resonance B) which is chemically shifted 3.4 ppm relative to water.

Adopting this model, the water and fat magnetization will be in phase every $t = n \times 4.6$ msec and exactly opposed phase every $t = n \times 4.6 + 2.3$ msec (assuming $B_0 = 1.5$ T).

In the conventional two-point method, one image is acquired in phase and one opposed phase:

$$y_{IP} = W + F$$
$$y_{OP} = W - F$$ (4.2)
Figure 4.4: Water (a) and fat (b) images reconstructed using the naive signal model. In the water image, leakage of fat signal is seen in the subcutaneous adipose tissue. Likewise, the water signal of muscle tissue leaks into the fat image.

Water and fat can then easily be separated voxel by voxel, according to:

\[
2W = y_{IP} + y_{OP} \\
2F = y_{IP} - y_{OP}
\]  

(4.3)

This signal model is naive, since it assumes that the image phase is due to chemical shift only. In practice, locally varying phase errors are always present, causing leakage of water signal into the fat image and vice versa [104, 105]. Water and fat images reconstructed using this method are shown in fig. 4.4.

Outside the scope of FWI, in phase and opposed phase images can also be useful in themselves [93, 99].

4.1.2 Simple model

Water and fat separation using the naive signal model results in unacceptable signal leakage between the water and fat images. To avoid this, phase errors can be incorporated into the signal model. This allows estimation and removal of the phase errors prior to separation of water and fat signal. Phase errors occur because the resonance frequency will be offset from the nominal frequency by an unknown frequency \( \omega \) that varies from voxel to voxel. Such off-resonance is caused by inhomogeneity in the \( B_0 \) field, including effects of imperfections in the applied field, and tissue susceptibility (although no specific assumption of the underlying causes are required).

Incorporating the off-resonance in the signal model, it becomes:

\[
y(t) = (W + Fe^{i\omega_B t})e^{i\omega t}
\]  

(4.4)

In contrast to \( \omega_B \), \( \omega \) is not known \textit{a priori}, and is spatially varying. In Dixon’s original method, \( e^{i\omega t} \) was removed by taking the absolute value of
the IP and OP images before addition and subtraction [93]. This approach is only adequate for voxels where \( W > F \), since the water image will always contain the largest of the two signal components, and the fat image will always contain the smallest [106].

A better strategy is to estimate \( \omega \) in each voxel. As a side effect, a map of the off-resonance (\( B_0 \) inhomogeneity) will be provided along with the water and fat images [104, 106]. Estimation of the off-resonance is simplified by increasing the number of spectral samples (echo times) [106]. A common approach is to obtain three complex samples of the spectral dimension, i.e. three time shifts. This solution, often called the three-point Dixon method, was first introduced by Kim et al. [104]. Images from a three-point acquisition are shown in fig. 4.5. Kim et al. [104], Lodes et al. [105], and Glover and Schneider [107] used the time shifts \( t = -T, t = 0, \) and \( t = +T \), where \( T = \pi/\omega_B \). This gives two opposed-phase and one in-phase image:

\[
\begin{align*}
y_{-T} &= (W - F)e^{-i\omega T} \\
y_0 &= (W + F) \\
y_{+T} &= (W - F)e^{+i\omega T}
\end{align*}
\]  

(4.5)

The off-resonance phase vector 
\[
e^{i\omega T} = \pm \sqrt{\frac{y_{+T}}{y_{-T}}}
\]  

(4.6)

can be analytically determined and removed from \( y_{+T} \) by division, or from \( y_{-T} \) by multiplication. Then, water and fat images can be obtained by adding and subtracting \( y_0 \) and either \( y_{+T} \) or \( y_{-T} \) [104].

Equation 4.6 gives two solutions for the off-resonance, of which only one can be correct. Choosing the false solution still gives a correct separation of water and fat signal, but an incorrect identification, meaning that the water signal will be interpreted as fat and the fat signal will be taken for water. In the water and fat images, this manifests as swap artifacts. It can be seen in

Figure 4.5: Magnitude of complex images acquired using a gradient recalled triple-echo acquisition at 1.5 T. The echo times were (a:) 3.1 msec, (b:) 6.0 msec, and (c:) 8.7 msec.
eq. 4.5 that adding multiples of $\pi/T$ to $\omega$ has the same effect as swapping $W$ and $F$. The off-resonance ambiguity problem is discussed in section 4.3. Images reconstructed using a three-point Dixon method with estimation of the off-resonance map are shown in fig. 4.6.

The ($-T, 0, T$) spectral sampling scheme of the three-point Dixon method was generalized by Glover to ($-\Delta t, 0, \Delta t$) or (0, $\Delta t, 2\Delta t$) for arbitrary $\Delta t$ [108], and further by Xiang and An to ($t_0, t_0 + \Delta t, t_0 + 2\Delta t$) for arbitrary $t_0$ and $\Delta t$ [101]. An and Xiang generalized to an $N$-point (0, $\Delta t$, ..., $(N-1)\Delta t$) sampling scheme [102], and Reeder et al. provided a solution for the general case were all the time shifts are arbitrary [109].

In paper I, the signal model given by eq. 4.4 was used, and the signal was sampled using a ($t_0, t_0 + \Delta t, t_0 + 2\Delta t$) sampling scheme.

Estimating complex-valued $W$ and $F$, and the real-valued $\omega$ from a complex triple-point acquisition, poses an over-determined problem, which may be solved in the least-squares (LS) sense. LS estimation in this context was first proposed by An and Xiang [102]. It is reasonable to use LS estimation, since it is the maximum likelihood estimator for Gaussian noise, which is typically assumed in complex MR images [110].

The signal model in eq. 4.4 can be put on matrix form, assuming $N$ time shifts $t_1, ..., t_N$:

$$y = BA \begin{bmatrix} W \\ F \end{bmatrix}$$

(4.7)

where $y = \begin{bmatrix} y(t_1) & \cdots & y(t_N) \end{bmatrix}^T$, 

$$B(\omega) = \begin{bmatrix} e^{i\omega t_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{i\omega t_N} \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} 1 & e^{i\omega t_1} \\ \vdots & \vdots \\ 1 & e^{i\omega t_N} \end{bmatrix}$$

(4.8)
The signal model is *separable* with linear terms $W$ and $F$ [111]. Provided the nonlinear term $\omega$, the least squares solution for $W$ and $F$ is given by:

\[
\begin{bmatrix}
W_{LS} \\
F_{LS}
\end{bmatrix} = A^+ B^H y
\]  

(4.9)

where $A^+ = (A^H A)^{-1} A^H$ is the pseudo-inverse and $^H$ denotes the conjugate transpose. The pseudo-inverse is common for all voxels and needs to be calculated only once. Section 4.3 describes different methods to determine $\omega$ in every voxel.

### 4.1.3 Simultaneous estimation of $T_2^*$ relaxation

In the simple signal model given by eq. 4.4, it is assumed that no relaxation occurs during the time shifts, i.e. $t \ll T_2^*$. This is a fair assumption in most cases. However, some tissues, such as cortical bone, have extremely short $T_2^*$. Tissue $T_2^*$ can also be shortened, for instance due to iron overload in the liver [112, 113]. In these cases, $T_2^*$ may be incorporated into the signal model. This results in a more correct water and fat estimation, and jointly provides a $T_2^*$-map, which may be useful in itself.

A simple approach to estimate $T_2^*$ along with water and fat was first described by Glover [108]. A more general method was presented by Yu et al. [114], who used the following signal model:

\[
y(t) = (W + Fe^{i\omega b}) e^{i(\omega - R_2^*) t}
\]  

(4.10)

where the decay $R_2^* = 1/T_2^*$ is assumed to be equal for the water and fat magnetization. Now, two nonlinear terms ($R_2^*$ and $\omega$) must be estimated in each voxel before solving for water and fat through eq. 4.9 using the modified model matrix:

![Figure 4.7: Axial slice across the liver. Water (a), fat (b), and $R_2^*$ (c) images were reconstructed simultaneously. Since the $R_2^*$ is quantitative, the gray scale bar indicates the rate of decay. Fast decay is seen in the intestine, and slow decay in the gall bladder and the aorta.](image)
A = \begin{bmatrix}
e^{−R_2^*t_1} & e^{(i\omega_B−R_2^*)t_1} \\
\vdots & \vdots \\
e^{−R_2^*t_N} & e^{(i\omega_B−R_2^*)t_N}
\end{bmatrix} \quad (4.11)

Now, $A^+$ varies from voxel to voxel, since it depends on $R_2^*$. An easy way to handle this is to discretize $R_2^*$ on a sufficiently fine grid and pre-calculate one $A^+$ for each discrete $R_2^*$ value.

$R_2^*$ can be found by minimization of the residual function, which will be described in more detail in section 4.3.1. The minimization with respect to $R_2^*$ is not associated with the same difficulties as $\omega$ is [115].

Since more parameters are to be estimated, it is beneficial to use more than three spectral samples. Typically, at least six points are used [114]. Fig. 4.7 shows water, fat, and $R_2^*$ images, reconstructed from a six-echo acquisition.

It has also been proposed to estimate independent $T_2^*$ values for water and fat [116].

This signal model is only valid for gradient echo acquisitions. In spin echo acquisitions, dephasing due to $T_2$ relaxation occurs during $TE$ (see fig. 4.3), while the time shift $t$ induces reversible relaxation due to $T_2'$ (see eq. 2.13).

### 4.1.4 Modeling multiple fat resonances

The assumption, that fat has a chemical shift of 1.3 ppm, is only valid for $\approx 60\%$ of the fat protons. Other fat protons may have a similar chemical shift such as 0.9 ppm or 2.0 ppm, but can also be found at 5.3 ppm, i.e. closer to the water resonance [65]. A table of the chemical shifts of fat is found on page 43.

A signal model that accounts for the multiple resonances of fat was proposed by Yu et al. [117]. The fat protons are assumed to be found at $M$ resonances $\omega_m = \gamma B_0(\delta_m − \delta_W)$ in relative amounts $\alpha_m$, normalized so that the $\alpha_m$ add up to unity. The model is given by:

$$y(t) = (W + F \sum_{m=1}^{M} \alpha_m e^{i\omega_m t})e^{(i\omega−R_2^*)t} \quad (4.12)$$

if used simultaneously with $T_2^*$ estimation. Note that $\omega_m$ and $\alpha_m$ are assumed to be known a priori and are common for all voxels. Thus, the multiple fat resonance model does not increase the number of parameters to be estimated, and comes at no expense but (slightly) increased complexity of the reconstruction algorithm. The values of $\omega_m$ and $\alpha_m$ may be found in the literature, or determined from spectroscopy measurements. The $\alpha_m$ values can also be...
determined from the data itself, using a self-calibration algorithm [117]. With this signal model, the model matrix becomes:

\[
A = \begin{bmatrix}
    e^{-R_2^* t_1} & \sum_{m=1}^{M} \alpha_m e^{(i\omega_m - R_2^*) t_1} & \cdots \\
    \vdots & \vdots & \ddots \\
    e^{-R_2^* t_N} & \sum_{m=1}^{M} \alpha_m e^{(i\omega_m - R_2^*) t_N}
\end{bmatrix}
\]  

(4.13)

Taking the multiple resonances of fat into account increases the accuracy of the model. This results in stronger suppression of fat signal in the water images [118] and more accurate estimates in quantitative applications [119]. The effect of multiple fat resonances is illustrated in fig. 4.8.

Studying this more accurate signal model, it appears that there are no perfectly opposed-phase images. Similarly, the only true in-phase image is obtained by setting \( t = 0 \).

This signal model was used in paper III.

\[\text{Figure 4.8: Axial slice across the liver from a triple-echo acquisition. The water images are reconstructed using both (a:) a single fat resonance and (b:) multiple fat resonances. The multiple fat resonance model achieves stronger and more uniform suppression of the fat signal, which can be seen in the subcutaneous adipose tissue (arrows).}\]
4.1.5 Four-component fat spectrum model

With reference to section 3.3, at least ten resonances are associated with fat. However, the fat spectrum can be characterized by only four components being linear combinations of the ten resonances. The first component is common to all triglycerides, regardless of fatty acid chain length or saturation degree. The second component represents bulk methylene, and is a function of the chain length. The third component is associated with the first double bond on each fatty acid, and the fourth component represents the additional double bonds. Introducing $a_m = e^{i\omega_m t}$ for resonance $m$, the four components are described by:

$$
\begin{align*}
 a_{F1} & = 9a_A + 6a_C + 6a_E + 2a_G + 2a_H + a_I \\
 a_{F2} & = 2a_B \\
 a_{F3} & = 4a_D + 2a_J \\
 a_{F4} & = 2a_F + 2a_J
\end{align*}
$$

The signal model can now be expanded by dividing $F$ into four components:

$$
y(t) = (W + F_1a_{F1} + F_2a_{F2} + F_3a_{F3} + F_4a_{F4})e^{(i\omega - R^*_{21})t}
$$

which gives the associated model matrix:

$$
A = \begin{bmatrix}
    e^{-R^*_{21}t_1} & a_{F1}e^{-R^*_{21}t_1} & a_{F2}e^{-R^*_{21}t_1} & a_{F3}e^{-R^*_{21}t_1} & a_{F4}e^{-R^*_{21}t_1} \\
    \vdots & \vdots & \vdots & \vdots & \vdots \\
    e^{-R^*_{2N}t_N} & a_{F1}e^{-R^*_{2N}t_N} & a_{F2}e^{-R^*_{2N}t_N} & a_{F3}e^{-R^*_{2N}t_N} & a_{F4}e^{-R^*_{2N}t_N}
\end{bmatrix}
$$

Since several parameters are to be estimated in each voxel, many echoes need to be acquired (six at minimum). Conveniently, all the fat components are linear in the signal model. However, components $F_1$, $F_3$ and $F_4$ are typically of low signal strength (see fig. 4.9). The images are taken from the study described in paper IV, in which this signal model was developed.

![Figure 4.9: Images of an oil-water phantom. Simultaneous estimation of water (a) and the fat components $F_1$ (b), $F_2$ (c), $F_3$ (d), and $F_4$ (e) from a 32-echo dataset.](image)
Techniques similar to the one described here were outlined in conference abstracts by Peterson et al. [120], and later by Peterson and Månsson [121] and by Bydder et al. [122].

4.2 Modified two-point Dixon techniques

Compared to three-point techniques, two-point datasets can in principle be acquired faster. Although, resolving the off-resonance phase errors has been reported to be more challenging [100, 123].

The first modification of Dixon’s original two-point technique was introduced by Kim et al. [104], by using complex rather than magnitude IP/OP images (see section 4.1.1). This technique is problematic, since it does not account for off-resonance phase errors [104]. Resulting water and fat images can be seen in fig. 4.4.

A more powerful modification of the two-point method taking phase errors into account was described in 1987 by Borrello et al. [124] and in 1988 by Brix et al. [125]. Similar methods were described ten years later by Skinner and Glover [123], and Coombs et al. [126]. These approaches use the simple model given by eq. 4.4, and acquire images in-phase and opposed-phase. Eq. 4.2 then becomes:

\[
y_{IP} = W + F \\
y_{OP} = (W - F)e^{i\omega T}
\]  

(4.17)

Estimating two complex-valued \(W \) and \(F\) and one real-valued \(\omega\) parameter from two complex valued samples \(y_{IP}\) and \(y_{OP}\) is an under-determined problem. However, one degree of freedom can be removed from the model, since water and fat should have a common phase \(\phi\) at \(t = 0\) [127]. This can be imposed by introducing \(W = we^{i\phi}\) and \(F = fe^{i\phi}\), where \(w\) and \(f\) are real-valued. Eq. 4.17 can be written:

\[
y_{IP} = (w + f)e^{i\phi} \\
y_{OP} = (w - f)e^{i(\phi + \omega T)}
\]  

(4.18)

Now, four real-valued parameters are to be estimated from two complex-valued samples. Since \(w\) and \(f\) are real-valued, \(e^{i\phi}\) represents the phase of \(y_{IP}\), and can be removed directly from \(y_{IP}\) and \(y_{OP}\). Likewise, \(\pm e^{i\omega T}\) represents the phase of \(y_{OP}/e^{i\phi}\), the sign depending on whether \(w\) or \(f\) is the larger component. This ambiguity resembles eq. 4.6 and can be resolved using the techniques described in section 4.3. When determined, this off-resonance phase vector is removed prior to addition and subtraction to obtain the water and fat images. Images reconstructed using this technique are shown in fig. 4.10.
4.2.1 Relaxing one time constraint

Xiang generalized the two-point technique to allow the time shifts $t_1$ and $t_2$, where $y_1 = y(t_1)$ is acquired in-phase and $y_2 = y(t_2)$ partially opposed phase (POP) [128]. After removing $e^{i\phi}$ given by the phase of $y_1$, the off-resonance phase vector $e^{i\omega(t_2-t_1)}$ can be determined analytically according to eq. 4.23.

Defining $y_1^*$ and $y_2^*$ as the signals after removing the phase errors, the real-valued matrices corresponding to the definitions in eq. 4.8 become

$$
y^* = \begin{bmatrix}
R(y_1^*) \\
I(y_1^*) \\
R(y_2^*) \\
I(y_2^*)
\end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix}
1 & R(a_1) \\
0 & I(a_1) \\
1 & R(a_2) \\
0 & I(a_2)
\end{bmatrix}
$$

(4.19)

where $R$ and $I$ represent taking the real and imaginary parts, respectively, and $a_n = e^{i\omega t_n}$. The LS estimates for $w$ and $f$ are given by:

$$
\begin{bmatrix}
w_{LS} \\
f_{LS}
\end{bmatrix} = A^+ y^*
$$

(4.20)

4.2.2 Relaxing both time constraints

In the general case of two time shifts without constraints, eq. 4.18 becomes:

$$
y_1 = (w + a_1 f)e^{i(\phi + \omega t_1)}
$$

$$
y_2 = (w + a_2 f)e^{i(\phi + \omega t_2)}
$$

(4.21)

---

**Figure 4.10**: Axial slice at the level of the kidneys. Water (a) and fat (b) images reconstructed using a modified two-point Dixon method. Uniform separation of water and fat signal is seen over the entire field of view.
Introducing the two phase vectors \( e^{i\theta} = e^{i(\phi + \omega t_1)} \) and \( e^{i\psi} = e^{i\omega(t_2 - t_1)} \), the phase errors can be removed from eq. 4.21 once \( e^{i\theta} \) and \( e^{i\psi} \) are known, in order to obtain \( \psi^* \). Then \( w \) and \( f \) can be found using eq. 4.20.

It can be shown that the fat signal fraction \( FF = f/(f+w) \) may be determined as:

\[
FF = \frac{c_1 \pm \sqrt{c_3}}{c_1 + c_2}
\]

where

\[
c_1 = |y_1|^2 (1 - \Re(a_2)) - |y_2|^2 (1 - \Re(a_1))
\]
\[
c_2 = |y_1|^2 (|a_2|^2 - \Re(a_2)) - |y_2|^2 (|a_1|^2 - \Re(a_1))
\]
\[
c_3 = |y_1|^2 |y_2|^2 |a_1 - a_2|^2 (\Im(a_1)|y_2|^2 - \Im(a_2)|y_1|^2)^2
\]

It now follows from eq. 4.21 that \( e^{i\psi} \) can be determined as:

\[
e^{i\psi} = \frac{y_2(1 + FF(a_1 - 1))}{y_1(1 + FF(a_2 - 1))}
\]

and that

\[
\theta = \frac{\angle y_1 (1 - a_2) - y_2 (1 - a_1)/e^{i\psi}}{a_1 - a_2}
\]

The ± ambiguity of eq. 4.22 gives two \( e^{i\psi} \) candidates through 4.23. Again, the process of choosing between these alternatives is described in section 4.3.

This flexible two-point Dixon method was described in paper II.

To efficiently remove noise from the resulting images, the smoothness assumption can be imposed by low-pass filtering of the \( e^{i\psi} \)-map and subsequently of the \( e^{i\theta} \)-map, prior to solving for water and fat through eq. 4.20.

It should be emphasized that \( a_1 \) and \( a_2 \) can be any arbitrary \textit{a priori} known complex numbers, equations 4.22–4.24 still being valid. For instance, the multiple resonance fat spectrum model with \( T^*_2 \) relaxation can be used with two-point methods by letting \( a_n = \sum_{m=1}^{M} \alpha_me^{(i\omega_m - R^*_2)n} \). Since there are not enough samples to estimate \( R^*_2 \), it must be assumed known \textit{a priori}. In paper II, the simpler model \( a_n = e^{(i\omega - \nu)n} \) was used, where \( \nu \) represents the spectral broadening of fat [129].
4.3 The off-resonance problem

This section deals with the determination of $\omega$, which is one of the major obstacles in FWI, since it is more or less ambiguous. This ambiguity is reflected in eqs. 4.6, 4.22, and 4.25. The ambiguity problem is best illustrated by the simple signal model (eq. 4.4). With this model, it is not possible to differentiate a voxel of only water from a voxel of only fat, since the observed resonance can be explained either by off-resonance or by chemical shift. Using the multiple fat resonance signal model (eq. 4.12), the water and fat spectral components are no longer identical, which should in principle resolve the ambiguity. In practice however, the ambiguity problem remains due to noise and model imperfections.

Some equations have already been given for analytical determination of the off-resonance phase factor $e^{i\psi}$ for the special cases of the two-point method (eq. 4.23) and the three-point method using the $(-T, 0, T)$ acquisition scheme (eq. 4.6). Analytical solutions for other cases have been described elsewhere [101, 130, 131, 132, 133].

A general solution for the three-point ($t_0, t_0 + \Delta t, t_0 + 2\Delta t$) acquisition scheme (referred to as constant time shift spacing) is given by:

$$\psi = \frac{y_2d_1 \pm \sqrt{y_2^2d_1^2 - 4y_1y_3d_2d_3}}{2y_1d_3} \quad (4.25)$$

where

\begin{align*}
    d_1 &= a_3 - a_1 \\
    d_2 &= a_2 - a_1 \\
    d_3 &= a_3 - a_2
\end{align*}

The argument operator $\mod$ of eq. 4.25 gives only a wrapped off-resonance $\omega = \psi/\Delta t$, i.e. it is only known within a principal period of $2\pi/\Delta t$ [134]. The wrapped $\omega$ can still be used to determine the $B$ matrix of eq. 4.9 to obtain the LS water and fat estimates. The choice of period will only affect the phase common to $W$ and $F$, and has no practical implication. Thus, estimating the wrapped off-resonance is sufficient for accurate separation of water and fat [100, 101, 124].

In paper I, a special case of eq. 4.25 was used to determine the off-resonance, valid only for the simple signal model in eq. 4.4. In contrast, eq. 4.25 is valid for a wider class of signal models, such as the multiple resonance fat model (eq. 4.12).
4.3.1 Residual function

Rather than using analytical solutions, the model fit residual can be used to find $\omega$ candidates. This solution is particularly useful when more than three time shifts are used, since the estimation problem becomes over determined.

Re-substituting $W_{LS}$ and $F_{LS}$ given by eq. 4.9 into eq. 4.7 enables calculation of the squared error residual $J$ as a function of $\omega$ only [102, 134, 135]:

$$J(\omega) = \| y - BA \begin{bmatrix} W_{LS} \\ F_{LS} \end{bmatrix} \|$$

(4.26)

where $\| \cdot \|$ denotes the L$_2$ norm. Simplifying eq. 4.26 gives:

$$J(\omega) = \| (I - AA^+) B^H y \|$$

(4.27)

This corresponds to a variable projection of $W$ and $F$, leaving only the nonlinear variable $\omega$ as a target for optimization [111, 130]. Inserting the $\omega$ that minimizes $J(\omega)$ into eq. 4.9 (through the $B$ matrix) gives the total LS solution for $W$ and $F$.

Fig. 4.11 shows the residual function $J(\omega)$ from several voxels in a six-echo dataset. Several local minima are present, reflecting the signal ambiguity. For the case of constant spacing $\Delta t$ of the time shifts, it can be shown that $J(\omega)$ becomes periodic in $\omega$ with a period of $\Omega = 2\pi/\Delta t$ [134]. In such a scenario, $J(\omega)$ needs to be considered in one period only [102]. This simplifies the problem by effectively reducing the solution space and avoiding non-trivial phase unwrapping [100]. The data in fig. 4.11 were acquired with constant echo time spacing. Thus, a single period of the residual function is shown for each voxel. It can be noted that two of the local minima within a period are typically much smaller than the other minima. This seems to be a consequence of the two-component model. It can be shown by simulation experiments that the multitude of local minima cannot be attributed to noise, multiple resonance fat modeling, or $T_1^*$ modeling. Rather, there is a connection to the number of acquired time shifts, $N$. Expanding on the calculations presented by Doneva et al. [136], it can be shown that the residual function has at maximum $N - 1$ minima within a period.

There are several ways to find the local minima of $J(\omega)$. For three-point acquisitions with constant time shift spacing, there are at maximum two local minima within a period, which can be found analytically [136]. In more general cases, the minima may be found by optimization such as Gauss-Newton [109] (requiring an initial guess) or golden section search [134] (requiring an initial search interval). The simplest method is probably to evaluate $J(\omega)$ at discrete point [102, 115, 130]. The local minima can then be identified by
Figure 4.11: Plots of the residual $J$ as a function of the off-resonance $\omega$ in seven voxels at different anatomical sites. The number of time shifts was $N = 6$.

comparing each point with its adjacent points. This strategy was used in papers III and IV.

Once the off-resonance map is determined, the water and fat separation can be done in each voxel separately, through eq. 4.9. However, determination of the off-resonance map requires the signal ambiguity to be resolved. To do this, some additional assumption must be made, for instance that the off-resonance is close to zero. The most common assumption for resolving the ambiguity is that the $B_0$ field, and hence the off-resonance map, varies smoothly across the field of view.

### 4.3.2 Voxel independent methods

These methods do not exchange information between voxels, but rely on resolving the ambiguity in each isolated voxel with the aid of some additional assumption.

The simplest strategy is to use only magnitude data, as originally proposed by Dixon [93]. This effectively eliminates the need for estimating any phase errors due to off-resonance, but also the possibility to distinguish water from fat as discussed in section 4.1.2. Nevertheless, this strategy has been used along with more advanced signal modeling for liver fat quantification, assum-
ing that $W > F$ in all voxels [129]. This approach has been shown to give more bias and noise in the estimates as compared to the use of complex data [137].

Another strategy is to assume that the off-resonance is close to zero. It is then reasonable to choose the off-resonance candidate being closest to zero, or the off-resonance phase vector $e^{i\psi}$ with the smallest phase. The candidates may have been obtained analytically, or may correspond to local minima in the residual function. A similar result is achieved by performing a line search minimization of the residual function, initialized at zero off-resonance [109]. This strategy would work fine for the voxels examined in fig. 4.11. The assumption of a small off-resonance is typically only valid for a small field of view. It is also necessary that the time shifts are chosen so that the false off-resonance candidates become well separated from the true candidate.

A third strategy of voxel independent off-resonance identification is to assume a small difference in phase between $W$ and $F$. This is a valid assumption, since $W$ and $F$ are defined as the water and fat magnetization vectors at $t = 0$, i.e. they should be in-phase [127]. If one of the components is below the noise level, their relative phase will be random. Therefore, this strategy only works for voxels with signal contribution from both water and fat protons. It is necessary to choose the time shifts so that the relative water and fat phase corresponding to false candidates is distinct from zero. For three-point acquisitions, this is fulfilled only for asymmetric time shifts, meaning that the first and third points must not be shifted symmetrically about an in-phase or opposed-phase point [101]. This strategy is used in paper I to identify reliable seed points.

Yet another voxel-independent strategy would be to choose the off-resonance candidate with the smallest residual $J$ [138]. In voxels with only water or only fat, this strategy will not work using the single fat resonance signal model. Even using the multiple fat resonance model, which was done in fig. 4.11, this strategy would only be able to accurately identify three of the seven highlighted voxels.

It has also been proposed to resolve the ambiguity by identifying water or fat as the dominant component, based on differences in $T_1$ relaxation acquired using double flip angles [139].

### 4.3.3 Region growing

The strategies for resolving the off-resonance ambiguity presented in this section and the following, are based on the assumption that the off-resonance map is spatially smooth.

The idea of region growing is to expand the region of resolved off-resonance one voxel at a time, starting from one or more seed points, until the entire image is resolved. The determination of the off-resonance in the present voxel is based on the off-resonance of previously resolved neighbor voxels.
Region growing in the context of FWI was first proposed by Borrello et al. [124]. Szumowski et al. described a 2D region growing scheme with one manually resolved seed voxel [140]. After a voxel has been resolved, its four closest neighbors are put on a voxel stack. The region growing continues by picking a voxel from the stack and picking the $\omega$ candidate closest to the resolved $\omega$ of the ‘parent’ voxel. If the $\omega$ difference to the parent voxel exceeds a certain threshold, that voxel is not put on the voxel stack. This scheme was originally proposed for the three-point Dixon method, but has also been used with the two-point method [126].

![Figure 4.12: Illustration of the multi-seeded region growing algorithm described in paper I. (a:) Calculate the quality factor $Q$ in all neighbor voxels. The voxel with the highest $Q$ is selected. (b:) Determine $\epsilon^\psi$ in the selected voxel as the candidate being most phase coherent with the already determined neighbor voxels. (c:) Update the quality factors and add new voxels to the queue. (figure concept: Henric Rydén)
Another region growing scheme for the two-point Dixon method was proposed by Ma [141]. This scheme considers two candidates of the off-resonance phase vector $e^{i\psi}$ in each voxel, instead of determining the absolute off-resonance. This binary choice is based on previously resolved voxels in a small area close to the present voxel, rather than a single parent voxel. Additionally, multiple voxel stacks are used, each corresponding to a certain range of the off-resonance phase difference compared to the neighbors. Voxels that are phase coherent with their neighbors are grown first, allowing confident regions to be grown before regions with large phase differences.

Yet another region growing scheme was described by Yu et al. [135]. Here, a Gauss-Newton minimization of the residual function is performed in each voxel. The initial guess is found by fitting a plane to already resolved voxels in a neighborhood. The seed voxel is picked automatically near the image centre, and minimized with an initial guess of zero off-resonance. The rest of the voxels are resolved along a predestinated square-spiral trajectory.

Region growing techniques in combination with multi-resolution coarse-to-fine propagation of the off-resonance map has been described both for the two-point [142] and the three-point methods [134]. Lu and Hargreaves [134] proposed single-seed region growing along a predestinated path at low resolution. The obtained off-resonance map is subsequently propagated to finer resolution levels, and refined at each level on voxel basis. A scheme of multiple single-seed region growing has also been proposed [102].

In paper I, a region growing scheme is described for the three-point Dixon method, that can be viewed as an advanced version of that proposed by Ma [141]. Since three samples are available, voxels containing both water and fat may be resolved. Such voxels are identified by thresholding operations and used as seeds in a 3D multi-seeded region growing scheme. Voxels are resolved by picking the $e^{i\psi}$ candidate out of two being most phase coherent with the already resolved neighbor voxels. The region growing is guided by a quality measure $Q$, which equals the absolute difference in phase coherence between the two candidates. This means that voxels with one coherent and one incoherent candidate are assigned high quality. If both candidates are equally coherent, the choice is not obvious. Such voxels are assigned a low quality. Neighbors of resolved voxels are put in one of several queues based on the $Q$ value. The next voxel to be resolved is picked from the non-empty queue with the highest $Q$ value. This scheme is repeated until all queues are empty and the entire off-resonance map is determined. This region growing scheme is illustrated in fig. 4.12.

### 4.3.4 Whole-image optimization

A sophisticated solution to the problem of resolving the off-resonance map is to formulate an energy function based on every voxel in the image, so that the
minimum energy corresponds to an off-resonance map that is spatially smooth and fits the data well in each voxel. This is also a means of regularizing the off-resonance map. The meaning of regularization here is to prevent data over-fitting by utilizing the prior knowledge of spatial smoothness, with the goal to decrease noise in the estimates. The whole-image minimization approach in the context of FWI was first described by Hernando et al. [130].

Introducing \( \mathcal{V} \) as the set of all voxels and \( \mathcal{N} \) as the set of neighbor voxel pairs, such an energy as a function of the off-resonance map \( \omega \) can be written:

\[
E(\omega) = \sum_{q \in \mathcal{V}} D_q(\omega_q) + \mu \sum_{(p,q) \in \mathcal{N}} w_{p,q} V_{p,q}(\omega_p, \omega_q)
\]

(4.28)

where \( D_q \) represents the data fidelity of voxel \( q \) (data cost), and \( V_{p,q} \) is some neighborhood consistency energy function (discontinuity cost). The weights \( w_{p,q} \) control the relative impact of the different neighbor pairs, and \( \mu \) controls the amount of regularization or, equivalently, the balance between data fidelity and spatial smoothness.

This kind of energy function resembles that of a Markov Random Field (MRF), where the energy equals the negative logarithm of the joint likelihood of the hidden variables \( \omega \). The minimum energy of eq. 4.28 corresponds to the estimate of \( \omega \) with maximum a posteriori likelihood. Markov Random Field models have previously been used in MRI for the related problem of phase unwrapping [143].

Hernando et al. [130], used \( D_q = J_q(\omega_q) \) (see eq. 4.27), \( V_{p,q} = |\omega_p - \omega_q|^2 \), and \( w_{p,q} = 1/d(p,q) \) where \( d(p,q) \) is the Euclidean distance between voxels \( p \) and \( q \). This energy formulation imposes smoothness on \( \omega \), since a larger off-resonance difference of a neighbor pair results in a greater energy \( E \). The energy formulation was later improved [115] by including \( \min(J''(\omega^{\min}_p), J''(\omega^{\min}_q)) \) in \( w_{p,q} \), where \( J''(\omega^{\min}) \) is the second derivative with respect to \( \omega \) at the smallest minimum \( \omega^{\min} \). Using such weights, the location of energy minima becomes independent of the arbitrary scaling of the signal \( y \).

The connection to Markov Random Fields is an advantage, as there is a rich literature on the subject, not least from the image analysis field of research. However, the minimization of energy functions like eq. 4.28 is non-trivial, and there is no guarantee to find the global optimum [144]. Hernando et al. [130] proposed to use the iterated conditional modes (ICM) technique [145]. An algorithm performing several graph cut moves [146] was later proposed by the same group [115]. A solution using belief propagation [147, 148] was described by Lu and Lu [149]. Other possible techniques for minimizing energies like eq. 4.28 include tree-reweighted message-passing (TRW-S) [150, 151], and recent graph cut techniques such as quadratic pseudo-boolean optimization (QPBO) [152] and fusion moves [153].
In paper II, the problem of choosing between two candidates of $e^{i\omega \Delta t}$ in each voxel was formulated in terms of the energy given by eq. 4.28 with $D_q = 0$, $V(\omega_p, \omega_q) = |e^{i\omega_p \Delta t} - e^{i\omega_q \Delta t}|^2$, and $w_{p,q} = m_{p,q}/d^2(p,q)$ where $m_{p,q}$ is the smallest of the two voxel magnitudes. The solution $\omega$ was restricted in each voxel to its two $e^{i\omega \Delta t}$ candidates. The problem was then solved using TRW-S. This formulation naturally takes into account the periodic nature of the two-point sampling scheme, and avoids the more complicated determination of the unwrapped off-resonance map.

In paper III, the energy proposed by Hernando et al. [115] was modified by using:

$$V(\omega_p, \omega_q) = \min \left( |\omega_p - \omega_q|^2, (\Omega - |\omega_p - \omega_q|)^2 \right)$$

(4.29)

With this modified discontinuity cost, $\omega$ can be restricted to a single period $\Omega$. This reduces the solution space without affecting the estimation of the other parameters. However, the graph-cut algorithm cannot be used since $V(\omega_p, \omega_q)$ is now non-submodular [144]. The problem was solved in two steps: first, the solution space was constrained to the two smallest minima in an $\Omega$ period in each voxel. This constrained problem was solved using QPBO [152]. The result given by QPBO was then used as initial guess when solving the unconstrained problem using the locally convergent ICM [145].

In paper IV, the same method was used for resolving the off-resonance map as in paper III.

4.3.5 Other methods

Several additional methods have been proposed to solve the off-resonance problem under the assumption of spatial smoothness. Similar to energy minimization, the off-resonance problem has been described in terms of solving a Poisson equation [154, 155]. Techniques performing iterative low-pass filtering and updating of the off-resonance candidates have also been used [101, 128, 131, 138]. Other approaches include region merging [133] and the use of normalized convolution [155].
4.4 Acquisition strategies

An important consideration for the acquisition scheme is how to choose the time shifts $t_n$, which has been discussed above. Other considerations in the acquisition strategy require special attention, not least since they have impact on the image weighting and acquisition time. In Dixon’s original paper [93], two spin echo images were acquired sequentially, according to fig. 4.3. This doubles the acquisition time compared to conventional imaging. A similar approach using gradient echo imaging (fig. 4.2) was described by Park et al. [156]. The duration between the acquisitions of the images makes this approach vulnerable to motion artifacts and image misregistration [157].

An alternative approach was described by Wang et al. [158], who acquired three gradient recalled images in an interleaved fashion to reduce such problems. The acquisition time can be reduced by acquiring all the time shifts within the same repetition [159]. A so-called flyback gradient with opposite polarity to the readout gradient can be used to refocus the signal between readouts. This approach has been described for spin echo [159] and gradient echo acquisitions [119]. Alternatively, the flyback gradient can be replaced by a readout gradient of opposite polarity, giving a so-called bipolar acquisition [160, 161, 162, 163]. Fig. 4.13 shows a gradient echo acquisition, where all echoes are sampled in the same repetition using alternating readout polarity. This is a very efficient data sampling strategy. Although, inconsistent phase errors between the polarities are caused by eddy currents and system imperfections [163]. A linear component along the fast encoding direction of such phase errors may be corrected for, by identifying a corresponding peak shift in k-space [162]. Alternatively, additional reference data may be acquired for correction along the fast encoding direction [161] or correction along both fast and slow encoding directions [163]. Acquiring all time shifts in the same

![Figure 4.13](image-url)  
*Figure 4.13:* Timing diagram for a gradient recalled multi-echo acquisition with bipolar readout gradients. The amplitude of the received signal diminishes between echoes, due to $T_2^*$ relaxation.
repetition is time-efficient, but interleaved acquisition has the advantage of complete freedom in the choice of $t_n$.

In the studies described in this thesis, multi-echo data of the same polarity was acquired using gradient recalled sequences.

FWI has been used with fast spin echo acquisitions [103, 118, 160, 164], GRASE [161], spin echo and stimulated echo acquisitions [165], SSFP [131, 166] and balanced SSFP [167, 168]. Spin echo acquisitions with long $TE$ enable $T_2$-weighted FWI [169].

It has been described how to perform the separation of water and fat in k-space after estimating and demodulating the off-resonance map in image space [162, 170]. With this approach, the duration of the signal readout can be taken into account, removing artifacts of chemical shift displacement [92]. Such effects can be substantial when using non-cartesian trajectories [132, 170].

Other work explores the combination of FWI with ultra-short echo time imaging [171] and compressed sensing [136, 172].

### 4.5 Quantitative water/fat separation

Given the complex water and fat images, the fat signal fraction may be calculated as [173]:

$$\text{FF} = \frac{|F|}{|F| + |W|}$$

(4.30)

Neglecting relaxation effects, this corresponds to the relative fraction of fat protons [173]. As opposed to $W$ and $F$, the intensity scale of $\text{FF}$ is not arbitrary. An $\text{FF}$-image thus serves as a quantitative mapping of the fat signal fraction in each voxel [174]. A fat fraction map is shown in fig. 4.14.

The notion of qualitative and quantitative FWI identifies a major criterion for applications. While the priority of quantitative FWI is to obtain as accurate fat fraction estimates as possible, the qualitative approaches strive to achieve sufficiently accurate water/fat separation for visual assessment.

*Figure 4.14:* Images of water (a), fat (b), and the derived fat fraction map (c). Here, the fat fraction is coded in pseudo-color.
Due to the utility of $FF$-maps for non-invasive estimation of the amount of organ fat [175], quantitative FWI has been met with great interest, not least in the assessment of fatty liver (steatosis) [176, 177]. FWI was used to detect liver steatosis already in 1984 [178]. A $FF$ of 5.6% has been proposed as a threshold for the classification of liver steatosis [179].

The non-invasive gold standard method for measuring organ fat is localized $^1$H spectroscopy [179]. This enables convenient comparison since MRS and MRI measure the same entity, and can be performed in the same examination. Excellent correlation between liver fat fraction measured by FWI and MRS has been demonstrated [119, 175, 180, 181, 182, 183, 184, 185, 186]. The liver fat fraction measured by MRS, in turn, has been shown to be highly reproducible [179] and correlates well with chemical analysis [62, 187] and steatosis assessed by histology grading [180, 188, 189] measured from biopsy samples. Measurement of $FF$ by quantitative FWI has been validated in phantom experiments [129, 137, 175, 190, 191] and has been shown to be highly reproducible [186], and to correlate well with chemical analysis [192] and histology grading of steatosis [176, 192]. Taken together, there is a large body of evidence for the clinical utility of quantitative FWI in the diagnosis of liver steatosis.

Several sources of confounding have been identified, which must be addressed to obtain an accurate measurement of $FF$. Since relaxation weighting is included in $W$ and $F$, accurate fat fraction estimation requires either proton density weighted images, or estimation and incorporation of the relaxation factors into the calculation of $FF$ [173].

To avoid effects from $T_2^*$, the most common approach is to use spoiled gradient echo acquisitions and include $T_2^*$ relaxation in the signal model [114, 129]. A short $T_2^*$ may, when not accounted for, result in inaccurate fat fraction estimates [114, 129, 137, 191]. In the liver, $T_2^*$ may be short if the concentration of iron is enhanced, which is common in patients with steatosis [119]. For accurate $T_2^*$ estimation, at least six echoes are typically used [114].

$T_1$ relaxation may also cause inaccurate fat fraction estimates. Bias from $T_1$ can be avoided either by estimating and compensating for $T_1$ [129, 173, 193], or by achieving low $T_1$-weighting using a low flip angle [129, 193].

Another source of error in fat fraction estimation is inaccurate signal modeling using the simple signal model with a single fat resonance [174]. As should be expected, accounting for all the fat resonances (eq. 4.12) gives a more correct signal model and more accurate fat fraction estimates [117, 119, 137, 191].

The fat fraction as calculated by eq. 4.30 will be $0 \leq FF \leq 1$. If the true fat content is zero, any noise in the fat estimate will result in an $FF > 0$. Thus, the
fat fraction will be biased by noise [193]. This can be avoided by calculating the fat fraction as:

\[
FF = \begin{cases} 
\frac{|W|}{|W+F|} & \text{if } |W| < |F| \\
1 - \frac{|W|}{|W+F|} & \text{if } |W| > |F|
\end{cases}
\]  

(4.31)

Yet another source of errors in the fat fraction estimates is eddy currents [194]. In multi-echo gradient recalled sequences, eddy currents manifest as phase errors, particularly in the first echo. This can be avoided by discarding the phase of the first echo and including only the magnitude of the first echo in the signal model [195].

Taking all these confounding factors into account, quantitative FWI enables accurate and precise assessment of liver steatosis [185, 186].

4.6 Performance analysis

An important consideration, is how noise in the acquired data propagates into the estimates. This can be investigated by direct measurement in the reconstructed images, by mathematical analysis of the signal model, or by simulation, employing a signal model and an estimation algorithm. These three methods are complementary, since the direct measurement method makes no assumption about the correctness of the signal model, mathematical analysis reveals properties of the signal model itself, and simulation reveals properties of a particular estimation method, assuming that the signal model is correct.

In FWI, analysis of the noise performance has been used to aid the choice of time shifts, and to evaluate the efficacy of estimation methods.

Glover [108] defined the effective number of signal averages (NSA) as the noise variance in the acquired images over the noise variance in the estimate. For three-point acquisitions, assuming known off-resonance and using the single fat resonance model, the optimal time shift spacing \( \Delta t \) was analytically determined as \( \Delta t = 2T/3 \), where \( T = \pi/\omega_B \), resulting in a maximum NSA of 3.0 [108]. Xiang [128] used the same method to examine two-point acquisitions, where a maximum NSA of 2.0 is achieved for \( \Delta t = T \).

Pineda et al. [196] included the off-resonance estimation into the analysis of three-point acquisitions, and demonstrated that the NSA also depends on the choice of the initial time shift, \( t_1 \). A maximum NSA of 3.0 for this more realistic situation is achieved for \( t_1 = -T/6 + kT \), where \( k \) is an integer. This was verified by Reeder et al. by direct measurement in an oil-water phantom [103]. Pineda et al. also performed simulation experiments to show that the maximum NSA was achieved by least-squares estimation of the off-resonance, water, and fat, using several combinations of time shifts [196]. This is not necessarily the case for certain analytical estimation schemes [197].
Yu et al. analyzed the maximum NSA for multi-echo gradient recalled acquisitions, incorporating $T_2^*$ in the signal model [114]. Due to $T_2^*$ dephasing, it was shown that a shorter initial echo time $t_1$ resulted in higher maximum NSA. Chebrolu et al. [198] analyzed the maximum NSA for three-point acquisitions using the multiple resonance fat model (eq. 4.12) without $T_2^*$ estimation, and found good agreement with direct measurement in an oil-water phantom. Hernando et al. [137] analyzed the bias and NSA for a wide variety of signal models using both mathematical analysis, simulation, and direct measurement in an oil-water phantom. The best performance was found using complex data and a multiple fat resonance signal model with $T_2^*$ estimation (eq. 4.12) or separate $T_2^*$ estimation for the water and fat components [116].

In paper I, the NSA in the three-point method with analytical determination of the off-resonance (eq. 4.25) and least squares estimation of water and fat was examined by simulation. In paper II, the maximum NSA in the two-point method with off-resonance estimation was analyzed as a function of the time shifts. Simulation experiments were also performed to examine the noise efficacy of the proposed estimation method. In paper IV, the noise performance was analyzed mathematically as a function of $\Delta t$. 
5. Contributions

This chapter summarizes papers I-IV.

5.1 Paper I

This work is entitled *Three-point Dixon method enables whole-body water and fat imaging of obese subjects.*

5.1.1 Aim

To describe and evaluate a fast FWI method suitable for 3D whole-body imaging of normal and obese subjects. 3D whole-body water and fat images are mainly of interest in studies of various diseases, in particular diabetes and obesity [199]. However, reconstructing water and fat images from such datasets is challenging. Both due to the large amounts of data and due to the severe field inhomogeneity caused by the large field of view.

5.1.2 Materials & methods

The described reconstruction algorithm requires the data to be obtained using three time shifts with constant time shift spacing. The simple signal model given by eq. 4.4 including off-resonance modeling was used. Two candidates of the off-resonance phase factor \( e^{i\psi} \) were obtained in each voxel through eq. 4.25.

Assuming the true phase between \( W \) and \( F \) to be zero, voxels with sufficient signal contribution from both water and fat can be resolved. Such voxels were identified by thresholding operations and used as seed points in a region growing scheme described in more detail in section 4.3.3. The region growing path is sensitive to the spatial gradient of the off-resonance map and to the signal strength, allowing “safe” regions with high signal strength and homogeneous \( B_0 \) field (small gradient of the off-resonance) to be determined before less reliable regions with field inhomogeneity or low signal. After determining and removing the off-resonance-induced phase errors, the least squares estimates for water and fat are found in each voxel through eq. 4.9.

The reconstruction algorithm was evaluated on 3D whole-body datasets acquired from 39 subjects with a wide range of body mass indexes (BMI: 19.8–
45.4 kg/m²). A 1.5 T clinical scanner was used with a triple-echo gradient recalled sequence. The data was acquired in 18 stacks using a continuously moving table [57]. This, together with the short repetition time (\(TR = 5.9\) msec) allowed whole-body data of resolution \(2.1 \times 2.1 \times 8.0\) mm³ to be acquired in only 5 min 15 sec.

Reconstruction of water and fat images was performed using the proposed reconstruction algorithm and two reference methods [134, 135], also based on region growing. The quality of the resulting water and fat images was subjectively and independently scored in five body regions on a four-grade scale by two radiologists.

The noise performance of the water and fat estimates was analyzed through Monte-Carlo simulation. The proposed analytical solution was compared to iterative optimization.

Figure 5.1: Water (a, b) and fat (c, d) images from one of the study subjects, reconstructed using the proposed multi-seeded region-growing scheme. The 3D whole-body images are shown in one coronal (a, c), and five axial slices (b, d).
5.1.3 Results

The proposed method was found to give water and fat images with superior quality compared to the reference methods, due to its better ability to determine the off-resonance. Uniform water/fat separation with few or without reconstruction artifacts was achieved for all subjects. Water and fat images from one of the study subjects are shown in fig. 5.1.

Both reference methods received lower grades in body regions of more complicated topology, such as the legs (separated by a background region) and the thorax (containing the signal void of the lungs). The proposed method performed well even in these regions, probably owing to the use of multiple rather than single seeds, and the dynamic rather than predetermined region growing path.

The average reconstruction time for the proposed algorithm was 1 min 53 sec. The Monte-Carlo simulation revealed no difference in noise performance between the proposed method and iterative optimization.

5.1.4 Conclusion

The proposed method enables fast 3D whole-body FWI of normal weight and obese subjects with few or no reconstruction errors and reasonable reconstruction times. The noise performance of the analytical solution is equal to that of iterative optimization, at least for the examined sets of echo times. Using constant spacing of the time shifts, estimating the wrapped off-resonance map is sufficient for accurate water and fat separation.
5.2 Paper II

This work is entitled *Two-point Dixon method with flexible echo times*.

5.2.1 Aim

To shorten the data acquisition times for FWI by removing the constraints on the time shifts in two-point FWI. This will allow faster image acquisition, but also more flexibility in terms of image resolution and readout bandwidth.

5.2.2 Materials & methods

By using real-valued water and fat estimates, the two complex samples in each voxel provide sufficient information to solve for water, fat, and the two phase error factors $e^{i\theta}$ and $e^{i\psi}$. The equations are given in section 4.2.2. Each voxel is assigned two candidates for the latter phase factor, due to the inherent ambiguity of the signal model. An extension of the signal model to include broadening of the fat resonance served as a simple way to account for multiple fat resonances.

The off-resonance phase factor ambiguity problem was resolved by whole-image optimization. The energy given by eq. 4.28 was used, setting $D_q = 0$, $V_{p,q} = |e^{i\omega_p\Delta t} - e^{i\omega_q\Delta t}|^2$, and $w_{p,q} = m_{p,q}/d^2(p,q)$. This energy function was formulated at low resolution, and minimized using the sequential tree-reweighted message-passing algorithm, TRW-S [151].

Feasibility of the proposed method was demonstrated in breath-hold imaging of the abdomen of three volunteer subjects. The water images obtained from dual-echo acquisitions were compared to single-echo images with conventional fat saturation (SPIR) and minimum $TR$, but otherwise identical imaging parameters. Data was acquired from both 1.5 T and 3.0 T clinical scanners.

The Cramér-Rao lower bound of the noise in the water and fat estimates was studied analytically and compared to the actual noise in the estimates given by the proposed method, as obtained by simulation experiments. The noise properties were studied as a function of the two echo times.

5.2.3 Results

The proposed method demonstrated high quality water and fat images. Better separation of water and fat signal was achieved using the broad fat resonance spectrum model. The suppression of fat signal in the water images was found superior compared to conventional fat saturation.

The acquisition times were comparable between the dual-echo and single-echo acquisitions. At 3.0 T, acquisition of the single echo with fat saturation was faster, while at 1.5 T, the dual-echo acquisition was the fastest. For the
protocols used in this study, removing the echo time constraints allowed acquisition time reductions of 26% and 43% at 1.5 T and 3.0 T respectively, as compared to the conventional in-phase/opposed-phase acquisition scheme.

The simulation experiment showed that the noise of the water and fat estimates achieved the Cramér-Rao lower bound. The noise analysis serves as a guide in choosing the echo times.

5.2.4 Conclusion

With the proposed technique, the usual constraints on the echo times in the two-point Dixon method can be relaxed, allowing greater flexibility in the acquisition. This flexibility can be used to shorten the acquisition time. The estimation procedure is noise efficient and compatible with more advanced signal models, including multiple fat resonances. The method is feasible in breath-hold abdominal imaging.
5.3 Paper III

This work is entitled *Three-dimensional water/fat separation and $T_2^*$ estimation based on whole-image optimization – application in breathhold liver imaging at 1.5 T*.

5.3.1 Aim

To describe and evaluate a flexible FWI method suitable for both qualitative and quantitative applications, and to develop an efficient algorithm for determining and regularizing the off-resonance map, feasible for 3D datasets.

5.3.2 Materials & methods

The signal model proposed by Yu et al. [117] given in eq. 4.12 was employed, using values of the amplitudes and chemical shifts of the fat resonances given by Hamilton et al. [66]. Simultaneous estimation of $T_2^*$ was included in the model [114], but can be omitted by assuming $R_2^* = 1/T_2^* = 0$. The complex water and fat estimates requires at least three echoes to be acquired. The ambiguous estimation of the off-resonance map was based on the whole-image optimization problem formulated by Hernando et al. [115], including a distance term [130] to account for non-isotropic voxels in 3D datasets. Constant time shift spacing was used to obtain a periodic residual function. This periodicity was taken into account in the problem formulation by using the discontinuity cost given by eq. 4.29. With this problem formulation, estimating a wrapped off-resonance map is sufficient, efficiently reducing the solution space. However, the graph-cut based algorithm [115] cannot be used, since the discontinuity terms are no longer submodular [144].

The residual function was evaluated at discrete points within a period in each voxel. The two local minima with the smallest residual were kept as candidates for the off-resonance map. This constrained binary optimization problem was solved by quadratic pseudo-boolean optimization (QPBO) [152]. In a second step, the unconstrained problem was solved approximately using iterated conditional modes (ICM) [130, 145].

The method was evaluated using a 1.5 T clinical scanner with a six-echo 3D spoiled gradient echo sequence to acquire data from ten volunteer subjects. The acquisition was performed within a single breathhold, and the field-of-view was set to cover most of the liver.

The reconstructed water and fat images were analyzed by two operators to locate any reconstruction errors.
5.3.3 Results

The image reconstruction was successful in all study subjects. Accurate separation of water and fat was found despite wraps in the off-resonance map. Images reconstructed by the proposed method are shown in fig. 5.3. Only minor reconstruction errors were found, primarily lung blood vessels being swapped into the fat image (average volume: $6.7 \pm 4.6 \text{ mL}$).

The image acquisition time was 15.9 sec, and the average reconstruction time was $13.0 \pm 0.9 \text{ sec}$ (datasize: $128 \times 96 \times 15 \text{ voxels}$).

5.3.4 Conclusion

The proposed method enables fast quantitative FWI with accurate separation of water and fat signal and simultaneous $T_2^*$ estimation. The method is feasible for 3D datasets. Estimating a wrapped off-resonance map is sufficient for accurate water/fat separation, and the voxel-level signal ambiguity is well captured by the two smallest local minima within a period of the residual function.

![Figure 5.3](image_url)

*Figure 5.3:* Axial slice of 3D images acquired in a single breathhold, covering the entire liver in one of the study subjects. Fat (a), water (b), and quantitative maps of fat fraction (c) and $R_2^*$ (d) are shown. The fat fraction image reveals a signal fat fraction of 11% in the liver, indicating steatosis.
5.4 Paper IV

This work is entitled *Model-based mapping of fat unsaturation and chain length by chemical shift imaging: phantom validation and in vivo feasibility*.

5.4.1 Aim

To develop and validate a method within the framework of FWI for quantitative mapping of the triglyceride characteristics fatty acid chain length ($CL$), number of double bonds ($ndb$), and number of methylene-interrupted double bonds ($nmidb$), and to demonstrate feasibility of the method *in vivo*.

5.4.2 Materials & methods

The signal model proposed by Yu et al. [117] was extended to include four components of the fat spectrum (eq. 4.15). The four fat components $F_1$–$F_4$ are linear combinations of the fat resonances, chosen to allow degrees of freedom in the model with respect to $CL$, $ndb$, and $nmidb$, in accordance to well-established assignment of the fat resonances [64, 66]. When estimated, the fat components were translated into $CL$, $ndb$, and $nmidb$ according to:

$$ CL = 4 + |(F_2 + 4F_3 + 3F_4)/3F_1| $$

$$ ndb = |(F_3 + F_4)/F_1| $$

$$ nmidb = |F_4/F_1| $$

This enabled spatial mapping of fatty acid chain length and degree of unsaturation from spoiled gradient multi-echo data. Estimation of the off-resonance and $T^*_2$ maps was performed using the method described in paper III.

At least six echoes are needed for the estimation. In this study, 32 echoes were acquired, which was the maximum allowed by the user interface of the scanners. Imaging was performed in a single 2D slice. To achieve a high signal-to-noise ratio, 32 signal averages were acquired.

The method was validated by imaging an oil phantom, containing ten food oils of varying fatty acid composition. The oils were contained in plastic cups placed in water stabilized with agar (see fig. 5.4 a). $CL$, $ndb$, and $nmidb$ were measured within regions-of-interests in the parametric images obtained using the proposed method. The reproducibility of the estimates was measured by calculating the intra class correlation (ICC) from repeated acquisitions. As a reference method, the fatty acid composition of the oils was also measured by gas-liquid chromatography (GLC), and translated into $CL$, $ndb$ and $nmidb$. The accuracy of the method was assessed by statistical tests of correlation and linear regression.
To demonstrate the feasibility of the method in vivo, the thigh of a volunteer subject was also imaged with repeated acquisitions. The triglyceride quantities were measured in three regions-of-interest in the subcutaneous adipose tissue. All imaging was performed using both 1.5 T and 3.0 T clinical scanners.

5.4.3 Results

Images of the oil phantom obtained by the proposed method are shown in fig. 5.4. The triglyceride quantities were relatively stable within each oil in the phantom, as well as within the adipose tissue of the volunteer subject.

The ICC for measuring $CL/ndb/nmidb$ of the oils were $0.964/0.997/0.999$ (1.5 T) and $0.998/0.999/0.999$ (3.0 T), indicating high reproducibility of all three quantities at both fieldstrengths.

The measurements of all three quantities were found to be significantly correlated between MRI and GLC at both fieldstrengths. The $r^2$ values for $CL/ndb/nmidb$ were $0.878/0.972/0.994$ (1.5 T) and $0.889/0.974/0.994$

![Figure 5.4: Resulting images from the oil phantom study. A photograph (a), water (b) and fat (c) images are shown along with quantitative maps of fat fraction (d), $R^*_c$ (e), off-resonance (f), and the triglyceride quantities $CL$ (g), $ndb$ (h), and $nmidb$ (i). The images g–i are masked to aid interpretation. Differences in triglyceride composition of the ten oils are clearly visible in g–i.](image-url)
(3.0 T). The linear regression revealed good accuracy for measuring \( ndb \) and \( nmidb \), while the \( CL \) measurements were over estimated with respect to GLC.

The in vivo measurements from the repeated acquisitions were in good agreement, and the three locations showed similar fatty acid composition, as would be expected. All estimates were slightly higher at 1.5 T than at 3.0 T. The measured values of \( ndb \) and \( nmidb \) were in agreement with literature values, while \( CL \) appeared over estimated.

5.4.4 Conclusion

This study demonstrated feasibility of quantitative mapping of \( CL, ndb, \) and \( nmidb \), both in phantom and in vivo, using standard clinical scanners. The estimates are highly reproducible and correlate with GLC when measured in phantom. The \( ndb \) and \( nmidb \) correlate strongly with GLC. The feasibility of the method in vivo makes it potentially useful in large-scale studies of triglyceride composition and its relation to diet and various diseases.
6. Discussion

Different FWI methods have been available for more than 25 years. The ability to handle off-resonance due to $B_0$ field inhomogeneity makes these methods an attractive alternative to conventional fat suppression methods, being potentially insensitive to off-resonance induced phase errors. Yet, the use of FWI has been relatively limited in clinical practice. This can be explained by several problems associated with FWI. First, the image acquisition time will be prolonged due to the requirement of multiple time shifts. Second, the model-fitting approach requires an additional post-processing step in the image reconstruction. The algorithms for handling off-resonance tend to get quite complicated, and the reconstruction times can be long. Third, the challenging estimation of the off-resonance map makes FWI vulnerable to reconstruction errors, which may cause unacceptable degradation of the image quality.

A renewed interest in FWI methods has been noted recently. Not least, this applies to quantitative FWI for diagnosis of liver steatosis. Widespread clinical use of FWI will require implementation directly on the MRI machines. This development has started recently among the major vendors.

6.1 Previous work

Some of the most important preceding work was described in chapter 4. The development of the three-point Dixon method [104] offered a way to handle the off-resonance problem. Concurrently, the problems of long acquisition times, complicated post-processing algorithms, and long reconstruction times were aggravated. In order to reduce the acquisition times, two-point methods with off-resonance correction are essential. Although such a method was introduced as early as 1987 [124], the next step to further minimize the acquisition time by relaxing the opposed-phase timing constraints was taken almost 20 years later [128].

Numerous methods have been proposed to solve the important off-resonance ambiguity problem. Still, no method seems to be generally accepted as the best one. Many of these methods require a certain choice of the number and timing of the spectral samples. Some methods are not described in sufficient detail to be reproducible, and many methods have not been rigorously evaluated. A collection of benchmark datasets would aid a comparison between the different methods.
6.2 Present work

The author hopes that this thesis contributes to the increased understanding and usage of FWI methods. It was intended to consult previous knowledge in the development of new algorithms. Consequently, off-resonance unwrapping or determination of the absolute off-resonance was avoided, since a wrapped off-resonance map is known to be sufficient for accurate water/fat separation [124]. The imposed requirement of constant time shift spacing is not believed to be a serious limitation in most cases.

Chapter 4 provided a unified framework for signal modeling. It should be noted that the four-component fat spectrum model (eq. 4.15) and the simple model with $T_2^*$ modeling (eq. 4.10) are special cases of the multiple fat resonance spectrum model (eq. 4.12). Further, the simple signal model (eq. 4.4) is a special case of eq. 4.10, and the naive signal model (eq. 4.1) is a special case of eq. 4.4. The use of the multiple fat resonance signal model, with or without $T_2^*$ estimation, is encouraged, as it enables more accurate estimates in quantitative applications and stronger fat suppression in qualitative applications. Since all fat resonances are linear in the model and assumed known a priori, additional signal samples are not required. Variation between tissues and between subjects is not taken into account using this signal model, since a fix fat spectrum model is assumed. It should be emphasized that the simple signal model also assumes a fix fat spectrum model, namely the inaccurate model of a single fat resonance.

All off-resonance correction algorithms were described and proven feasible in 3D. This has the advantage of generality, 2D images forming a special case. In addition, the use of 3D information might be beneficial, since the number of neighbor voxels increases, and for topological reasons.

Three subproblems of FWI were identified as: 1) Estimation of off-resonance candidates; 2) Solving the off-resonance ambiguity problem; 3) Estimation of water and fat signal components. These problems were treated separately, the last one consequently being solved by least squares estimation.

The reconstruction algorithms were developed with fast performance in mind, and reconstruction times were measured in the experiments. Reconstruction speed is regarded an important feature, crucial to the practical utility of the methods. For implementation directly on the MRI machines, reconstruction speed is even more important. It is desirable that the operator is able to view the images directly after acquisition, in order to decide if complementary images are required.

Paper I described a solution to the off-resonance problem for three-point acquisitions, being suitable in 3D whole-body applications. The authors were not aware of any such method being described in the literature, since the challenging datasets imposed special requirements. The method was required to utilize 3D information, and to handle complicated topology. Not least, time
efficiency of the reconstruction was required due to the large datasets. The method was evaluated for a particular acquisition scheme in several subjects, to provide confidence in the reconstruction accuracy. The evaluation turned out in favor for the proposed method, compared to the two reference methods. Although, the reference methods were not developed for whole-body applications and were not described for 3D, so the comparison was not completely fair. However, this highlighted the need for a fast 3D reconstruction algorithm.

A fast 3D FWI reconstruction algorithm was also described in paper III, but for the more general case of three or more time shifts. In principle, the algorithm described in paper I could have been used with minor modifications. However, it was desired to avoid the dependence of thresholds used to identify the seed points in the region growing method. The whole-image energy minimization approach [130] was found to be more attractive from a theoretical point of view. In addition, regularization of the off-resonance map is enabled within this framework. This serves as a means to reduce noise in the water and fat images in a well-controlled manner. The experiments showed that fast 3D reconstruction by whole-image energy minimization was indeed possible. In addition, few reconstruction errors were found in the resulting images, even for a challenging dataset with a very large field of view.

In paper II, a two-point Dixon method with a fully unconstrained choice of the two time shifts was described. This method provides more flexibility in the acquisition scheme compared to previous methods. In particular, shorter acquisition times can be achieved with this method. Thus, it addresses one of the major drawbacks of FWI, allowing water and fat images to be acquired faster than previously possible. The conventional IP/OP sampling scheme will not work well with the methodology described in paper II. However, this is a consequence of the strategy for resolving the off-resonance ambiguity, rather than the estimation of the off-resonance candidates. For the special case of IP/OP sampling, some other method for resolving the off-resonance map can be employed, such as region-growing [141].

After paper II was accepted for publication, a similar method was published by Eggers et al. [200], including equations equivalent to 4.22 and 4.23. However, no analytical expression was given for \( e^{i\theta} \), as in eq. 4.24. Instead, two alternative solutions were proposed. The first solution involves an optimization in each voxel to find \( w \) and \( f \) (although not in the least squares sense). The second solution first removes \( e^{i\psi} \) from \( y_2 \), and then finds the least-squares solution for the complex-valued \( W \) and \( F \). Since there are two complex known and two complex unknown variables, the residual will be zero in this case.

The solution given in paper II is advantageous over the first solution given by Eggers et al., since it does not require optimization and allows least-squares estimation. An additional advantage of calculating \( e^{i\theta} \) analytically, is that smoothing can be performed of both the \( e^{i\psi} \)-map and the \( e^{i\theta} \)-map. This avoids the enhanced interface problem illustrated in fig. 5 of ref. [200]. It also enables
higher NSA in the water and fat images, compared to the complex approach [200].

Given the development of flexible time shifts, the two-point method might turn out to be the method of choice for future qualitative applications.

Paper IV described a fully novel method for quantitative mapping of fatty acid chain length, and degree of unsaturation and polyunsaturation. This is quite different from conventional FWI. However, the method was described within the FWI framework, using a relatively minor modification of the signal model (compare equations 4.15 and 4.12). The methods ability to measure the degree of unsaturation and polyunsaturation was quite accurate with respect to gas-liquid chromatography when measured in an oil phantom. The same accuracy was not seen for the mapping of chain length. Similarly, the degree of unsaturation and polyunsaturation measured in vivo were in reasonable agreement with literature values, as opposed to the chain length. The acquisition time is quite prolonged compared to conventional FWI. Both due to the increased number of required samples, and due to the increased number of signal averages motivated by the relatively weak signal of the smaller fat components. The feasibility of the method on 1.5 T and 3.0 T scanners promises more widespread use of such triglyceride measurements, since corresponding spectroscopy methods [65] are difficult at 1.5 T and 3.0 T.

6.3 Limitations

An evident limitation common to all methods described in this thesis, was the requirement of a constant time shift spacing. This served as a means to impose periodicity of the apparent off-resonance. When acquiring multiple echoes in a single $TR$, which was done in all the described studies, the use of constant echo time spacing is natural. Related to this matter, a wrapped off-resonance map was estimated in all the studies. If an absolute off-resonance map is desired, one can be obtained in a separate unwrapping step.

A source of confounding common to all the described studies is the assumption that all signal emerges from protons in water or fat molecules. MR-visible resonances also originate from proteins, peptides, and small mobile metabolites [61]. However, such species must be present in large amounts to confound the signal from the much more abundant water and fat molecules. MR visible lipids other than fat include cholesterol esters, which can be considered negligible due to the low concentration and broad linewidth [62]. Other sources of model errors are non-Lorentzian lineshapes and susceptibility shifts of fat in different compartments [201].

Eddy currents are a known source of confounding in quantitative FWI that was not addressed. Further, all resonances were assumed to have a common $T_2^*$ value for simplicity. However, water and fat are known to have different $T_2^*$ values [129], as are the different fat resonances [66].
Water and fat signal separation in k-space [170] was not performed. Separation in k-space is known to remove chemical shift displacement artifacts [92]. In all the studies, strong gradient amplitudes were used. Therefore, chemical shift displacement was not considered to be a problem.

A major drawback of FWI that was not addressed in this thesis is the requirement of advanced post-processing algorithms. All the described algorithms are indeed rather advanced. Although, care was taken to describe the methods in great detail, and as clearly as possible.

The evaluation of the methods was also subject to some limitations. In papers I and III, the methods were evaluated in several subjects, but only at 1.5 T, and only using a single acquisition scheme in each of the studies. Strictly, the accuracy of these methods cannot be generalized to other acquisition schemes or other field strengths. In paper III, only qualitative but not quantitative accuracy was evaluated. The flexible two-point method was evaluated at both 1.5 T and 3.0 T using a few different acquisition schemes, but only four datasets were evaluated in total. The method described in paper IV was not validated in vivo.

The measurement of reconstruction time is problematic, since it is heavily dependent on the computer used for reconstruction. In the described studies, the reconstruction algorithms were run on a standard laptop computer. Thus, the measured reconstruction times are indicative of what is to be expected using offline implementation.

6.4 Future work

Further evaluation of the developed methods is of special interest, both general evaluation of method performance in larger study cohorts, and evaluation for specific applications. In particular, the flexible two-point Dixon method needs to be evaluated in more subjects and in different anatomical locations, investigating several sets of time shifts. Further, the method described in paper III needs to be evaluated with respect to quantitative accuracy, for instance using an oil-water phantom [190]. The method introduced in paper IV needs to be validated in vivo, for example by GLC of biopsy samples. It is also desirable to investigate sources of error in the estimation of fatty acid chain length in order to improve the estimation.

Specific applications of special interest for future investigation include dynamic imaging of the liver using the flexible two-point technique, and precise quantitative measurements of the fat signal fraction in the myocardium. Using rapid FWI in MR angiography is another area of interest, where improved fat suppression might enable reduction of contrast agent doses.

There are some applications where modified FWI methods could be developed as an alternative to conventional fat suppression. One such application is diffusion-weighted imaging, where conventional fat suppression might leave
residual fat signal causing artifacts due to the large chemical shift displacement [202]. The same problem is also apparent in fMRI. Another interesting area of FWI method development is absolute temperature mapping based on the temperature dependence of the chemical shift between water and fat [203].

6.5 Implications

This thesis advances the field of model-based water/fat separation by chemical shift imaging. The developed methods enable fast qualitative 3D whole-body fat-water imaging and fast quantitative 3D fat-water imaging, with few reconstruction errors. Faster image acquisition in qualitative fat-water imaging than previously possible can be achieved by the described flexible two-point Dixon method. Finally, non-invasive quantitative mapping of fatty acid chain length and degree of unsaturation and polyunsaturation is rendered possible.

All methods are feasible using standard clinical MRI machines. The developed methods are already being used in several research projects (see Related work) which include studies of diet, obesity and the metabolic syndrome.
Acknowledgements

I am most thankful to all who have supported me in the work with this thesis. I would especially like to express my gratitude to:

**Dr. Joel Kullberg** for being an excellent tutor, and for teaching me everything about research and MRI. No one could ask for a more supportive, knowledgeable and available supervisor.

**Prof. Håkan Ahlström** for providing me the opportunity to join his fine research group. Thank you for sharing your extensive experience, and for confident tutorship.

**Dr. Lars Johansson** for encouraging me to solve problems instead of creating new ones. Thank you for appreciating what is important in science and life, such as great jokes.

I would also like to thank my friends and colleagues at the MRI lab, including Richard Nordenskjöld, Elin Lundström, Anders Lundberg, Gunilla Arvidsson, Francisco Ortiz-Nieto, Dr. Jan Weis (sorry for treating spectroscopy as a subheading in the MRI chapter), Arvid Rudling, Firas Mosavi, Arvid Morell, Johanna Swärd, Goran Abdul-Qadhr, Raquel Themudo, Dr. Morten Bruvold, and Britt-Mari Bolinder. I thank Dr. Tomas Bjerner and Christina Lundberg for their enjoyable travel companionship. I am indebted to Christl Richter-Frohm and Håkan Pettersson for administrative support, and Anna Maria Lundins stipendiefond for generous travel grants. I am grateful to Dr. Brian Welch, Dr. Calum Avison and Marcus Björk for fine collaboration. I thank Dr. Ulf Risérus for interesting discussions and Siv Tengblad for excellent help with GLC analysis, even during her own vacation.

I would like to thank all my friends and family, especially my parents Anita Berglund and Lennart Berglund for teaching me to appreciate the value of knowledge. I owe gratitude to all my teachers during these 20 years of education, and I thank my friends and brothers in Orphei Drängar for singing so well.

Finally, I wish to express my love towards my wife Minna. Thank you for wanting to share your life with me.

Uppsala, September 2011
Bibliography


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 701

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-158111